Should substance use disorders be considered as categorical or dimensional?

Bengt Muthén

University of California, Los Angeles, CA, USA

ABSTRACT

Aims This paper discusses the representation of diagnostic criteria using categorical and dimensional statistical models. Conventional modeling using categorical or continuous latent variables in the form of latent class analysis and factor (IRT) analysis has limitations for the analysis of diagnostic criteria. **Methods** New hybrid models are discussed which provide both categorical and dimensional representations in the same model using mixture models. Conventional and new models are applied and compared using recent data for *Diagnostic and Statistical Manual of Mental Disorders* version IV (DSM-IV) alcohol dependence and abuse criteria from the National Epidemiologic Survey on Alcohol and Related Conditions. Classification results from hybrid models are compared to the DSM-IV approach of using the number of diagnostic criteria fulfilled. **Results** It is found that new hybrid mixture models are more suitable than latent class and factor (IRT) models. **Conclusion** Implications for DSM-V are discussed in terms of reporting results using both categories and dimensions.

Keywords Classification, diagnosis, factor analysis, latent classes, substance abuse, substance dependence.

Correspondence to: Bengt Muthén, University of California, Los Angeles, CA, USA. E-mail: bmuthen@ucla.edu

INTRODUCTION

This paper discusses the representation of diagnostic criteria using categorical and dimensional modeling. The choice between categorical and dimensional views of disorders has created a long-standing debate in psychiatry. In the context of traditional *Diagnostic and Statistical Manual of Mental Disorders* (DSM) diagnosis the categorical view dominates, because it meets clinical needs and the needs of reporting for health-care planners and insurance companies. Recent interest, however, focuses on the possibility of dimensional approaches where a quantitative score, or scores, can be used for research purposes. This raises questions about which approach is most suitable for a particular domain of disorders and for which particular purpose, as well as if and how one can translate between categorical and continuous representations.

To be able to answer the question posed in the title of this paper, it is important to bring together critical thinking both in areas of psychiatric measurement and statistical analysis. This paper aims to contribute to statistical analysis, presenting the research frontier in terms of psychometric modeling. To give subject-matter experts a chance to understand the current analytical possibilities, it is necessary to give an overview of relevant methods, including particularly promising novel approaches that combine categorical and dimensional representations.

The current psychiatric debate about categorical and dimensional views has a counterpart in psychometrics and statistics in general, where the corresponding choice is between using categorical and continuous latent variables. Categorical latent variables (also called latent class variables and finite mixture components) are used to find homogeneous groups of individuals using latent class analysis or, with longitudinal data, to describe across-time changes in group membership using latent transition analysis. Continuous latent variables (also called traits, factors and random effects) are used to study underlying dimensions by explaining correlations among outcomes in item response theory and factor analysis or, with longitudinal data, to describe individual differences in development in growth modeling (also called repeated measures analysis, multi-level analysis).

Conventional modeling using categorical or continuous latent variables has limitations for the analysis of diagnostic criteria and symptom items. In latent class analysis, which uses categorical latent variables, the latent classes ignore possible within-class heterogeneity such as individual differences in severity, and the categorical nature of the latent variable causes relatively low power for genetic analysis such as linkage analysis. In factor analysis, which uses continuous latent variables, there is no model-based classification and it may be difficult to find natural cut points or thresholds for diagnostic purposes. Novel psychometric developments, using hybrids of categorical and continuous latent variable models, aim to circumvent these limitations and provide a useful bridge between the two modeling traditions. Two such hybrids will be discussed here: latent class factor analysis and factor mixture analysis.

The aim of this paper is to present new methodology for studying categories and dimensions rather than trying to reach substantive conclusions. Readers interested in substantive aspects of the debate may consult the large set of papers in psychology and psychiatry, including Meehl [1], Widiger & Clark [2], De Boeck, Wilson & Acton [3] and Markon & Krueger [4]. Early methods for clustering in alcohol studies (e.g. [5,6]) are also not covered in this paper. These methods have shortcomings [7] and are inferior to the statistically more rigorous latent class analysis approach (see [8] and references therein).

This paper begins with a brief, non-technical overview of the two conventional models of latent class analysis and factor analysis from the perspective of analyzing diagnostic criteria and symptom items. In the context of factor analysis, a brief description is also given of a reporting system used for educational achievement testing, where issues of categories and dimensions similar to those in psychiatry have been discussed. The section 'Hybrid latent variable analysis applied to diagnostic criteria' introduces the hybrid models of latent class factor analysis and factor mixture analysis. The following sections provide some general considerations for the analysis of diagnostic criteria and apply the various models to recent data on DSM-IV alcohol dependence and abuse criteria in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). The final section concludes with a summary of the assets and liabilities of the different analytical approaches.

CONVENTIONAL LATENT VARIABLE ANALYSIS APPLIED TO DIAGNOSTIC CRITERIA

This section gives a brief overview of latent class analysis (LCA) and factor analysis (FA). LCA uses categorical latent variables and FA uses continuous latent variables. The presentation is non-technical, using model diagrams and examples. References to literature with both technical and application focus are provided for further studies.

Categorical representation: latent class analysis (LCA)

Figure 1 describes LCA. Figure 1a considers analysis results in terms of profiles for the four items listed along the x-axis. Here, the example of dichotomous diagnostic criteria for attention deficit hyperactive disorder (ADHD) is used with the first two items representing different aspects of inattentiveness and the next two items representing different aspects of hyperactivity. The picture shows four hypothetical classes of individuals who are homogeneous within classes and different across classes. The class membership is not known, but latent (unobserved) and to be inferred from data using the LCA model. In this sense, LCA has the same aim as cluster analysis. Class 1 consists of individuals who have a high probability of endorsing both types of items ('combined class'), class 2 consists of individuals who show low inattentiveness and high hyperactivity probability ('hyperactive-only class'), class 3 consists of individuals who show high inattentiveness and low hyperactivity probability ('inattentiveness-only class') and class 4 consists of individuals

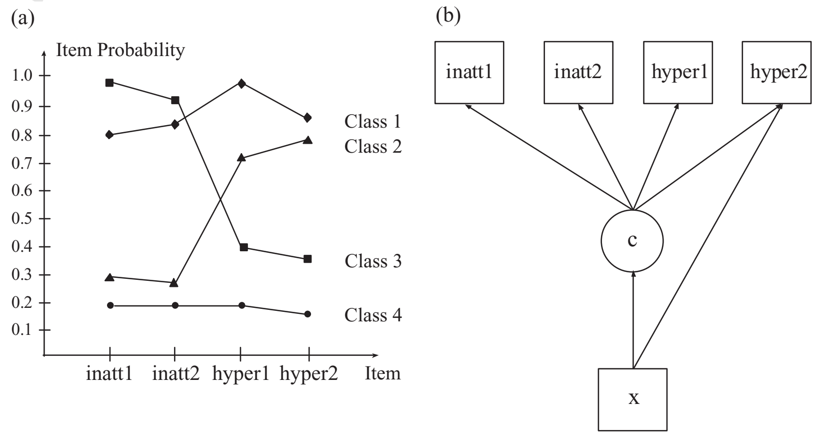

Figure 1 Latent class analysis: (a) item profiles; (b) model diagram

who have low probabilities for all types of items ('unaffected class'). It is seen that the item profiles are distinct and even show two classes with crossing profiles. In a general population sample, the prevalence is the largest for the normative class 4, whereas it is found typically that the hyperactive-only class is the least prevalent in that hyperactivity is observed most often in conjunction with inattentiveness. The class probability may be regressed on background variables (covariates) such as family history of ADHD to estimate how elevated the prevalence is for each of the affected classes 1–3 for individuals with a positive family history as compared to having no such family history.

Figure 1b shows a corresponding model diagram. The boxes at the top represent the four observed items, the circle in the middle represents the categorical latent variable c with four classes, and the box at the bottom represents a covariate x, such as family history. LCA with covariates has four key sets of parameters: (1) the influence of c on each of the items (as shown in the left-hand picture); (2) the prevalence for the classes of c; (3) the influence of x on c; and (4) the direct influence of x on an item. The fourth type of parameter is useful to study measurement non-invariance. As an example, consider a covariate such as gender or age. It is often the case that males and females and old and young differ in their responses on certain items, even when they belong to the same latent class. For example, in the class of combined inattentiveness and hyperactivity, expressions of hyperactivity are more common among younger individuals. A proper model needs to allow for such partial measurement non-invariance. When the covariate has a genetic content, such item non-invariance may be of particular interest in that certain criteria may show especially strong heritability. A fifth type of parameter is also possible, allowing for correlations between items within class, e.g. due to similar question wording. Such relaxation of the independence of the items within class can affect the class formation. Given an estimated model, each individual's probability of class membership can be estimated and the person may be classified into his/her most likely class.

For an overview of LCA methods and applications see, for example, Hagenaars & McCutcheon [9]. In terms of statistical specifications for LCA, both the influence of c on an item and the influence of x on c are modeled using logistic regression and can therefore be expressed in common terms of odds, odds ratios, probabilities and logits. The decision on the number of classes to be used in the analysis is perhaps the most difficult part of LCA, but a combination of statistical and substantive consideration is usually satisfactory. Muthén [10] put LCA into a broader latent variable modeling framework. Muthén & Muthén [11] discussed several applications including LCA of antisocial behavior items in the National Longitudinal Survey of Youth (NLSY), a survey of individuals in early adulthood, where in addition to a normative class they found three classes of individuals with clearly different profiles of antisocial acts: property offense, person offense and drug offense. Rasmussen *et al.* [12] applied LCA to DSM-IV ADHD symptoms in Australian twin data and found an eight-class solution where only some classes were congruent with DSM-IV subtypes. While these studies did not show parallel profiles for all classes, the parallel profiles outcome is often seen in LCA with alcohol use disorder criteria; see, for example, Bucholz *et al.* [13] for Collaborative Study on the Genetics of Alcoholism (COGA) data and Muthén [10] for NLSY data, but has also been found in other cases such as with schizophrenia [14].

Dimensional representation: factor analysis (FA)

Consider a different version of Fig. 1a where the profiles are parallel. Parallel profiles obtained by LCA may be seen as an indication that the construct under study is unidimensional. This view would suggest a factor analysis (latent trait) representation instead of LCA. Factor analysis is discussed next.

Figure 2 describes FA. This analysis is often referred to as latent trait analysis, or item response theory (IRT) modeling, particularly when a single factor is used. For this situation, Fig. 2a shows how the probability of endorsing an item increases as a function of the factor f. Different items have different functions, represented by logistic regressions with different intercepts and slopes. Below the f axis is shown the distribution of the factor, assumed typically to follow a normal distribution.

Figure 2b shows the corresponding model diagram. The factor f is assumed to describe all the correlations among the items. The model has a set of four key types of parameters similar to those of LCA: (1) the two measurement parameters for the influence of f on each item (logit intercept and slope); (2) the mean and variance of the factor distribution (typically standardized to 0, 1); (3) the influence of the covariate x on f; and (4) the direct influence of x on an item. The interpretations of the parameters are similar to those of LCA although for the influence of x on f a regular linear regression specification is used, not a logistic regression, because the dependent variable (f) is continuous. A fifth type of parameter is also possible, allowing for correlations between items within class. Given an estimated factor model, each individual's factor score can be estimated. The estimated precision of this estimate, referred to typically as information curves, can also be assessed.

Analysis and reporting of national general population surveys is one important area of interest for DSM-V considerations. In this context it is interesting to note that FA is used routinely for reporting on national trends in edu-

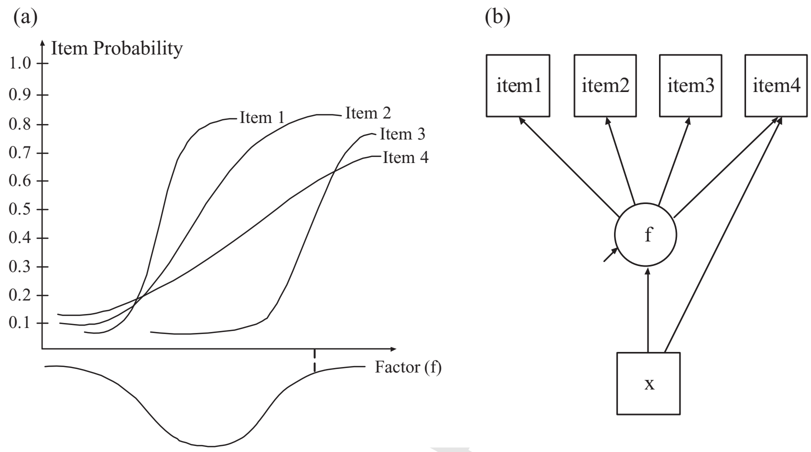

Figure 2 One-dimensional factor analysis: (a) item response curves; (b) model diagram

cational achievement in the survey National Assessment of Educational Progress (NAEP; [15]). The basis for the reporting is a dimensional model such as the one shown in Fig. 2b, where the items in a particular domain such as mathematics are assumed to follow a unidimensional factor model. Different sets of students are given different test forms randomly in order to cover more content domains, which implies that for a given content domain any one student responds to a limited set of items. Because the limited set of items does not produce sufficiently precise factor score estimates, it is necessary to bring in more information in the form of a large set of covariates. Although Fig. 2b shows only one covariate, NAEP achievement analysis uses over 100 covariates including detailed demographic information. The dissemination of information to the public as seen in newspaper reports, however, is not in terms of scores on the factor, but in terms of regions of proficiency that are easier to understand: basic, proficient, and advanced. In this way, a categorization is made of the dimensional factor. The regions are related to the percentiles of the estimated factor distribution with current choices levels being approximately the 30th, 80th and 95th percentiles (Mislevy, personal communication). The factor percentiles are anchored to performance on items discriminating well at the percentile. The choice of relevant percentiles is made in special standard setting sessions with panels of judges basing their judgement on what might be expected of students at a given grade level and subject domain. In sum, NAEP reporting has a dimensional foundation augmented by substantively based categories. This is in contrast with analyses providing model-based categories to be discussed later.

It is interesting to consider a procedure similar to that of NAEP to be used for analysis and reporting of national trends with respect to substance use disorders. If support for dimensional modeling of substance use disorder criteria is found, it might be possible to track national trends using categories such as unaffected, abuse and dependence, where those category boundaries are anchored in FA scores.

For an overview of methods for FA in the form of unidimensional traits see, for example, the item response theory text of Hambleton & Swaminathan [16]. Muthén [17] discusses general multi-factorial FA, including the use of covariates. FA in the form of both unidimensional and multi-dimensional models has been suggested in mental health applications at many points in time: neuroticism in Duncan-Jones, Grayson & Moran [18]; depression in Muthén [17,19] and Gallo, Anthony & Muthén [20]; and alcohol in Muthén [21,22], Muthén, Grant & Hasin [23], Harford & Muthén [24] and Krueger *et al.* [25]. The experience with latent trait modeling in education has been very positive, but it remains to be seen if this methodology is the most suitable or the only one needed for mental health applications.

HYBRID LATENT VARIABLE ANALYSIS APPLIED TO DIAGNOSTIC CRITERIA

Recent methodological developments have made efforts to use a combination of categorical and continuous latent variables to understand more clearly various substantive phenomena. Two key models are latent class factor analysis and factor mixture modeling. Following is a brief description of these analyses and how they relate to the conventional techniques.

Latent class factor analysis (LCFA)

With parallel item profiles, the notion of a dimension influencing the item responses can be formalized into a latent class factor analysis model. This modeling is described in pictorial form in Fig. 3. Figure 3a shows a distribution for a factor (latent trait) *f* and Fig. 3b shows a model diagram. The distribution of the factor is shown as a histogram in Fig. 3a, indicating a strongly non-normal

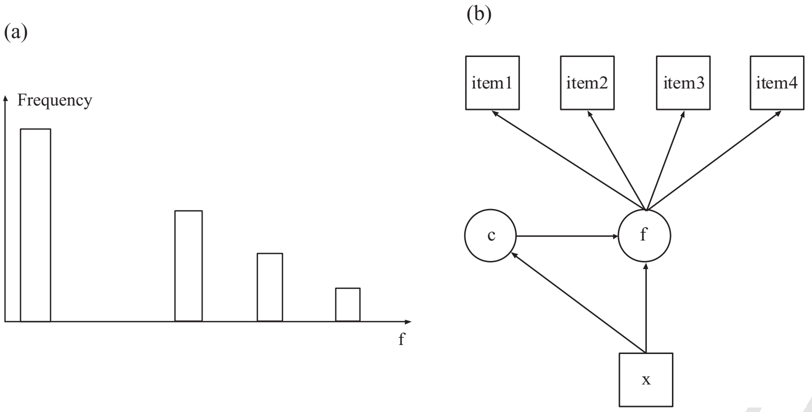

Figure 3 Latent class factor analysis: (a) factor distribution; (b) model diagram

distribution where most individuals are at the unaffected point. The discrete distribution makes for a very flexible description of the factor distribution and is referred to as a non-parametric representation, in that it does not assume a specific statistical distribution such as the normal. Although the points of the distribution are occupied by individuals in different latent classes, it is up to the analysis interpretations in light of auxiliary variables (correlates) and substantive theory to decide if these classes can be seen as substantively different categories or simply representing a single, non-normal distribution.

LCFA has five key types of parameters: (1) the influence of *f* on the items is represented by logistic regressions as in the FA model, so that each item has an intercept and a slope; in line with FA, these measurement parameters do not change across the classes; (2) the influence of *c* on *f* is analogous to regression with dummy variables so that the mean of *f* changes across the classes of *c*, giving rise to the distances between the histogram bars seen in Fig. 3a; (3) the class probabilities give the height of the histogram bars in Fig. 3a; (4) the influence of the covariate *x* on *c* indicates how the class probabilities change as a function of *x*, i.e. how the distribution of *f* is changed by *x*; and (5) the influence of *x* on *f* indicates that *f* may have within-class variation as a function of *x*; this within-class influence can be allowed to vary across class. In line with LCA and FA, LCFA can also have direct influence from *x* to items and items can have residual correlations. Given an estimated model, two types of individual estimates are obtained. First, probabilities for membership in each class are provided. Secondly, factor score estimates are obtained, both for the most likely class and mixed over all classes.

LCFA combines the strengths of both LCA and FA, providing a categorical *and* dimensional representation. Unlike LCA, LCFA provides a factor-analytical interval-scaled dimension with quantitative scores on the factor f. The LCFA model is also considerably more parsimonious than LCA. Using the example of 11 dependence and abuse criteria, four classes, and no *x* variables, LCA uses 47 parameters (corresponding to 11× four item probabilities and three class probabilities) while LCFA uses only 27 parameters (corresponding to 11× two item intercepts and slopes, four factor means of which two are fixed to set the metric, and three class probabilities). The relative parsimony of LCFA can make it more powerful in detecting the influence of covariates.

Factor mixture analysis (FMA)

A second hybrid model, factor mixture analysis (FMA), can be seen as a generalization of LCA, FA, and LCFA. FMA will be discussed only briefly here for lack of space, but may be suitable for applications where there are reasons to believe that there is within-class variation in the item probabilities across individuals due to a common source of influence within class, e.g. representing degree of severity of alcohol dependence. This causes within-class correlation among the items because they are all influenced by this common factor. FMA can be specified to have measurement invariance or not across the latent classes for the logistic regression intercepts and slopes. With measurement invariance the latent classes share the same dimensions, whereas without measurement invariance the dimensions are not comparable across classes. From an LCA perspective, FMA without measurement invariance is a more general clustering technique because it relaxes the LCA specification of zero within-class correlation (no severity variation). From an FA perspective, FMA adds latent classes corresponding to groups of individuals who behave differently. Measurement invariance may or may not be a suitable specification. With measurement invariance, FMA is a generalization of LCFA by allowing for within-class variation around the factor means represented by the *x*-axis values of the histogram bars in Fig. 3a.

LCFA and FMA draw on statistical methods described in Asparouhov & Muthén [26]. For applications to diagnostic criteria for alcohol and tobacco disorders, see Muthén & Asparohov [27] and Muthén, Asparohov & Rebollo [28]. For related modeling without covariates, see Wilson [29], Heinen [30], Vermunt [31] and Formann & Kohlman [32]; with mental health applications in De Boeck, Wilson & Acton [3] and Krueger *et al*. [33]. Even without covariates, LCFA and FMA do not seem to have been used widely and seem very worthwhile to explore further in mental health contexts.

GENERAL ANALYSIS CONSIDERATIONS

Although the discussion in this paper centers on dichotomous outcomes, it should be noted that the outcomes could be of any type: dichotomous (binary), ordinal (ordered polytomous), nominal (unordered polytomous), continuous, limited-dependent (censored-normal), counts, etc. and any combination of such outcomes. This holds true for both categorical and continuous latent variable models. In other words, the type of observed outcome does not necessarily affect the choice between categorical and continuous latent variables. The variety of observed outcome types that can be analyzed together makes it possible, for example, to combine information on dichotomous diagnostic criteria with different information such as quantitative biological measures. As one example, the Windle & Scheidt [34] analysis of alcoholic subtypes could be carried out fruitfully by LCA.

Another consideration related to variables is exemplified by the choice between analyzing symptom items and aggregating their information into diagnostic criteria. An even higher level of aggregation is considered when analyzing diagnoses of dependence for several domains such as alcohol, tobacco, marijuana and depression. Such different levels of aggregation may uncover different features related to categories and dimensions and the differences need to be understood.

In studying mental health phenomena, especially in general population samples, it is typically the case that a large proportion of the sample exhibits none of the symptoms. Proper modeling should include specifications that reflect this. This is possible using an added latent class, a 'zero class'.

Many of the models discussed here cannot be chosen between based on only statistical criteria. For example, it is well known that LCA and FA models often fit the data similarly [35]. Subject-matter considerations play an important role in choosing among models used for different purposes, including considering auxiliary variables in the form of antecedents, concurrent events and distal events (predictive validity; [6]). Typically, with these models maximum-likelihood estimation is used, where the log likelihood (logL) can be seen as an overall assessment of the fit between the model and the data when comparing models. LogL can, however, be made larger simply by adding more parameters to the model and therefore Bayesian information criterion (BIC) and ABIC (sample-size adjusted BIC) statistics are used to combine logL with a penalty for using many parameters. A good model has both a high logL value and low BIC and ABIC values. A likelihood ratio test referred to as LMR [37] provides testing of k–1 versus k classes, and bootstrapped likelihood ratio tests are also possible [38]. In models with categorical latent variables, the entropy (with a 0–1 range, 1 being optimal) gives a measure of how well the latent classes can be distinguished. This is based on individual posterior class probabilities, which can be used for classification into most likely class. The Mplus program [39] provides a very general latent variable modeling framework for maximum-likelihood estimation where the models discussed are special cases. Some of the new models draw on techniques in Asparouhov & Muthén [26].

This method overview, by necessity, omits a host of related new and old developments, and the longitudinal data models of latent transition analysis and growth mixture modeling. An overview of these techniques is given in Muthén [36]. It also omits the work by Meehl & colleagues [40,41] on techniques for distinguishing between categories and dimensions. The taxometric approach of Meehl involves graphical displays, resulting in a useful descriptive and exploratory device. The approach is, however, limited because in line with LCA, it assumes that there is conditional independence among the items within each class, and furthermore it is applicable only to situations with two latent classes [42].

APPLICATION TO NESARC ALCOHOL DEPENDENCE AND ABUSE

This section illustrates the different modeling techniques presented above using data on alcohol dependence and abuse from the NESARC [43]. NESARC is a nationally representative face-to-face survey of 43 093 respondents carried out in 2001–02. NESARC uses a complex survey design with stratification, 435 primary sampling units and oversampling of black and Hispanic households. Within each household, one person was selected randomly for interview, with young adults [18–24] oversampled at the rate of 2.25. The analyses to be presented concern a subsample of 13 067 male current drinkers (respondents who reported drinking five or more drinks on a single occasion one or more times in the past year). The analyses focus on the seven alcohol dependence criteria and the four alcohol abuse criteria, which were derived from a set of 32 past-year symptom item questions designed to operationalize DSM-IV.

The analysis steps will correspond to the order in which the methods were presented: LCA, FA, LCFA and FMA. All analyses were carried out using the Mplus program [39]. The estimation takes into account the NESARC complex survey features of stratification, clustering and sampling weights [44]. Mplus set-ups are available on request from the author.

Results for LCA

As a first step, the 11 alcohol criteria in NESARC were explored in the male current drinker sample using LCA with two to five classes. Table 1 shows model fit in terms of the maximum logL, BIC, sample-size adjusted BIC (ABIC) and LMR. The LCA results at the top part of the table suggest that a four-class solution is preferred. The increase in logL levels off when going from four to five classes and BIC is at its optimum at four classes. Although ABIC suggests five classes, LMR points to four classes.

Figure 4 shows the item profiles of the regular four-class LCA model. The *x* axis lists the seven alcohol dependence criteria and the four alcohol abuse criteria, while the *y* axis shows the probability of endorsing an item. It is seen that this is an example of parallel profiles, suggesting an ordering among the classes from low to high. The estimated class percentages are (going from class 1 with the highest endorsement probabilities to class 4 with the lowest endorsement probabilities): 1%, 5%, 17% and 77%.

Table 1 Latent class analysis, factor analysis, latent class factor analysis, and factor mixture analysis model results, NESARC, male current drinkers, $n = 13\,067$.

No. classes (c), no. factors (f)	*logL*	*No. par*	*BIC*	*ABIC*	*LMR*
LCA					
2*c*	–25.887	23	51.993	51.590	0.0000
3*c*	–25.100	35	50.532	50.420	0.0000
4*c*	**–24.989**	**47**	**50.424**	**50.274**	**0.0025**
5*c*	–24.947	59	50.452	50.265	0.1028
FA					
1*f*	**–25.033**	**22**	**50.274**	**50.204**	**–**
2*f*	–24.991	32	50.285	50.183	–
LCFA					
4*c*	**–25.012**	**27**	**50.279**	**50.193**	**0.0000**
5*c*	–25.006	29	50.287	50.195	0.1520
FMA					
2*c*, 1*f*	**–24.961**	**35**	**50.254**	**50.143**	**–**

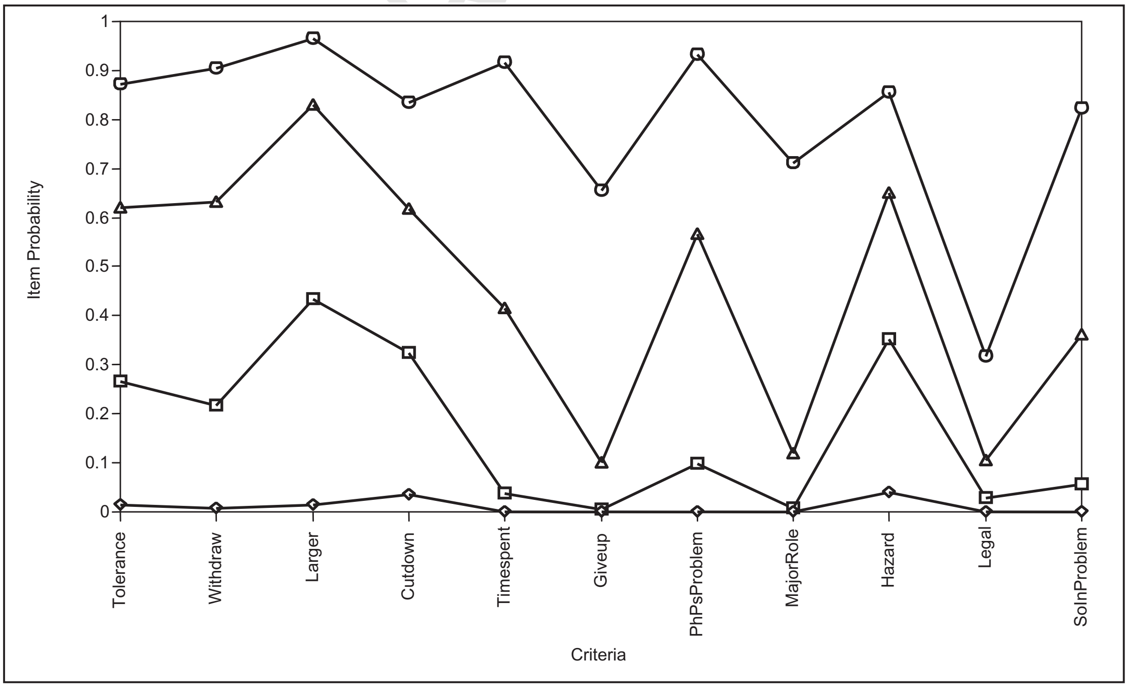

Figure 4 Latent class analysis profiles: ○ Class 1, 1.1%; △ Class 2, 4.8%; □ Class 3, 17.5%; ◇ Class 4 76.6%

The entropy for this model is 0.83, suggesting good classification qualities.

Results for FA

The model fitting results for FA are given in Table 1, both for a single factor and for two factors. The two-factor solution is an exploratory factor analysis solution with minimum restrictions on the factor loadings. The fit statistics of Table 1 indicate that little is gained by adding a second factor. The second factor is measured by the last two abuse criteria, but the two factors are highly correlated (0.95) and it appears that it is not meaningful to consider two separate factors. The item slopes for the factor indicate how well an item discriminates between different levels of the factor. The one-factor model shows similar slopes for most criteria, but has lower slopes for the fourth dependence criterion 'Persistent desire or unsuccessful effort to cut down or control drinking' (cut down) and the second and third abuse criteria 'Recurrent drinking in situations where alcohol use is physically hazardous' (hazard) and 'Recurrent alcohol-related legal problems' (legal).

Results for LCFA

Given the parallel profiles found for the four-class LCA, as well as the unidimensionality of the FA, it is natural to fit a four-class LCFA. This model adds a factor to the regular LCA in line with Fig. 3. The model fit statistics for this model are given in Table 1. Although logL is worse than for the regular four-class LCA, this difference is not large and the parsimony of the LCFA relative to the LCA is reflected by LCFA having considerably better BIC and ABIC values. It is interesting to note that the LCFA model fits better in terms of logL than the one-factor FA, although the difference is not large and BIC and ABIC values are rather close. LCFA does, however, have clear advantages to FA in terms of practical utility as described earlier, in that it provides not only dimensional information but also a classification of individuals. The LCFA slopes in the regression of the items on the dimension have values close to those of the FA.

The LCFA estimated class percentages and entropy remain the same as for LCA. The dimensional aspect of the model is reflected in the estimated class-varying factor means, i.e. the quantitative scores on the single dimension (in the order of class 4, class 3, class 2, class 1): 0, 1, 1.5 and 1.9 (the first two values are fixed to set the metric of the scale). The 11 criteria give rise to 2048 possible outcome patterns of which 50 had a frequency of at least 10 in the analysis sample. The LCFA implies that the large number of response patterns for the 11 criteria has been reduced to only four significantly different types of patterns and these types of patterns can be given these quantitative scores along a single dimension. These scores are well estimated in terms of having small standard errors. Their relative difference indicates that the last two steps are smaller than the first one.

Interpretation of the classes is aided by using the individual estimated class probabilities to classify each individual into his most likely class. For class 1, the response patterns have all dependence criteria met and have most abuse criteria met. Class 2 has mainly one abuse criterion met, 'Recurrent drinking in situations where alcohol use is physically hazardous' (hazard), and this may have to do with the high prevalence of drunken driving. Class 3 is heterogeneous. The unaffected class, class 4, consists of those meeting none of the criteria as well as responses with only one criterion met.

Alternative classifications: LCFA versus number of criteria met

The LCFA classification can be contrasted with the DSM-IV method of diagnosis requiring meeting at least three of the seven dependence criteria and at least one of the four abuse criteria. Basing diagnosis on the number of criteria fulfilled makes several implicit assumptions: (1) the criteria are equivalent (for example, it does not matter which three criteria are fulfilled for a dependence diagnosis); (2) a single dimension (factor) underlies all the criteria; and (3) the same interpretation and metric can be attached to the single dimension in all parts of its range. The LCFA results show that assumption (1) is not met in these data, given different logistic intercepts and slopes for the different items. The other two assumptions are, however, in line with the LCFA model.

Because LCFA specifies a unidimensional model for the 11 criteria, it is of interest to consider a classification based on the sum of all 11 criteria instead of a division into dependence and abuse criteria. Table 2 shows how this alternative classification relates to the LCFA

Table 2 Total number of criteria met versus latent class factor analysis diagnosis.

Total	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Total*
11	11.27	0	0	0	11.27
10	35.93	0	0	0	35.93
9	35.76	0	0	0	35.76
8	40.52	23.07	0	0	63.58
7	8.35	91.14	0	0	99.49
6	0	129.20	0	0	129.20
5	0	197.69	0	0	197.69
4	0	134.22	175.18	0	309.39
3	0	5.32	419.43	0	424.76
2	0	0	856.16	0	856.16
1	0	0	524.40	1 195.78	1 720.18
0	0	0	0	9 183.59	9 183.59
Sum	131.83	580.63	1975.17	10 379.37	13 067.00

classification (frequencies are computed using sampling weights). It is seen that given the LCFA model, the number of criteria met is only a crude approximation. For example, the class 1 diagnosis should be made if at least eight of the 11 criteria are met, but 31 (8 + 23) individuals would be misclassified. The class 2 diagnosis should be made if between five and seven of the 11 criteria are met and the class 3 diagnosis should be made if between two and four of the 11 criteria are met, but both classifications would involve a large degree of misclassification relative to LCFA. The class 4 diagnosis should be made if 0 or 1 criteria are met, but this would include 524 individuals who are in class 3. Although a classification based on number of criteria met is possible and transparent, the classification based on the LCFA model uses more information than merely the sum of criteria and also has a statistical modeling rationale.

Results for FMA

The bottom part of Table 1 shows model fitting results for a two-class FMA model with one factor. This model appears to fit the data better than all the previous ones. The FMA version reported here is the one that focuses on a clustering of subjects, not a representation with measurement invariance and a single dimension for all individuals. The model has class-varying thresholds (intercepts) and factor variances, and class-invariant factor loadings. The non-invariant thresholds imply that the items measure a different construct for the two classes so that, within each class, a separate dimensional representation is obtained. A class with very low probabilities of endorsing items contains 81% of the subjects. This can be compared to the 70% who do not endorse any of the 11 criteria, but the class also contains individuals who endorse one or two criteria. The high 19% class contains individuals who have varying degrees of problematic alcohol involvement. Relative to the low class, the item probability profile for this high class is characterized by especially elevated endorsement probabilities for five of the seven dependence criteria (the first four criteria tolerance, withdrawal, larger, cut down, and the seventh criterion physical and psychological problems), but also for the third abuse criterion (hazard). The factor dimension for this high class may be useful for creating severity scores for this group of individuals.

CONCLUSIONS

This paper describes several powerful latent variable approaches to investigating categories and dimensions of substance abuse and other mental disorders. These should be very useful techniques for investigating psychiatric measurement instruments in the process of formulating the DSM-V. Some techniques have been in use for a long time and have been much explored in mental health settings, such as latent class analysis (LCA) and factor analysis (FA; latent trait analysis) for cross-sectional data and latent transition analysis (LTA) and growth modeling for longitudinal data. Methods that combine categories and dimensions are more recent developments that have seen little application to mental health: latent class factor analysis (LCFA), factor mixture analysis (FMA) and, with longitudinal data, growth mixture analysis (GMA; [11,36,45–47]). LCA and LTA fit well with the need to provide categories of individuals, but cannot supply dimensional assessment. FA supplies dimensional assessment but no categories. In contrast, the newer hybrid models of LCFA, FMA and GMA provide both categories and dimensions. These techniques may be particularly promising for applications to substance use disorders in that such disorders have often been found to have dimensional aspects (see, e.g. [22,25]). As shown by the hybrid models, the fact that dimensions are found does not imply that categories cannot be provided as well. In sum, the answer to the question in the title of the paper is that one does not have to choose categories *or* dimensions, but can consider categories *and* dimensions.

In the NESARC, data on the 11 alcohol dependence and abuse criteria were found to be fit equally well by a four-class, one-dimensional LCFA as by a one-dimensional FA (latent trait model), but the LCFA model provides a richer representation of the data. A similar four-class LCFA was also found for the 32 symptom items underlying the 11 criteria. Furthermore, three-class LCFA models were found to fit NESARC data on marijuana dependence and abuse criteria as well as tobacco dependence criteria.

The NESARC data were used to compare the LCFA classification into dependence and abuse with the number of criteria met. Instead of the DSM-IV requirement of at least three of seven dependence criteria for a dependence diagnosis and at least one of four abuse criteria for an abuse diagnosis, cut-points based on the total number of criteria met were considered. They were found to provide only a crude approximation to the classification based on LCFA.

Hybrid models can be used in analyses with different aims. As opposed to FA, they can be used to produce model-based national prevalence rates in categories such as alcohol dependence and abuse. As opposed to LCA, they can be used for research analyses such as genetic analysis to attain high power due to using a more parsimonious model with a dimensional character; for ideas along these directions, see Muthén, Asparohov & Rebollo [28]. Translations between categories and dimensions are achieved because the categories are formed on the dimensions. Hybrid modeling with longitudinal data appears particularly powerful in uncovering different pathways of problematic development.

Acknowledgements

This research was supported by grant K02 AA 00230 from NIAAA. I thank Tihomir Asparouhov for stimulating modeling discussions, Tom Harford for preparing the NESARC data and for advice on analysis variables, and Karen Nylund, Maija Burnett and Danqing Yu for helpful displays of the results.

References

1. Meehl P. Bootstrap taxometrics: solving the classification problem in psychopathology. *Am Psychol* 1995; **50**: 266–75.
2. Widiger T. A., Clark L. A. Toward DSM-V and the classification of psychopathology. *Psychol Bull* 2000; **126**: 946–63.
3. De Boeck P., Wilson M., Acton G. S. A conceptual and psychometric framework for distinguishing categories and dimensions. *Psychol Rev* 2005; **112**: 129–58.
4. Markon K. E., Krueger E. Categorical and continuous models of liability to externalizing disorders. *Arch Gen Psychiatry* 2005; **62**: 1352–9.
5. Morey L. C., Skinner H. A. Empirically derived classifications of alcohol-related problems. *Recent Dev Alcohol* 1986; **4**: 145–68.
6. Morey L. C., Skinner H. A., Blashfield R. K. A typology of alcohol abusers: correlates and implications. *J Abnorm Psychol* 1984; **93**: 408–17.
7. Blashfield R. K. *The Classification of Psychopathology: neo-Kraepelinian and quantitative approaches*. New York: Plenum; 1984.
8. Vermunt J. K., Magidson J. Latent class cluster analysis. In: Hagenaars J. A., McCutcheon A., editors. *Applied Latent Class Analysis*, pp. 89–106. Cambridge: Cambridge University Press; 2002.
9. Hagenaars J. A., McCutcheon A. *Applied Latent Class Analysis*. Cambridge: Cambridge University Press; 2002.
10. Muthén B. Latent variable mixture modeling. In: Marcoulides G. A., Schumacker R. E., editors. *New Developments and Techniques in Structural Equation Modeling*, pp. 1–33. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.
11. Muthén B., Muthén L. Integrating person-centered and variable-centered analysis: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res* 2000; **24**: 882–91.
12. Rasmussen E. R., Neuman R. J., Heath A. C., Levy F., Hay DA, Todd RD. Replication of the latent class structure of attention-deficit/hyperactivity disorder (ADHD) subtypes in a sample of Australian twins. *J Child Psychol Psychiatr* 2002; **43**: 1018–28.
13. Bucholz K. K., Heath A. C., Reich T., Hesselbrock V. M., Kramer J. R., Nurnberger J. I. *et al.* Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multi-center family study of alcoholism. *Alcohol Clin Exp Res* 1996; **20**: 1462–71.
14. Nestadt G., Hanfelt J., Liang K. Y., Lamacz M., Wolyniec P., Pulver A. E. An evaluation of the structure of schizophrenia spectrum personality disorders. *J Pers Disord* 1994; **8**: 288–98.
15. Beaton A. E., Zwick R. Overview of the National Assessment of Educational Progress. *J Educ Stat* 1992; **17**: 95–109.
16. Hambleton R. K., Swaminathan H. *Item Response Theory: principles and applications.* Boston, MA: Kluwer Nijhoff; 1985.
17. Muthén B. Latent variable modeling in heterogeneous populations. Presidential address to the Psychometric Society. *Psychometrika* 1989; **54**: 557–85.
18. Duncan-Jones P., Grayson DA, Moran PAP. The utility of latent trait models in psychiatric epidemiology. *Psychol Med* 1986; **16**: 391–405.
19. Muthén B. Dichotomous factor analysis of symptom data. In: Eaton W. O., Bohrnstedt G. W., eds. Latent Variable Models for Dichotomous Outcomes: analysis of data from the Epidemiological Catchment Area Program. *Soc Methods Res* 1989;18: 19–65.
20. Gallo J. J., Anthony J. C., Muthén B. Age differences in the symptoms of depression: a latent trait analysis. *J Gerontol* 1994; **49**: 251–64.
21. Muthén B. Latent variable modeling in epidemiology. *Alcohol Health Res World* 1992; **16**: 286–92.
22. Muthén B. Psychometric evaluation of diagnostic criteria: application to a two-dimensional model of alcohol abuse and dependence. *Drug Alcohol Depend* 1996; **41**: 101–12.
23. Muthén B., Grant B., Hasin D. The dimensionality of alcohol abuse and dependence: factor analysis of DSM-III and proposed DSM-IV criteria in the National Health Interview Survey 1988. *Addiction* 1993; **88**: 1079–90.
24. Harford T., Muthén B. The dimensionality of alcohol abuse and dependence: a multivariate analysis of DSM-IV symptom items in the National Longitudinal Survey of Youth. *J Stud Alcohol* 2001; **62**: 150–7.
25. Krueger R. F., Nichol P. E., Hicks B. M., Markon K. E., Patrick C. J., Iacono W. *et al.* Using latent trait modeling to conceptualize an alcohol problems continuum. *Psychol Assess* 2004; **16**: 107–19.
26. Asparouhov T., Muthén B. Maximum-likelihood estimation in general latent variable modeling. Draft 2004.
27. Muthén B., Asparouhov T. Item response mixture modeling: application to tobacco dependence criteria. *Addict Behav* 2006; **31**: 1050–66.
28. Muthén B., Asparouhov T., Rebollo I. Advances in behavioral genetics modeling using Mplus: applications of factor mixture modeling to twin data. Special Issue: Advances in Statistical Models and Methods. *Twin Res Human Genet* 2006; **9**: 313–24.
29. Wilson M. Saltus. A psychometric model of discontinuity in cognitive development. *Psychol Bull* 1989; **105**: 276–89.
30. Heinen T. *Latent Class and Discrete Latent Trait Models: similarities and differences*. Thousand Oaks, CA: Sage; 1996.
31. Vermunt J. K. *Log-Linear Models for Event Histories.* Thousand Oakes, CA: Sage 1997.
32. Formann A. K., Kohlmann T., Three-parameter linear logistic latent class analysis. In: Hagenaars J. A., McCutcheon A., editors. *Applied Latent Class Analysis*, pp. 183–210. Cambridge: Cambridge University Press; 2002.
33. Krueger R. F., Markon K. E., Patrick C., Iacono W. G. Externalizing psychopathology in adulthood: a dimensional-spectrum conceptualization and its implications for DSM-V. *J Abnorm Psychol* 2005; **114**: 537–50.
34. Windle M., Scheidt D. M. Alcoholic subtypes: are two sufficient? *Addiction* 2004; **99**: 1508–19.
35. Bartholomew D. J., Knott M. *Latent Variable Models and Factor Analysis*, 2nd edn. London: Arnold; 1999.

36. Muthén B. Latent variable analysis: growth mixture modeling and related techniques for longitudinal data. In: Kaplan D., editor. *Handbook of Quantitative Methodology for the Social Sciences*, pp. 345–68. Newbury Park, CA: Sage; 2004.
37. Lo Y., Mendell N. R., Rubin D. B. Testing the number of components in a normal mixture. *Biometrika* 2001; **88**: 767–78.
38. McLachlan G. J., Peel D. *Finite Mixture Models*. New York: Wiley & Sons; 2000.
39. Muthén L., Muthén B. *Mplus User's Guide*, 4th edn. Los Angeles, CA: Muthén & Muthén; 1998–2006.
40. Waller N. G., Meehl P. E. *Multivariate Taxometric Procedures.* Thousand Oaks, CA: Sage; 1998.
41. Beauchaine T. P. Taxometrics and developmental psychopathology. *Dev Psychopathol* 2003; **15**: 501–27.
42. McDonald R. P. A review of multivariate taxometric procedures: distinguishing types from continua. *J Educ Behav Stat* 2003; **28**: 77–81.
43. Grant B. F., Dawson D. A., Stinson F. S., Chou S. P., Dufour M. C., Pickering RP. The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States 1991–1992 and 2001–2002. *Drug Alcohol Depend* 2004; **74**: 223–34.
44. Asparouhov T. Sampling weights in latent variable modeling. *Struct Equation Model* 2005; **12**: 411–34.
45. Muthén B., Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 1999; **55**: 463–9.
46. Muthén B., Brown C. H., Masyn K., Jo B., Khoo S. T., Yang C. C. *et al.* General growth mixture modeling for randomized preventive interventions. *Biostatistics* 2002; **3**: 459–75.
47. Muthén B., Khoo S. T., Francis D., Boscardin C. K. Analysis of reading skills development from kindergarten through first grade: an application of growth mixture modeling to sequential processes. In: Reise, S. R., Duan, N., editors. *Multilevel Modeling: methodological advances. issues, and applications*, pp. 71–89. Mahwah, NJ: Lawrence Erlbaum Associates; 2002.

