The Dimensionality of Alcohol Abuse and Dependence: A Multivariate Analysis of DSM-IV Symptom Items in the National Longitudinal Survey of Youth*

THOMAS C. HARFORD, PH.D.,1 AND BENGT O. MUTHÉN, PH.D.1

Department of Social and Behavioral Sciences, School of Public Health, Boston University, Boston, Massachusetts

ABSTRACT. Objective: This article examines the factor structure of 22 symptom items used to configure the criteria of DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition) alcohol abuse and dependence and relates the factor structure to background characteristics. Method: Data for this study were drawn from the National Longitudinal Study of Labor Market Experience in Youth (NLSY). The symptom items were related to the covariates using the statistical technique of structural equation modeling generalized to dichotomous outcomes. The present model is a special case of structural equation modeling, a multiple causes and multiple indicators (MIMIC) model, in which one or more latent variables (i.e., alcohol abuse and dependence) intervene between a set of observed background variables predicting a set of observed response variables (i.e., DSM-IV symptom items). Results: The results of the structural equation analysis provide further support for two dimensions underlying the DSM-IV symptom items. Although the two-factor dimension bore a strong resemblance to the DSM-IV conceptions of abuse and dependence, there were notable differences in the item content of the symptom items for each dimension. The dependence dimension drew upon items related to the abuse criteria for continued drinking despite social problems and recurrent drinking resulting in failure to fulfill role obligations. The abuse dimension drew upon items related to the abuse criterion for hazardous drinking and the dependence criterion for larger amounts over time. The two factors were shown to have different relationships to the background variables. Alcohol dependence was related to family history of alcoholism and educational status. Age was not related to dependence and inversely related to alcohol abuse. Conclusions: Findings from this study replicate the two-dimensional model for DSM-IV criteria found in other studies and provide further support for the validity of alcohol dependence in general population samples. A major implication of the factor structure in the present study relates to the different classification of cases that would otherwise be obtained with DSM-IV criteria. These departures were shown to affect abuse, which retained only 40% of DSM-IV diagnoses, more strongly than dependence, which retained 91% of DSM-IV diagnoses. (J Stud Alcohol 62: 150-157, 2001)

GUIDED BY FORMULATIONS of alcohol abuse and dependence in the DSM-III-R (Diagnostic and Statistical Manual of Mental Disorders, Third Edition Revised: American Psychiatric Association [APA], 1987) and the concept of the Alcohol Dependence Syndrome (Edwards and Gross, 1976), several studies have operationalized these criteria for prevalence estimation in the general population (Grant et al., 1994; Kessler et al., 1994; Regier et al., 1988). Whereas clinical studies have found support for a unidimensional structure of the alcohol dependence syndrome (see Allen et al., 1993), general population studies on the factor structure of alcohol dependence and other alcohol-related problems have not always yielded consistent results. In a secondary analysis of the 1984 national survey on alcohol problems, Hasin et al. (1994) examined six indicators of the alcohol dependence syndrome and five indicators of other alcohol problems. Using a two-factor confirmatory model, they found that the model fitted the data reasonably well, but they noted a very high correlation between the two factors. A one-factor confirmatory solution also provided a reasonably good fit to the data, which led Hasin et al. to conclude that one single factor or dimension underlies alcohol dependence and the many types of alcohol-related problems.

Drawing upon the 1988 National Health Interview Survey, Muthén et al. (1993) examined the dimensionality of both the DSM-III-R and DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; APA, 1994) criteria. Basing their investigation on the work of Edwards and Gross (1976), which posits that increasing severity of alcohol dependence is associated with clustering of the dependence criteria, the authors examined the factor structure of the DSM criteria across subgroups that are defined in terms of progressively greater risk for alcohol dependence—current drinkers, heavy drinkers, and high-risk drinkers based on background characteristics. They reported that a two-factor model significantly improved the model fit for the one-factor solution in all three subgroups and each validation sample. Similar analysis at the symptom level

Received: July 31, 2000. Revision: October 21, 2000.

*This research was supported in part by National Institute on Alcohol Abuse and Alcoholism grant R21-AA10948-01A1 to Bengt Muthén and by grant K02 AA 00230-01 to Bengt Muthén.

†Send correspondence to Thomas C. Harford, Ph.D., 279 North Main Street, Wolfeboro, NH 03894. Bengt O. Muthén is with the Graduate School of Education and Information Studies, University of California, Los Angeles, Los Angeles, CA.
(Mutén, 1995) indicated that a two-factor solution best described the item correlations and bore a strong resemblance to the solution for DSM criteria reported by Mutén et al. (1993). One factor was represented by “persistent efforts to cut down or control drinking,” “give up important activities in favor of drinking” and “continue to drink despite physical or psychological problems,” together with symptom items related to tolerance and withdrawal. A second factor was represented by two criteria: “drinking larger amounts over a longer period of time” and “recurrent drinking in situations where alcohol use is physically hazardous.”

In a community study of heavy drinkers, Haisin et al. (1997b) found that their DSM-IV diagnosis for alcohol dependence was significantly associated with “criterion” variables including alcohol consumption, treatment seeking, family history and suicidal ideation. DSM-IV alcohol abuse, when compared to no diagnosis, was not associated with these criterion variables. Basing their analysis on the 1992 national survey on alcohol use disorders, Hasin and Paykin (1999a) reported that DSM-IV diagnosis of alcohol dependence, when compared with diagnosis of alcohol abuse and no diagnosis, was significantly associated with each of the following external criteria: average daily alcohol consumption, suicidal ideation/attempt, treatment/help seeking, blackouts and family history of alcoholism. Comparisons between a diagnosis of alcohol abuse and no diagnosis yielded significantly greater odds of alcohol abuse for average consumption and blackouts, a marginal relationship for treatment/help seeking, and no significant differences for suicidal ideation and family history of alcoholism. Two prospective studies have indicated that the course of alcohol dependence differs from that of abuse. In a short-term (1 year) prospective follow-up of the community study of heavy drinkers, Haisin et al. (1997c) found chronicity to be the most common outcome for dependence but not abuse. Similar findings were reported in a 4-year follow-up with current drinking men (Hasin et al., 1990).

In summary, both factor analytic and concurrent validity studies provide support for the validity of alcohol dependence in general population samples but raise some questions regarding the concept and measurement of alcohol abuse. In view of variations in the factor structure of DSM symptom items reported in other studies (Hasin et al., 1994) and because the two-factor solutions reported by Mutén et al. (1993) and Mutén (1995) depart somewhat from DSM-IV conceptions of alcohol abuse and dependence, continued exploration of factor structure of symptom items in general population samples is warranted. The present study involves a secondary analysis of a nationally representative sample of young adults at an age when the prevalence of alcohol use disorders is highest. The major objectives of the present study are an examination of the factor structure of 22 symptom items used to configure DSM-IV diagnoses of alcohol dependence and abuse and to assess the relationships between the obtained factors and a set of criterion variables within a structural modeling framework.

Method

Sample

Data for this study were drawn from the National Longitudinal Survey of Labor Market Experience in Youth (NLSY; Ohio State University, 1994). The NLSY, initiated in 1979, comprised a representative sample of men and women aged 14 through 21 on January 1, 1979. Procedures and methods used were designed to produce a database that represented the entire population born in 1957 through 1964, with an oversampling of black, Hispanic and economically disadvantaged nonblack and non-Hispanic youth and a cross-sectional sample designed to represent the military population (Frankel et al., 1983). The military sample (N = 1,280) was dropped in 1985. The sample has been reinterviewed between January and May of each year since 1979, with a retention rate of 90.9% in 1989. The present study draws upon those 1989 respondents who reported the use of alcohol in the last 30 days. Of the total sample of current drinkers (N = 6,648), 5,984 provided complete information on each of the variables described below.

Measures

Alcohol dependence. Twenty-two symptom-item questions designed to operationalize the DSM-III-R (APA, 1987) were applied to DSM-IV (APA, 1994) criteria. Because the symptom items were developed prior to publication of DSM-IV, several criteria are represented by a single item and it was not possible to measure the abuse criterion related to alcohol-related legal problems. The criteria and symptom items are provided in Table 1. The diagnoses derived from these symptom items have demonstrated good test-retest reliability (Grant et al., 1995) and population validity (Harford and Grant, 1994). The symptom items were obtained from respondents who reported the use of alcoholic beverages in the past 30 days. For each symptom, respondents were asked whether this happened three or more times in the past year, two or more times, once, in lifetime other than the past year, or never. Because there were few responses in the first two categories, the items were dichotomized to one or more times in the past year compared to not in the past year.

Family history of alcoholism. In 1988, respondents were asked which of their relatives had been alcoholics or problem drinkers at any time during their lives. Positive responses were recorded for each type of blood relative and classified into four mutually exclusive groups: family history negative (FHN); family history positive involving alcoholism in only second- and/or third-degree blood relatives
(grandparents and other blood relatives) (FH23); family history positive involving alcoholism in only first-degree blood relatives (FH1); family history positive involving alcoholism in both first-degree and second- or third-degree blood relatives (FH123). Three dummy variables were created (FH23, FH1, FH123) with FHN as the referent.

Early onset of drinking. In 1982, respondents were asked, “How old were you when you first started drinking?” (probe: “For example, having two or more drinks a week”). Early onset of drinking was dichotomized as starting to drink at age 15 or younger = 1 and all others = 0.

Alcohol consumption. In 1989 respondents were asked if they had ever drank an alcoholic beverage and, if “yes,” whether they had drunk any alcoholic beverages during the past month. Current drinkers in this study are defined as those using any amount of alcohol in any frequency during the past month. Respondents also reported the number of days on which they drank alcohol and the average number of drinks consumed on drinking days during the past month (the month preceding their interview). A measure of episodic heavy drinking was based on the frequency of occasions during the past month on which six or more drinks were consumed (never, once, 2-3 times, 4-5 times, 6 or more times). Heavy drinkers in this study are defined as those using six or more drinks per occasion one or more times during the past month.

Other variables. The respondent’s age (in years), education, marital status and race were included in the analysis. Two dummy variables were constructed for marital status where 1 = never married and 0 = all others and 1 = separated or divorced and 0 = all others. The referent category for marital status was married. Two dummy variable were constructed for education where 1 = less than a high school education (dropouts) and 0 = all others and 1 = some college and 0 = all others. The referent category for education was high school graduation. Two dummy variables were constructed for minority status where 1 = black and 0 = nonblack and 1 = Hispanic and 0 = non-Hispanic.

Statistical analysis

The DSM-IV symptom items were related to the covariates using the statistical technique of structural equation modeling (see, e.g., Jöreskog and Sörbom, 1979) generalized to dichotomous outcomes (see, e.g., Muthén, 1979, 1989). The present model is a special case of structural equation modeling, a multiple causes and multiple indicators (MIMIC) model, in which one or more latent variables (i.e., alcohol abuse and dependence) intervene between a set of observed background variables, predicting a set of observed response variables (i.e., DSM-IV symptom items). As a preliminary step, the 22 DSM-IV symptom items were analyzed using exploratory factor analysis. The results of
the exploratory factor analysis were used to specify the factor structure in the MIMIC model. The model was estimated with and without the drinking measures and separately for current and heavy drinkers. The analyses were implemented in the computer program Mplus (Muthén and Muthén, 1998). The default estimator for analysis is a robust weighted least squares estimator.

**Results**

The demographic and background characteristics of current and heavier drinkers are presented in Table 2. As shown, approximately 14% and 6% of current drinkers and 23% and 12% of heavy drinkers were classified under DSM-IV alcohol abuse and dependence, respectively. The mean (SD) frequency of heavy episodic drinking was 1.2 (1.6) occasions, the mean number of drinking days (frequency) was 7.3 (7.3) and the mean average number of drinks (average quantity) was 2.9 (2.2). The drinking measures were correlated as follows: episodic-frequency, 0.52; episodic-average quantity, 0.54; frequency-average quantity, 0.17.

**Exploratory factor analysis.** The factor loadings of the exploratory factor analysis of the DSM-IV symptoms items for current and heavy drinkers are presented in Table 3. Among current drinkers, the two-factor solution ($\chi^2 = 841.1$, 126 df) improved the fit over the one-factor solution ($\chi^2 = 1,465$, 127 df), although the two factors are highly correlated ($r = 0.71$). For ease of presentation, the two factors are labeled “alcohol dependence” and “alcohol abuse” as opposed to “DSM-IV alcohol dependence” and “DSM-IV alcohol abuse.” As shown in Table 3, Factor 1 (alcohol dependence) was well measured by items from the following criteria: A1 (social problems), A3 (recurrent), D1 (tolerance), D2 (withdrawal), D4 (CUTDOWN), D5 (GIVENUP), and D7 (continued). Factor 2 as (alcohol abuse) was well measured by items from the criteria A2 (hazard) and D3 (larger) and one item (SICKAFTFR) from the criterion D2 (withdrawal). One item, DIFSTOP, loaded equally on both factors, and the loadings for the three other items, HUNGOVER, DRNKSHAK and SPNTTIME, are not substantially higher on one factor. The elimination of these items in confirmatory factor analyses yielded stable and comparable solutions. As shown in Table 3, the analysis for heavy drinkers replicates the measurement model estimates very closely.

**DSM-IV misclassification.** Because the factor structure in the present study departs somewhat from DSM-IV diagnoses and to provide the reader with an appropriate context for interpretation of the structural equation models to be presented, the symptom items were reconfigured with respect to the present two-dimensional model, with abuse defined by two criteria (A1 and D3) and dependence by the remaining criteria. As expected, the prevalence of dependence in current drinkers changed from DSM-IV 6.4% to 4.3% and the prevalence of abuse changed from DSM-IV 14.1% to 34.3%. Cross-tabulation of these two measures (not shown) indicated that alcohol dependence captured 90.5% of the DSM-IV alcohol dependence cases, with the remainder (9.5%) falling in the DSM-IV alcohol abuse category. Alcohol abuse captured only 39.3% of the DSM-IV alcohol abuse category, with the remainder consisting of 6.4% former DSM-IV alcohol dependence cases and 54.4% cases with no DSM-IV diagnosis. Thus, one could expect the pattern of covariate associations with alcohol dependence, but not necessarily with alcohol abuse, to be consistent with other studies.

**MIMIC analysis.** In the MIMIC model, the exploratory factor analysis of the symptom items was specified in a confirmatory factor analysis framework based on the two-factor solution described above. There were several direct effects associated with gender and race, and their inclusion in the structural model provided important model improvements. Table 4 shows the structural regression estimates for each of the two factors, alcohol dependence and alcohol abuse, for the sample of current drinkers. These estimates are interpreted as partial regression coefficients, controlling for the effects of all variables in the model just as in ordinary, linear multiple regression. With respect to the latent variable for alcohol dependence, there are statistically significant effects for the following demographic characteristics: race/ethnicity, marital status and education. Being black or Hispanic, not currently married and of lower educational status are associated with increased probability of alcohol dependence. Table 4 further shows that family history, early age at onset of alcohol use and each of the three current drinking measures are significantly and positively associated with alcohol dependence. Given that the
variables are standardized, the coefficient 0.17 for the age at onset of drinking, for example, refers to a 17% increase in alcohol dependence when age at onset is 15 years or younger as compared to age 16 years or older. Also, the coefficient 0.39 for family history3 (FH123) refers to a 39% increase in alcohol dependence when family history is present among first- and second- and/or third-degree blood relatives as compared to when there is no type of positive family history. In order to determine whether the inclusion of drinking measures confounded the relationships between demographic characteristics and alcohol dependence and to address the correlations among the drinking measures, the model was rerun four times, removing the three drinking measures and including each drinking measure separately. The only significant effect to emerge related to gender. When the drinking measures were removed, male gender was significantly related to the probability of alcohol dependence. There were slight variations in the size of estimates for the other variables; however, they did not change the inferences made from the model.

With respect to the latent variable for alcohol abuse, there were statistically significant effects for the following demographic characteristics: age, race/ethnicity, marital status and education. Being younger, nonblack and non-Hispanic, and not currently married are associated with increased probability of alcohol abuse. School dropouts, as compared to high school graduates, are less likely to be associated with alcohol abuse. Table 4 further shows that family history, early age at onset of alcohol use and each of the three current drinking measures are significantly and positively associated with alcohol dependence. The inclusion of drinking measures confounded the relationships between gender and educational status and alcohol abuse. When the drinking measures were removed (not shown) male gender was significantly related to the probability of alcohol abuse, and respondents with some college experience, as compared to high school graduates, were less likely to be associated with alcohol abuse.

The structural regression estimates for alcohol dependence and abuse for the subsample of heavy drinkers (not
shown) replicate the findings for current drinkers very closely, especially for family history and the drinking measures. The demographic associations with alcohol dependence are somewhat more attenuated in the sample of heavy drinkers as compared to current drinkers.

**Discussion**

There are several caveats and limitations in the present study which should be considered when assessing the findings presented above. First, this study was not designed to test hypotheses about DSM symptoms and criteria, but, rather, represents a secondary analysis of a large, national longitudinal survey on the labor market experiences of young people aged 14-22 in 1979. Information on drinking experiences, the focus of the present study, was included for select years and not on a systematic basis. Consequently, the examination of criterion variables related to alcohol abuse and dependence is constrained by availability and not by theoretically imposed relevant constructs. Second, the interviewers were not familiar with the criteria or trained to probe for additional information regarding the symptom items. The measurement of withdrawal symptoms, for example, does not provide clear distinctions between clinical withdrawal symptoms versus indicators or short-term physiological consequences of bouts of heavy drinking, and this may explain why this symptom loaded on the alcohol abuse factor rather than on the dependence factor. Third, respondents were not provided with formal criteria of alcoholism or problem drinking for family history and their reports were not validated by other sources. We do not know how the respondents decided that their relatives were alcoholics or problem drinkers. They may have based their judgments on difficulties experienced with relatives rather than on the presence or absence of alcohol use disorders. And, fourth, because there was no provision for temporal clustering of symptoms and criteria, DSM-IV diagnostic estimates were available only for 1989. Some respondents may have met a lifetime diagnosis in prior years with remission in 1989.

In addition to these limitations, comparisons with other studies in the general population are constrained for two reasons. First, the findings are limited to specific age cohorts, aged 26-32 years in 1989. And, second, the methodology used in the present study contrasts with conventional risk factor analysis in which a dichotomous diagnostic variable is regressed on a set of covariates or hypothesized antecedents. Here, two continuous latent variables, alcohol dependence and abuse, that explain the covariation among a set of criteria (i.e., symptom items) are regressed on the covariates.

One finding from this study of young adults, 25 to 32 years of age, replicates the two-dimensional model for DSM-IV criteria and symptom items found in other general population samples (Muthén, 1995; Muthén et al., 1993). Factor 1, alcohol dependence, is well defined by symptom items related to five of the DSM-IV dependence criteria: tolerance, withdrawal, unsuccessful efforts to control drinking, reduced activities in favor of drinking and continuing to drink despite physical and/or psychological problems. Factor 1 in the present study also draws upon two of the DSM-IV abuse criteria: continued to drink despite social problems caused by drinking and failure to fulfill role obligations. Factor 2, alcohol abuse, is defined by two items related to the DSM-IV abuse criterion of recurrent drinking in hazardous situations, two items related to the DSM-IV dependence criterion for drinking larger amounts over a longer period of time and one withdrawal symptom item related to being sick or vomiting after drinking. The abuse factor bears some resemblance to the discrete first illness phase in a study of symptom sequencing in 369 clinical cases by Langenbucher and Chung (1995). In their study, mean age of onset and survival-hazard methods revealed the presence of three discrete stages, with the first illness phase defined by the DSM-IV dependence criterion of "larger/longer" and the abuse criteria of hazardous drinking and recurrent social/legal problems.

The validity of alcohol dependence has been well established through numerous studies including general population samples (Hasin and Paykin, 1999b; Hasin et al., 1997b,c), but a general population psychometric study of alcohol disorders using variables external to the diagnosis for validation produced equivocal results for abuse (Hasin
According to DSM-IV, alcohol abuse is diagnosed only in the absence of dependence and is characterized by use of alcohol in hazardous situations or when associated with tangible consequences (i.e., legal problems, continued use despite social or interpersonal problems, or failure to fulfill major role obligations). Only one of the above four criteria must be met for a diagnosis, and studies have indicated that the majority of respondents meet the diagnosis based on the single criterion of hazardous use, which is nearly always reflected by the symptom item "driving after drinking too much" (Hasin and Paykin, 1999b; Hasin et al., 1997c). Muthén et al. (1993) noted in their two-dimensional model of DSM criteria that the dependence factor was represented by criteria with lower prevalence, which they interpreted as reflecting more severe problems as compared to the higher prevalence criteria related to alcohol abuse. Inspection of the proportions of respondents endorsing the symptom items in the present study (Table 2) indicates relatively higher prevalence of the abuse items compared to the dependence items. An examination of the symptom item content further differentiates the two factors in the present study. The items constituting the second factor, interpreted as abuse, reflect excessive and hazardous drinking experiences and not explicit negative consequences associated with drinking. It is not clear whether abuse in this study reflects a residual category based on symptoms with high prevalence or some abuse-like condition as described by Saunders and Lee (2000). Prospective studies have indicated that respondents classified as DSM-IV alcohol abusers do not necessarily progress to alcohol dependence (Hasin et al., 1990; Hasin and Paykin, 1999b).

The inclusion of background covariates, which were similarly related in the samples of current and heavier drinkers, provides further evidence for validity of alcohol dependence in general population studies. The findings indicated that alcohol dependence was significantly associated with family history, early age at onset of alcohol use and patterns of current alcohol consumption, outcomes consistent with criterion outcomes used in other general population studies (Grant and Dawson, 1997; Hasin et al., 1997b; Hasin and Paykin, 1999a; Muthén, 1995).

With respect to demographic associations, the absence of gender effects for alcohol dependence is consistent with the literature (Grant and Harford, 1990; Harford et al., 1991; Hilton, 1987; Midanik et al., 1996). Although the prevalence of heavy drinking and alcohol problems is higher among men than women, Hilton (1987) has shown that the association between gender and problems disappears once the frequency of heavy drinking is controlled.

The inclusion of background covariates serves to further differentiate the boundaries between dependence and abuse. Alcohol abusers were younger, white, not currently married and less likely to be school dropouts. As with alcohol dependence, however, alcohol abuse was significantly associated with family history, early age at onset of alcohol use and patterns of current alcohol consumption. Interpretations of the relationships between alcohol abuse and covariates and comparisons with other studies are constrained by the fact that the measurement of alcohol abuse in the present study departs from DSM definitions. An analysis of the validity of the DSM-IV alcohol abuse category by Hasin and Paykin (1999a) indicated that, whereas abusers were younger and white, as in the present study, they reported higher frequencies of heavy, episodic drinking in the past year. Family history of alcoholism or alcohol problems, however, was unrelated to alcohol abuse.

Age was significantly and negatively related to alcohol abuse but unrelated to alcohol dependence. The inverse association between age and alcohol abuse is consistent with DSM-IV symptom sequences reported by Langenbucher and Chung (1995) and with surveys in the general population which indicate that heavy drinking and alcohol-related problems peak in the early twenties and then decline. The lack of association between age and alcohol dependence, as measured in the present study, may reflect departure from normative growth trajectories based on symptom severity and patterns of alcohol consumption. The symptom items that defined the abuse factor in the present study draw heavily on the DSM-IV criteria for use in hazardous situations and larger amounts of alcohol than intended, circumstances that reflect single episodes of heavy drinking as opposed to consistent and frequent patterns of heavier use over time.

Acknowledgments

The authors wish to thank the reviewers for their helpful comments on the revision of this article.

References


OHIO STATE UNIVERSITY. National Longitudinal Surveys, Columbus, OH: Center for Human Resource Research, Ohio State University, 1994.
