Mapping Common Psychiatric Disorders

Structure and Predictive Validity in the National Epidemiologic Survey on Alcohol and Related Conditions

Carlos Blanco, MD, PhD; Robert F. Krueger, PhD; Deborah S. Hasin, PhD; Shang-Min Liu, MS; Shuai Wang, PhD; Bradley T. Kerridge, MA; Tulshi Saha, PhD; Mark Olfson, MD, MPH

Context: Clinical experience and factor analytic studies suggest that some psychiatric disorders may be more closely related to one another, as indicated by the frequency of their co-occurrence, which may have etiologic and treatment implications.

Objective: To construct a virtual space of common psychiatric disorders, spanned by factors reflecting major psychopathologic dimensions, and locate psychiatric disorders in that space, as well as to examine whether the location of disorders at baseline predicts the prevalence and incidence of disorders at 3-year follow-up.

Design, Setting, and Patients: A total of 34,653 individuals participated in waves 1 and 2 of the National Epidemiologic Survey on Alcohol and Related Conditions.

Main Outcome Measures: The distance between disorders at wave 1, calculated using the loadings of the factors spanning the space of disorders as coordinates. This distance was correlated with the adjusted odds ratios for age, sex, and race/ethnicity of the prevalence and incidence of Axis I disorders in wave 2, with the aim of determining whether smaller distances between disorders at wave 1 predicts higher disorder prevalence and incidence at wave 2.

Results: A model with 3 correlated factors provided an excellent fit (Comparative Fit Index=0.99, Tucker-Lewis Index=0.98, root mean square error of approximation=0.008) for the structure of common psychiatric disorders and was used to span the space of disorders. Distances ranged from 0.070 (between drug abuse and dysthymia) to 1.032 (between drug abuse and avoidant personality disorder). The correlation of distance between disorders in wave 1 with adjusted odds ratios of prevalence in wave 2 was −0.56. The correlation of distance in wave 1 with adjusted odds ratios of incidence in wave 2 was −0.57.

Conclusions: Mapping psychiatric disorders can be used to quantify the distances among disorders. Proximity in turn can be used to predict prospectively the incidence and prevalence of Axis I disorders.


Preparations for DSM-5 bring to the fore underlying interrelationships among psychiatric disorders. Examination of these interrelationships should ideally inform the structure of the diagnostic classification system and guide research strategies and treatment approaches for each disorder. Disorders deemed to be more closely related to one another may share etiologic factors, clinical course, or treatment response.

Several studies have used latent variable techniques to investigate relationships among psychiatric disorders. In a seminal study, Krueger used 10 psychiatric diagnoses available in the National Comorbidity Survey to identify 2 correlated factors, one corresponding to externalizing disorders and the other to internalizing disorders. The latter could further be decomposed into 2 lower-order factors called anxious-misery or distress and fear. Other studies have confirmed this model, with small variations (such as whether the internalizing factor can be decomposed), and extended it to document its structural stability over time and cross-cultural validity in Western and non-Western countries. Indeed, the underlying structure of genetic risk for these disorders closely resembles the observed, phenotypic structure of their co-occurrence patterns. Using cross-sectional data from the World Health Organization World Mental Health Survey Initiative, Kessler and colleagues showed that 2 factors (one representing internaliz-
orders and their symptoms to co-occur and respond to simi-
generalized anxiety disorder (GAD) than to substance use
order (MDD) might be considered closer to dysthymia or
cal manifestations, likelihood of co-occurrence, etiology,
can be considered more closely related regarding their clini-
matoform and antagonism dimensions. Taken together,
these findings suggest that there are a limited number of
common causal pathways for the common disorders. An
implicit assumption of research in this area, which is re-
lected in the structure of psychiatric classifications and reflects clinicians’ experience, is that some disorders are
more similar to one another than others; therefore, they
be considered more closely related regarding their clinical
manifestations, likelihood of co-occurrence, etiology, or
treatment response. For example, major depressive dis-
order (MDD) might be considered closer to dysthymia or
generalized anxiety disorder (GAD) than to substance use
disorders (SUDs), based on the tendencies for these dis-
orders and their symptoms to co-occur and respond to simi-
lar treatment strategies.

We sought to build on prior work by developing a for-
mal measure of similarity among common disorders. One
means of operationalizing the similarity of disorders to
one another is to conceive of them as existing along 1 or
more dimensions in space with values (represented as co-
ordinates) on each dimension corresponding to the role
each dimension plays in the occurrence of the disorder.
Comparing the location of disorders to one another along
these dimensions then provides a formal measure of their
proximity. We sought to examine the proximity among
disorders using data from a large nationally representa-
tive sample of US adults who were followed up 3 years
later and assessed for a broad range of Axis I and Axis II
disorders. Based on the prior literature, we expected that
disorders included within the same DSM chapters (eg, mood disorders) would tend to be closer to each other
than to disorders from other chapters.

The pattern of disorder development over time pro-
vides a means of evaluating the validity of the disorder
map. For example, if GAD was closer to MDD than to
SUDs at baseline, we would expect that the correlation
between having GAD at baseline and MDD at follow-up
would be higher than the correlation between GAD and
SUDs at follow-up. After assigning each disorder to a po-
sition on the dimensions, we tested the validity of these
positions by evaluating correlations with prevalent and
incident disorders at follow-up. The results provide an
empirical map of common mental disorders that may help
guide etiologic research and clinically focus the differ-
ential diagnosis assessment on neighboring disorders.

**STATISTICAL ANALYSES**

**Identification of Dimensions**

Entering all Axis I and Axis II disorders included in wave 1 of the
NESARC as indicators, we used exploratory factor analysis (EFA) to identify the potential latent factors (ie, dimensions).
Exploratory factor analysis was chosen over confirmatory factor analyses (CFA) to allow for cross-loadings. Factor
selection was guided by several considerations including the scree test, fit indices, parallel analyses, and interpretability. Each
factor was subsequently interpreted as a latent dimension, which could be represented spatially by an axis. Those axes spanned
a vectorial space that was subsequently used to calculate the distance between each disorder and each axis, as well as be-
tween all pairs of disorders.

The default estimator for the analysis was mean and variance-
adjusted weighted least squares, a robust estimator that does not assume normally distributed variables and provides the best
option for modeling categorical data. The fit indices used for model evaluation were the Comparative Fit Index (CFI), the
Tucker-Lewis Index (TLI), the root mean square error of approx-
imation (RMSEA), and the number of free parameters in the
model. Values of CFI and TLI greater than 0.95 and values of
RMSEA less than 0.06 are commonly used guidelines for in-
ferring reasonably good model fit. A number of options are
available for rotating the factor structure extracted in an EFA. We used the oblique geomin rotation, which is the recommended approach in Mplus (www.statmodel.com). All analyses were conducted using Mplus, which uses a sandwich estimator to correct standard errors for the nonindependence of the data resulting from the complex design of the NESARC.

Coordinates of the Disorders

In factor analysis, the loadings are the regression coefficients of the indicators (disorders) on the latent variables. They estimate the amount of change on the latent response variables underlying the binary indicators given a unit change on the factor. Higher factor loadings indicate that a particular indicator (ie, disorder) shares larger amounts of variance with the common factor (ie, dimension) that is being defined by the set of indicators included in the model. Thus, the loadings of the disorders on the factors can be used to create a system of coordinates that locates each disorder in the space spanned by the factors. High positive values of the coordinates indicate strong association of a disorder with that dimension and that the dimension is an important component in the occurrence of the disorder. Low values on a dimension would indicate that the dimension plays a smaller role in the disorder’s occurrence, whereas negative values indicate an inverse association with the disorder.

Distance Between Disorders in Wave 1 and Prevalence and Incidence of Axis I Disorders in Wave 2

To estimate the proximity of 2 disorders, we calculated the Euclidean distance between them. The Euclidean distance between 2 points in a space (or hyperplane) is the square root of the sum of squares of its coordinates, applying a generalization for higher dimensional spaces of the well-known Pythagorean theorem: \[ \text{distance} = \sqrt{a^2 + b^2}, \] where \(a\) and \(b\) are vectors whose elements are the loadings of the disorders on the factors.

To assess the predictive validity of this measure, we first estimated, the adjusted odds ratios (AORs) for age, sex, and race/ethnicity of the 12-month prevalence and incidence of Axis I disorders at wave 2. Pearson correlation coefficient was then used to assess the strength of association between distance among disorders in wave 1 and the AORs of the 12-month prevalence and incidence of disorders at wave 2. Because smaller distances between pairs of disorders at wave 1 would indicate a higher degree of similarity between those 2 disorders, we hypothesized that the closer 2 disorders were at wave 1, the better they would predict the incidence and prevalence of Axis I disorders at wave 2 as measured by the AORs.

Complementary Analyses

To assess the predictive value of alternative approaches, we conducted a CFA for the 3-factor and 4-factor models in which disorders were constrained to load exclusively on the factor on which they had the largest loading in the EFA, as well as a bifactor model with a general psychopathology factor loading on all disorders in addition to disorder-specific factors. We then calculated the correlation between the distances derived from those models and the prevalence and incidence of Axis I disorders at wave 2. We also calculated the correlation between the prevalence and incidence of Axis I disorders in wave 2 and the inverse of the odds ratios (ORs) between pairs of disorders in wave 1. The inverse of the ORs was used rather than the OR per se to ensure that the correlation would be negative, thus comparable to our correlation with distance measures (because greater ORs indicate stronger association, whereas larger distances indicate weaker association in the multidimensional space).

RESULTS

The eigenvalues for the first 5 factors were 7.97, 2.32, 1.87, 1.58, and 1.03. The 95th percentiles of the eigenvalues for the 5 factors derived from the parallel analyses were 1.04, 1.04, 1.03, 1.03, and 1.02, which supported the extraction of up to 5 factors. Fit indices indicated that a 1-factor model had poor fit (CFI=0.91, TLI=0.89, RMSEA=0.019). A model with 2 factors provided a reasonably good fit (CFI=0.95, TLI=0.94, RMSEA=0.014), whereas models with 3 (CFI=0.99, TLI=0.98, RMSEA=0.008), 4 (CFI=0.99, TLI=0.99, RMSEA=0.006), and 5 (CFI=1.00, TLI=0.99, RMSEA=0.005) factors provided excellent and comparable fit. In the 2-factor model, factor 1 had the highest loadings on SUDs, pathological gambling, and antisocial personality disorders, whereas factor 2 had higher loadings on mood and anxiety disorders and the other 6 personality disorders assessed in wave 1. The correlation between the 2 factors was 0.44 (Table 1).

In the models with 3 to 5 factors, the first factor replicated the first factor of the 2-factor model, whereas the remaining factors constituted an expansion of the other factor. In the 3-factor model, factor 2 of the 2-factor model subdivided into a factor that included bipolar disorder, social anxiety disorder, specific phobia, and all personality disorders except antisocial personality disorder, and another factor that included MDD, dysthymia, GAD, and panic disorder. In the 4-factor model, factor 2 of the 2-factor model subdivided into 3 factors: a first factor included MDD, dysthymia, and bipolar disorder; a second factor included panic disorder, social anxiety disorder, GAD, and specific phobia; and a third factor included all the personality disorders except antisocial personality disorder. The 5-factor model was very similar to the 4-factor model but with avoidant and dependent personality disorders splitting into a separate factor (Table 1).

Based on these results, the 3-factor, 4-factor, and 5-factor models were further evaluated. However, because the results for 3 models were very similar (possibly because of the relatively high correlations among factors with higher loadings from internalizing disorders), only the results of the 3-factor model (which is graphically represented in the Figure) are presented. Detailed results of the 4-factor and 5-factor models are presented in eTable 1 and eTable 2 (http://www.jamapsych.com).

COORDINATES AND DISTANCE BETWEEN DISORDERS

The coordinates for each disorder can be derived from Table 1. For example, in the 3-factor model, the coordinates for alcohol abuse were 0.476, −0.217, and −0.036, whereas for MDD, they were 0.037, 0.019, and 0.784. There was broad variation in the pattern of coordinates.
Some disorders, such as alcohol dependence and avoidant personality disorder, had coordinates with large values in 1 dimension but low in the others. Other disorders, such as bipolar disorder and histrionic personality disorder, had coordinates with moderately large values in 1 dimension and intermediate values in the others. Table 2 and Table 3 present the Euclidian distances among all pairs of disorders in the 3-factor model. Because Euclidian distances are symmetric, only the cells under the matrix diagonal are presented. Although there was a wide range of distances, with the smallest distance being between drug abuse and alcohol dependence (0.070) and the largest between drug abuse and dysthymia (1.032), the overall pattern indicated that disorders traditionally considered to be more closely related (e.g., alcohol dependence and drug dependence) tended to have smaller distances separating them than did disorders considered less similar such as alcohol dependence and MDD.

CORRELATION BETWEEN DISTANCE IN WAVE 1 AND PREVALENCE AND INCIDENCE IN WAVE 2

In wave 2, the largest AORs for prevalence were for the disorders with themselves, followed by disorders that were at short distances in wave 1 (Table 4). The correlation coefficients between the distance between pairs of disorders in wave 1 and AORs of prevalence in wave 2 was $-0.56$ ($P < .001$), indicating that larger distances in wave 1 were related to weaker associations in wave 2. A similar pattern was found for incidence; although in the case of incidence, by definition, the AOR of the disorder with itself was not computed because the disorder was already present in wave 1 (Table 5). The correlation between distance in wave 1 and the AORs of incidence in wave 2 was $-0.57$ ($P < .001$), indicating a strong inverse relationship between distance and incidence. The correlation coefficients for prevalence and incidence were also $-0.56$ and $-0.57$, respectively, using the 4-factor model, whereas they were $-0.54$ and $-0.55$, respectively, for the 5-factor model. Overall, the results indicate that the closer 2 disorders were in wave 1, the better predictors they were of the prevalence and incidence of each other in wave 2.

### Table 1. Factor Analyses of Axis I and Axis II Disorders in Wave 1 of the National Epidemiologic Survey on Alcohol and Related Conditions

<table>
<thead>
<tr>
<th>Index</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFI</td>
<td>0.954</td>
<td>0.986</td>
<td>0.993</td>
<td>0.997</td>
</tr>
<tr>
<td>TLI</td>
<td>0.943</td>
<td>0.980</td>
<td>0.989</td>
<td>0.994</td>
</tr>
<tr>
<td>RMSEA</td>
<td>0.014</td>
<td>0.008</td>
<td>0.006</td>
<td>0.005</td>
</tr>
</tbody>
</table>

**Disorder**
- AA: alcohol abuse
- AD: alcohol dependence
- DD: drug dependence
- DA: drug abuse
- D: dysthymia
- PD: personality disorder
- PG: pathological gambling
- RMSEA: root mean square error of approximation
- TLI: Tucker-Lewis Index

**Factor correlation**

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>1</th>
<th>2</th>
<th>1</th>
<th>2</th>
<th>1</th>
<th>2</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Factor 2</td>
<td>0.44</td>
<td>1.00</td>
<td>0.49</td>
<td>1.00</td>
<td>0.28</td>
<td>1.00</td>
<td>0.23</td>
<td>1.00</td>
<td>0.47</td>
<td>1.00</td>
</tr>
<tr>
<td>Factor 3</td>
<td>0.25</td>
<td>0.42</td>
<td>1.00</td>
<td>0.30</td>
<td>0.46</td>
<td>1.00</td>
<td>0.47</td>
<td>1.00</td>
<td>0.31</td>
<td>0.62</td>
</tr>
<tr>
<td>Factor 4</td>
<td>0.40</td>
<td>0.40</td>
<td>0.64</td>
<td>1.00</td>
<td>0.31</td>
<td>0.31</td>
<td>0.62</td>
<td>1.00</td>
<td>0.18</td>
<td>0.43</td>
</tr>
<tr>
<td>Factor 5</td>
<td>0.18</td>
<td>0.43</td>
<td>0.55</td>
<td>1.00</td>
<td>0.18</td>
<td>0.43</td>
<td>0.55</td>
<td>1.00</td>
<td>0.18</td>
<td>0.43</td>
</tr>
</tbody>
</table>

**Abbreviations:** AA, alcohol abuse; AD, alcohol dependence; BD, bipolar disorder; CFI, Comparative Fit Index; DA, drug abuse; DD, drug dependence; GAD, generalized anxiety disorder; MDD, major depressive disorder; NID, nicotine dependence; OCPD, obsessive compulsive personality disorder; PD, personality disorder; PG, pathological gambling; RMSEA, root mean square error of approximation; SAD, social anxiety disorder; TLI, Tucker-Lewis Index.

### COMPLEMENTARY ANALYSES

The fit indices for the bifactor model with a general factor and 3 disorder-specific factors were 0.982 for CFI, 0.977 for TLI, and 0.009 for RMSEA (Table 3), and the correlation of distance with the prevalence and incidence of Axis I disorders in wave 2 were $-0.56$ and $-0.57$, respectively, whereas a bifactor model with 4 disorder-
specific factors produced implausible parameter estimates (e.g., standard errors considerably larger than parameter estimates on some of the disorder-specific factors). The fit indices for the 3-factor CFA model were 0.959 for CFI, 0.953 for TLI, and 0.0013 for RSMEA, whereas for the 4-factor CFA model, they were 0.969 for TLI, and 0.010 for RSMEA. Although the fit of these CFA models was slightly worse than for the corresponding EFA models, it was still good for both the 3-factor and 4-factor CFA models. However, for the 3-factor model, the correlation of distance with prevalence of Axis I disorders in wave 2 was only $r^2 = 0.42$, whereas the correla-

![Figure](https://jamanetwork.com/)

**Figure.** Three-dimensional representation of the space among psychiatric disorders. Each disorder is represented in the color of the factor for which it has higher loadings. Although the factors are correlated (see Table 1 and the “Results” section), they are represented as orthogonal to facilitate visualization. 1 indicates alcohol abuse; 2, alcohol dependence; 3, drug abuse; 4, drug dependence; 5, nicotine dependence; 6, major depressive disorder; 7, bipolar disorder; 8, dysthymia; 9, panic disorder; 10, social anxiety disorder; 11, specific phobia; 12, generalized anxiety disorder; 13, pathological gambling; 14, avoidant personality disorder; 15, dependent personality disorder; 16, obsessive compulsive personality disorder; 17, paranoid personality disorder; 18, schizoid personality disorder; 19, histrionic personality disorder; 20, antisocial personality disorder.

**Table 2. Euclidian Distances Among Axis I and Axis II Disorders in Wave 1 of the NESARC in the 3-Factor Model**

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Alcohol Abuse</th>
<th>Alcohol Dependence</th>
<th>Drug Abuse</th>
<th>Drug Dependence</th>
<th>Nicotine Dependence</th>
<th>MDD</th>
<th>Bipolar Disorder</th>
<th>Dysthymia</th>
<th>Panic Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol abuse</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>0.43</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug abuse</td>
<td>0.39</td>
<td>0.07</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug dependence</td>
<td>0.58</td>
<td>0.20</td>
<td>0.27</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicotine dependence</td>
<td>0.41</td>
<td>0.19</td>
<td>0.24</td>
<td>0.21</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDD</td>
<td>0.90</td>
<td>0.90</td>
<td>0.94</td>
<td>0.83</td>
<td>0.74</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>0.74</td>
<td>0.61</td>
<td>0.66</td>
<td>0.49</td>
<td>0.43</td>
<td>0.59</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysthymia</td>
<td>0.98</td>
<td>0.99</td>
<td>1.03</td>
<td>0.92</td>
<td>0.84</td>
<td>0.10</td>
<td>0.69</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Panic disorder</td>
<td>0.75</td>
<td>0.71</td>
<td>0.76</td>
<td>0.62</td>
<td>0.52</td>
<td>0.40</td>
<td>0.22</td>
<td>0.49</td>
<td>0.00</td>
</tr>
<tr>
<td>SAD</td>
<td>0.91</td>
<td>0.91</td>
<td>0.96</td>
<td>0.83</td>
<td>0.73</td>
<td>0.61</td>
<td>0.35</td>
<td>0.69</td>
<td>0.30</td>
</tr>
<tr>
<td>Specific phobia</td>
<td>0.68</td>
<td>0.71</td>
<td>0.75</td>
<td>0.66</td>
<td>0.52</td>
<td>0.56</td>
<td>0.26</td>
<td>0.65</td>
<td>0.21</td>
</tr>
<tr>
<td>GAD</td>
<td>0.88</td>
<td>0.83</td>
<td>0.89</td>
<td>0.73</td>
<td>0.65</td>
<td>0.37</td>
<td>0.31</td>
<td>0.44</td>
<td>0.13</td>
</tr>
<tr>
<td>PG</td>
<td>0.39</td>
<td>0.41</td>
<td>0.41</td>
<td>0.47</td>
<td>0.32</td>
<td>0.90</td>
<td>0.49</td>
<td>1.00</td>
<td>0.60</td>
</tr>
<tr>
<td>Avoidant PD</td>
<td>0.98</td>
<td>0.93</td>
<td>0.97</td>
<td>0.83</td>
<td>0.75</td>
<td>0.76</td>
<td>0.35</td>
<td>0.83</td>
<td>0.40</td>
</tr>
<tr>
<td>Dependent PD</td>
<td>0.90</td>
<td>0.81</td>
<td>0.86</td>
<td>0.72</td>
<td>0.65</td>
<td>0.79</td>
<td>0.28</td>
<td>0.88</td>
<td>0.40</td>
</tr>
<tr>
<td>OCPD</td>
<td>0.82</td>
<td>0.81</td>
<td>0.84</td>
<td>0.75</td>
<td>0.64</td>
<td>0.83</td>
<td>0.36</td>
<td>0.91</td>
<td>0.45</td>
</tr>
<tr>
<td>Paranoid PD</td>
<td>0.90</td>
<td>0.80</td>
<td>0.85</td>
<td>0.70</td>
<td>0.65</td>
<td>0.86</td>
<td>0.31</td>
<td>0.95</td>
<td>0.47</td>
</tr>
<tr>
<td>Schizoid PD</td>
<td>0.85</td>
<td>0.81</td>
<td>0.85</td>
<td>0.74</td>
<td>0.65</td>
<td>0.80</td>
<td>0.32</td>
<td>0.89</td>
<td>0.42</td>
</tr>
<tr>
<td>Histrionic PD</td>
<td>0.73</td>
<td>0.62</td>
<td>0.66</td>
<td>0.56</td>
<td>0.50</td>
<td>0.93</td>
<td>0.36</td>
<td>1.02</td>
<td>0.54</td>
</tr>
<tr>
<td>Antisocial PD</td>
<td>0.51</td>
<td>0.32</td>
<td>0.35</td>
<td>0.29</td>
<td>0.21</td>
<td>0.85</td>
<td>0.38</td>
<td>0.95</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Abbreviations: GAD, generalized anxiety disorder; MDD, major depressive disorder; NESARC, National Epidemiologic Survey on Alcohol and Related Conditions; OCPD, obsessive compulsive personality disorder; PD, personality disorder; PG, pathological gambling; SAD, social anxiety disorder.
We found that the patterns of associations of psychiatric disorders could be described by a relatively small number of underlying dimensions. Our findings are consistent with previous studies, which have documented that 2 to 3 correlated factors (ie, dimensions) capture the patterns of comorbidity among the most common psychiatric disorders.\cite{8,10,12,33} The dimensions identified in our study are similar to those identified in previous investigations (which include an externalizing dimension and a variable number of internalizing dimensions), supporting the robustness of these findings across different samples and extending them by including 7 personality disorders. By using the loadings to align the disorders along those dimensions, it was possible to obtain an estimate of the relative importance of each dimension for each disorder and create a space in which psychiatric disorders could be placed.

Disorders considered nosologically related (eg, those grouped in the same DSM-IV chapters) had generally

In a large nationally representative sample, interrelationships among disorders were well described by 3 correlated dimensions, 2 of which included Axis I and Axis II disorders. When the factors were used to span a space and their loadings used as the coordinates in that space, disorders included in the same DSM-IV diagnostic groups tended to have smaller distances among themselves than from disorders in other diagnostic groups. Furthermore, proximity across disorders significantly predicted prevalence and incidence in wave 2.

### Table 3. Additional Euclidian Distances Among Axis I and Axis II Disorders in Wave 1 of the NESARC in the 3-Factor Model

<table>
<thead>
<tr>
<th>Specific phobia</th>
<th>GAD</th>
<th>PG</th>
<th>Avoidant PD</th>
<th>Dependent PD</th>
<th>OCPD</th>
<th>Paranoid PD</th>
<th>Schizoid PD</th>
<th>Histrionic PD</th>
<th>Antisocial PD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific phobia</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAD</td>
<td>0.29</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PG</td>
<td>0.48</td>
<td>0.73</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoidant PD</td>
<td>0.33</td>
<td>0.39</td>
<td>0.69</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent PD</td>
<td>0.32</td>
<td>0.44</td>
<td>0.58</td>
<td>0.14</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCPD</td>
<td>0.29</td>
<td>0.50</td>
<td>0.50</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paranoid PD</td>
<td>0.38</td>
<td>0.51</td>
<td>0.56</td>
<td>0.22</td>
<td>0.08</td>
<td>0.18</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizoid PD</td>
<td>0.29</td>
<td>0.46</td>
<td>0.53</td>
<td>0.17</td>
<td>0.10</td>
<td>0.07</td>
<td>0.13</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Histrionic PD</td>
<td>0.43</td>
<td>0.63</td>
<td>0.36</td>
<td>0.44</td>
<td>0.30</td>
<td>0.28</td>
<td>0.24</td>
<td>0.29</td>
<td>0.00</td>
</tr>
</tbody>
</table>
| Antisocial PD   | 0.49| 0.66| 0.22        | 0.65         | 0.52 | 0.51        | 0.50        | 0.52         | 0.31         | 0.00

Abbreviations: GAD, generalized anxiety disorder; MDD, major depressive disorder; NESARC, National Epidemiologic Survey on Alcohol and Related Conditions; OCPD, obsessive compulsive personality disorder; PD, personality disorder; PG, pathological gambling; SAD, social anxiety disorder.

### Table 4. Adjusted Odds Ratios of the Prevalence of Axis I and Axis II Disorders in Wave 1 With Prevalence of Axis I in Wave 2

<table>
<thead>
<tr>
<th>Wave 1</th>
<th>Alcohol Abuse</th>
<th>Alcohol Dependence</th>
<th>Drug Abuse</th>
<th>Drug Dependence</th>
<th>Nicotine Dependence</th>
<th>MDD</th>
<th>Bipolar Disorder</th>
<th>Dysthymia</th>
<th>Panic Disorder</th>
<th>SAD</th>
<th>Specific phobia</th>
<th>GAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol abuse</td>
<td>6.56*</td>
<td>2.73*</td>
<td>2.53*</td>
<td>1.89*</td>
<td>1.71*</td>
<td>0.91</td>
<td>1.07</td>
<td>0.70</td>
<td>0.96</td>
<td>0.74</td>
<td>1.03</td>
<td>1.09</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>2.22</td>
<td>10.91</td>
<td>3.48</td>
<td>4.32</td>
<td>3.01</td>
<td>1.77</td>
<td>2.62</td>
<td>1.43</td>
<td>2.76</td>
<td>2.31</td>
<td>1.35</td>
<td>2.21</td>
</tr>
<tr>
<td>Drug abuse</td>
<td>2.47</td>
<td>4.46</td>
<td>15.33</td>
<td>7.62</td>
<td>3.23</td>
<td>1.23</td>
<td>2.93</td>
<td>0.46</td>
<td>2.78</td>
<td>3.02</td>
<td>1.63</td>
<td>1.77</td>
</tr>
<tr>
<td>Drug dependence</td>
<td>1.51</td>
<td>7.18</td>
<td>6.45</td>
<td>22.23</td>
<td>6.52</td>
<td>1.71</td>
<td>6.95</td>
<td>1.85</td>
<td>4.69</td>
<td>4.05</td>
<td>2.19</td>
<td>4.74</td>
</tr>
<tr>
<td>Nicotine dependence</td>
<td>1.49</td>
<td>3.13</td>
<td>2.69</td>
<td>4.51</td>
<td>19.72</td>
<td>1.66</td>
<td>2.81</td>
<td>2.50</td>
<td>2.99</td>
<td>2.37</td>
<td>1.96</td>
<td>2.10</td>
</tr>
<tr>
<td>MDD</td>
<td>1.04</td>
<td>1.78</td>
<td>1.83</td>
<td>2.84</td>
<td>1.71</td>
<td>5.16</td>
<td>2.45</td>
<td>8.22</td>
<td>2.90</td>
<td>4.03</td>
<td>2.04</td>
<td>3.92</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>0.96</td>
<td>3.12</td>
<td>2.18</td>
<td>2.92</td>
<td>2.49</td>
<td>19.26</td>
<td>1.36</td>
<td>5.33</td>
<td>5.14</td>
<td>3.09</td>
<td>4.61</td>
<td></td>
</tr>
<tr>
<td>Dysthymia</td>
<td>0.86</td>
<td>1.76</td>
<td>2.84</td>
<td>3.94</td>
<td>2.34</td>
<td>5.07</td>
<td>3.07</td>
<td>15.53</td>
<td>3.60</td>
<td>5.75</td>
<td>2.86</td>
<td>4.44</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>0.91</td>
<td>2.86</td>
<td>2.70</td>
<td>5.12</td>
<td>3.01</td>
<td>2.52</td>
<td>6.11</td>
<td>1.98</td>
<td>10.09</td>
<td>6.54</td>
<td>3.94</td>
<td>4.44</td>
</tr>
<tr>
<td>SAD</td>
<td>0.76</td>
<td>1.73</td>
<td>1.38</td>
<td>4.13</td>
<td>1.79</td>
<td>2.33</td>
<td>5.22</td>
<td>2.71</td>
<td>4.53</td>
<td>12.78</td>
<td>3.54</td>
<td>4.46</td>
</tr>
<tr>
<td>Specific phobia</td>
<td>0.87</td>
<td>1.74</td>
<td>1.90</td>
<td>1.75</td>
<td>1.68</td>
<td>1.94</td>
<td>2.80</td>
<td>2.01</td>
<td>3.11</td>
<td>3.85</td>
<td>4.58</td>
<td>2.53</td>
</tr>
<tr>
<td>GAD</td>
<td>1.07</td>
<td>2.10</td>
<td>2.21</td>
<td>4.43</td>
<td>2.78</td>
<td>3.00</td>
<td>8.73</td>
<td>6.73</td>
<td>7.28</td>
<td>6.48</td>
<td>2.96</td>
<td>7.72</td>
</tr>
<tr>
<td>PG</td>
<td>0.50</td>
<td>6.68</td>
<td>1.99</td>
<td>2.29</td>
<td>3.93</td>
<td>2.45</td>
<td>5.82</td>
<td>4.64</td>
<td>1.98</td>
<td>0.86</td>
<td>2.59</td>
<td>2.26</td>
</tr>
<tr>
<td>Avoidant PD</td>
<td>0.66</td>
<td>2.35</td>
<td>1.31</td>
<td>5.76</td>
<td>2.05</td>
<td>2.43</td>
<td>8.53</td>
<td>4.22</td>
<td>4.75</td>
<td>13.05</td>
<td>3.19</td>
<td>5.41</td>
</tr>
<tr>
<td>Dependent PD</td>
<td>0.36</td>
<td>3.53</td>
<td>1.61</td>
<td>14.82</td>
<td>4.48</td>
<td>2.07</td>
<td>10.73</td>
<td>4.27</td>
<td>7.99</td>
<td>13.07</td>
<td>3.68</td>
<td>11.28</td>
</tr>
<tr>
<td>OCPD</td>
<td>1.09</td>
<td>1.79</td>
<td>1.55</td>
<td>2.23</td>
<td>1.64</td>
<td>1.57</td>
<td>4.06</td>
<td>2.23</td>
<td>2.97</td>
<td>3.15</td>
<td>2.00</td>
<td>3.08</td>
</tr>
<tr>
<td>Paranoid PD</td>
<td>0.95</td>
<td>2.58</td>
<td>1.90</td>
<td>3.96</td>
<td>2.62</td>
<td>1.85</td>
<td>7.01</td>
<td>2.82</td>
<td>4.44</td>
<td>5.27</td>
<td>2.79</td>
<td>4.44</td>
</tr>
<tr>
<td>Schizoid PD</td>
<td>0.78</td>
<td>2.07</td>
<td>2.13</td>
<td>5.36</td>
<td>2.56</td>
<td>2.36</td>
<td>5.20</td>
<td>5.06</td>
<td>3.38</td>
<td>5.45</td>
<td>2.47</td>
<td>4.04</td>
</tr>
<tr>
<td>Histrionic PD</td>
<td>1.26</td>
<td>3.67</td>
<td>3.04</td>
<td>4.29</td>
<td>2.91</td>
<td>1.36</td>
<td>6.45</td>
<td>1.90</td>
<td>4.24</td>
<td>4.27</td>
<td>2.53</td>
<td>3.85</td>
</tr>
<tr>
<td>Antisocial PD</td>
<td>1.15</td>
<td>2.69</td>
<td>3.14</td>
<td>6.10</td>
<td>3.19</td>
<td>1.27</td>
<td>4.20</td>
<td>1.23</td>
<td>3.46</td>
<td>2.40</td>
<td>2.12</td>
<td>2.69</td>
</tr>
</tbody>
</table>

Abbreviations: GAD, generalized anxiety disorder; MDD, major depressive disorder; NESARC, National Epidemiologic Survey on Alcohol and Related Conditions; OCPD, obsessive compulsive personality disorder; PD, personality disorder; PG, pathological gambling; SAD, social anxiety disorder.

*Significant result.
smaller distances among themselves than did those grouped in different categories. Furthermore, proximity at baseline, as measured by the distances obtained from the EFA, was highly correlated with the incidence and prevalence of disorders 3 years later, and it provided a stronger measure of association than alternative approaches such as using CFA to estimate distances or using the ORs at baseline to predict the prevalence and incidence of disorders at follow-up. Our findings are generally consistent with the existing DSM-IV chapter headings and suggest that, in addition to having face validity, these categories provide information about the underlying similarities across disorders and have prognostic validity. These patterns are also consistent with findings by Krueger et al.9 and Vollebergh et al.10 who found substantive within-subject stability (also called rank or differential stability) on the scores in the latent factors. By using the concept of proximity in the space of psychiatric disorders and a prospective design, our results extend and generalize recent findings from the World Mental Health Survey Initiative.33 Using a retrospective design, that study found that 2 latent variables, one representing an internalizing dimension and the other representing an externalizing dimension, predicted the onset of psychiatric disorders during the life course. Taken together, these findings may help inform development of the DSM-5 and other future classifications of psychiatric disorders.1,3

To our knowledge, our study is only the third to include a substantial number of Axis I and Axis II disorders,14,34 and the first to use a nationally representative sample. In the 3-factor model, the dimensions identified by our analyses did not neatly correspond with Axis I and Axis II disorders but rather corresponded to the more general processes identified in prior structural psychiatric nosologic studies. In accord with results from clinical samples,32 disorders considered nosologically related, such as social anxiety disorder and avoidant personality disorder, had similar loadings across the dimensions and were therefore close to one another.35 The same was true of antisocial personality disorder and Axis I disorders characterized by high impulsivity such as pathological gambling and SUDs. These findings are consistent with current conceptualizations that question the categorical distinction between Axis I and Axis II disorders36-38 and that relate psychiatric disorders with personality traits described in the broader psychologic literature.39,40

By contrast, our 4-factor model revealed a dimension that loaded on all personality disorders assessed in wave 1 except antisocial personality disorder. This model may be more in line with recent findings from the Norwegian Twin Panel Study,34 which found that 4 factors (2 of which loaded mainly on personality disorders) provided a good fit for 25 Axis I and Axis II disorders. However, even in that study, the pattern of intercorrelations among factors, as well as the loadings of antisocial personality disorder, borderline personality disorder, and dysthymia, suggested some overlap rather than clear separation between Axis I and Axis II disorders. Furthermore, our 4-factor model did not substantially improve the fit of the model, change the distances among disorders, or modify the strength of association between the distance among disorders and the prevalence and incidence of wave 2 disorders. The 5-factor model was very similar to the 4-factor model and, although it had a slightly better fit, it also had slightly lower correlations with the incidence and prevalence of Axis I disorders in wave 2. The bifactor model also provided a good fit to the data, consistent with promising results in a recent twin study of adolescent psychopathology.31 As our understanding of the structure of mental disorders continues to advance, it is possible that for certain applications, the 4-factor or 5-factor...
rather than discrete entities.8,10,45 Our approach helps sup-
ceptualizations of psychiatric disorders as continuous
lineating a space, our study is consistent with recent con-
shown substantial cross-loadings among factors. By de-
defining the genetic space, our findings are consistent with previous
molecular genetics48 research, suggesting that disorders
are best conceptualized as multidimensional, having fac-
ents of both internalizing and externalizing disorders.30,43
Prior epidemiologic11,46 and twin studies12,33 have also shown
stantial cross-loadings among factors. By de-
delineating a space, our study is consistent with recent con-
ceptualizations of psychiatric disorders as continuous
rather than discrete entities.8,10,12 When variables are posi-
tively correlated, increases in one variable are associ-
ated with increases in the other. Therefore, a disorder such
as drug dependence that loads predominantly on 1 di-
msion (externalizing disorders) does not decrease but
rather increases the risk for disorders, such as social anxiety
order or GAD, that load predominantly on other
dimensions. Our findings are consistent with previous
studies that have documented that the presence of 1 psy-
chiatric disorder increases the risk for most other psy-
chiatric disorders.22,46,47 From the etiologic point of view, our analyses con-
verge with findings from genetic epidemiology12,33 and
molecular genetics48 research, suggesting that disorders
that are closer to each other are more likely to share li-
abilities than those located further apart. Furthermore,
because each disorder loads on several dimensions (ie, liabilities), our findings suggest that disorders can de-
develop through multiple etiologic paths, although those
dimensions with higher loadings may be more common
paths.3 For example, the risk for nicotine depend-
ence may be particularly high in individuals with a pro-
ensity toward externalization, but it may also be in-
creased in individuals with some internalizing disorders
such as depression.40 Alternatively, it is also possible that
several dimensions may have to be affected before a full-
blown disorder manifests itself. For example, bipolar dis-
order may involve abnormalities in facets related to both
internalizing and externalizing dimensions. Also con-
verging with the current findings, recent research has
started to investigate whether proximity among disor-
ders is the same when examining the structure of their
shared genetic vs environmental variance. For example,
the Norwegian Twin Panel Study34 found that from a ge-
etic perspective, dysthymia is closely related to personal-
ity disorders (ie, close in the genetic space), whereas
its environmental risk factors place it much closer to ma-
or depression. By contrast, alcohol abuse or depen-
dence were environmentally closer to internalizing dis-
orders, but they shared genetic risk factors with Axis I
externalizing disorders. Continuing to uncover the sources of
proximities among disorders could help identify com-
mon and specific etiologic pathways for those disor-
ders, suggest which neurobiological mechanisms and brain
structures or circuits may be shared across disorders, and
help guide treatment research. Alternatively, treat-
ments that have efficacy for different disorders may sug-
gest previously undetected etiologic links among disor-
ders. The convergence of multimodal maps50 that integrate
data from genetics, neuroimaging, treatment response, and
possibly other domains should lead to a more empirically
based classification of psychiatric disorders.51
From the clinical perspective, knowledge of the prox-
imity among disorders may help narrow the differential
diagnoses as more closely related disorders are more likely
to co-occur. Because disorder proximity also provides in-
formation about which disorders are most likely to co-
ccur in the future, spatial mapping may help guide pre-
ventive interventions. Proximity among disorders can also
inform treatment development. Disorders that are close
to each other may have aspects similar to the one being
treated, thus they may need to be targeted or prevented.
Alternatively, the location of each disorder may suggest
how treatments that work for one disorder may need to be
modified to address other facets of a neighboring dis-
order. For example, the efficacy of selective serotonin
reuptake inhibitors in the treatment of social anxiety dis-
order may suggest their efficacy for the treatment of avoid-
ant personality disorder.
Our study should be understood in the context of sev-
eral limitations. First, although the study included a broad
range of disorders, some disorders, such as obsessive com-
pulsive disorder or eating disorders, were not assessed in the
NESARC. Inclusion of these disorders might have re-
sulted in identification of additional dimensions. Second,
the NESARC did not collect information on individuals
younger than age 18 years. It is possible that the space of
common psychiatric disorders differs between adults and
young people. Data from the National Comorbidity Survey–
Adolescent Supplement may be helpful in examining this
hypothesis. Third, our study used factor analysis and Eu-
clidian distances to create a space of disorders and calcu-
late the distances among disorders. An important avenue
for future research would be the use of other distances or
statistical techniques, such as multidimensional scaling,20
exploratory structural equation modeling,52 or self-
organizing maps,53 to examine whether alternative ap-
proaches can yield additional insights into the structure or
etiology of psychiatric disorders.
Despite these limitations, this study helps to
advance our understanding of the structure of several
DSM-IV Axis I and Axis II disorders. Psychiatric disor-
ders can be conceptualized as existing in a space with
a limited number of fundamental dimensions. Proxim-
ity in that space can be used to assess nosologic simi-
arity, provide a measure of the likelihood of current and
future co-occurrence across disorders, and may yield clues about shared etiology and treatment response. We hope this information can be useful in
the development of new classification systems and
guide research on etiology and treatment.
Submitted for Publication: January 5, 2012; final revision received June 18, 2012; accepted June 20, 2012. Published Online: December 24, 2012. doi:10.1001/jamapsychiatry.2013.281

Author Affiliations: College of Physicians and Surgeons, Department of Psychiatry (Drs Blanco, Hasin, Liu, Wang, and Olfson), Mailman School of Public Health, Department of Epidemiology (Dr Hasin), Columbia University; New York State Psychiatric Institute (Drs Blanco, Hasin, and Olfson), New York, New York; Department of Psychology, University of Minnesota, Minneapolis (Dr Krueger); Department of Epidemiology and Biostatistics, School of Public Health, University of Maryland, College Park (Mr Kerridge); and Intramural Laboratory of Epidemiology and Biometry, National Institute on Alcohol Abuse and Alcoholism, Rockville (Dr Saha), Maryland.

Correspondence: Carlos Blanco, MD, PhD, Department of Psychiatry, College of Physicians and Surgeons, Columbia University, 1051 Riverside Dr, Unit 69, New York, NY 10032 (cb255@columbia.edu).

Author Contributions: Dr Blanco had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of Interest Disclosures: Dr Olfson has worked on grants from Eli Lilly and Company and Bristol-Myers Squibb to Columbia University.

Funding/Support: The National Epidemiologic Survey on Alcohol and Related Conditions was sponsored by the National Institute on Alcohol Abuse and Alcoholism, with supplemental support from the National Institute on Drug Abuse. Work on this article was supported by grants DA019606, DA020783, DA023200, DA023973, CA133050, and MH082773 to Dr Blanco, AA014223 and AA018111 to Dr Hasin, and F31DA025377 to Mr Kerridge from the National Institutes of Health, and the New York State Psychiatric Institute (Drs Blanco, Hasin, and Olfson).


REFERENCES

32. Reiss SP, Morzot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. Qual Life Res. 2007;16(suppl 1):19-31.
33. Kendler KS, Aggen SH, Knudsen GP, Reynolds E, Neale MC, Reichborn-Kjennrud...


