Nicotine dependence, abuse, and craving: dimensionality in an Israeli sample

Dvora Shmulewitz, Ph.D. [1,2]
Katherine M. Keyes, Ph.D. [1,3]
Melanie M. Wall, Ph.D. [1,2,4]
Efrat Aharonovich, Ph.D. [1,2]
Christina Aivadyan, B.A. [1]
Eliana Greenstein, M.A. [1]
Baruch Spivak, M.D. [5]
Abraham Weizman, M.D. [5,6,7]
Amos Frisch, Ph.D. [5,6]
Bridget F. Grant, Ph.D., Ph.D. [8]
Deborah Hasin, Ph.D. [1,2,3]

Affiliations
[1] New York State Psychiatric Institute, New York, New York 10032, USA
[2] Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA
[3] Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York 10032, USA
[4] Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, New York 10032, USA
[5] Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel, 69978
[6] Felsenstein Medical Research Center, Petach Tikva, Israel, 49100
[7] Research Unit, Geha Mental Health Center, Petach Tikva, Israel, 49100
[8] Laboratory of Epidemiology and Biometry, National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland 20892, USA

Correspondence
Deborah S. Hasin, Ph.D., Columbia University College of Physicians and Surgeons, Department of Psychiatry, 1051 Riverside Drive #123, New York, NY 10032.
Phone: (212) 543-5035, Fax: (212) 543-5913, E-mail: dsh2@columbia.edu

Running Head
Dimensionality of nicotine diagnostic criteria in Israel

Word count: 3,496

Conflict of Interest
None of the authors or researchers has any connection with the tobacco, alcohol, pharmaceutical or gaming industries or any body substantially funded by one of these organizations.

Abstract

Aims. Evidence-based changes planned for DSM-5 substance use disorders (SUDs) include combining dependence and three of the abuse criteria into one disorder and adding a criterion indicating craving. Because DSM-IV did not include a category for nicotine abuse, little empirical support is available for aligning the nicotine use disorder criteria with the DSM-5 criteria for other SUDs. Design. Latent variable analyses, likelihood ratio tests (LRT) and bootstrap tests were used to explore the unidimensionality, psychometric properties and information of the nicotine criteria. Setting, Participants. A sample of household residents selected from the Israeli population register yielded 727 lifetime cigarette smokers. Measurements. DSM-IV nicotine dependence criteria and proposed abuse and craving criteria, assessed with a structured interview. Findings. Three abuse criteria (hazardous use, social/interpersonal problems, and neglect roles) were prevalent among smokers, formed a unidimensional latent trait with nicotine dependence criteria, were intermixed with dependence criteria across the severity spectrum, and significantly increased the diagnostic information over the dependence-only model. LRT results also supported including the abuse criteria ($X^2_3=259.63$, $p<0.0001$). A craving criterion was shown to fit well with the other criteria. Conclusion. Similar to findings from research on other substances, nicotine dependence, abuse, and craving criteria formed a single factor. The results support alignment of nicotine criteria with those for alcohol and drug use disorders in DSM-5.

Keywords: Item Response Theory, nicotine use disorders, nicotine dependence, DSM-IV, DSM-5, Israel

Introduction

In the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5), several changes have been proposed to the substance use disorder (SUD) diagnoses [1]. One change involves combining DSM-IV dependence and three abuse criteria (hazardous use; social/interpersonal problems related to use; neglect of roles to use) into a single disorder. This change is justified by many studies showing that the dependence and three abuse criteria for alcohol and drug use disorders were unidimensional, that combining dependence and three abuse criteria appeared to result in a more informative criteria set, and that differential item or test functioning by population subgroups was not pronounced enough to affect the overall diagnosis [2-15]. Legal problems (a DSM-IV abuse criterion) was not retained due to low prevalence [3,5,8,12], poor fit with other criteria [7,11,13] and little added SUD information [8,9,12]. Another proposed change is the addition of craving, which is in the International Classification of Disease, 10th edition [16], is considered by some to be a central feature of SUDs [17,18], and is unidimensional with DSM-IV criteria for alcohol use disorders [19,20].

DSM-IV included a diagnosis for nicotine dependence [21], whose criteria are unidimensional [22-26]. However, based on early expert opinion that the substance abuse criteria were not relevant to nicotine or not found without dependence [27], DSM-IV did not include a nicotine abuse category. Given the changes planned for the other SUDs, the DSM-5 workgroup addressed whether nicotine use disorder (NUD) criteria could be aligned with criteria for other SUDs by adding abuse and craving criteria, resulting in a consistent criterion set across all substances. The three abuse criteria now have face validity, perhaps due to less permissive attitudes towards smoking that might engender interpersonal problems related to smoking, or to job neglect due to restricted workplace smoking. Furthermore, with removal of DSM-IV's hierarchical relationship of abuse to dependence, nicotine abuse criteria should not be excluded due to co-occurrence with dependence. Nicotine craving is included in some nicotine dependence scales [28], is unidimensional with other DSM-IV nicotine criteria [26], and is considered by some as central to dependence [29]. Evidence supporting the addition of abuse and craving criteria to NUD would include unidimensionality, increased information, and lack of differential item or test functioning between population subgroups, similar to the evidence supporting the changes made for other SUDs.

Because DSM-IV did not include nicotine abuse and craving, little data were available to address these issues. However, a unique data source was found in an Israel general population study on genetic and environmental influences on drinking, smoking, and related traits, building on previous work [30-34]. The Israel study used the same diagnostic instrument as several other item response theory (IRT) studies that showed unidimensionality of substance dependence and abuse criteria [3,6,8,11,12,19,22,24,35], with additional nicotine items modeled on DSM-IV substance abuse in order to measure nicotine "abuse". This study already contributed findings supporting the proposed changes in SUD criteria [12]. Now, we address four questions using these data: (1) Are the DSM-IV nicotine dependence criteria unidimensional, similar to U.S. results? (2) Do the "abuse" criteria fit with the dependence criteria in a single dimension and add significantly to the diagnostic information? (3) Does evidence support adding a craving criterion? (4) Is differential item or test functioning (e.g., different probability of criterion endorsement or total test scores between demographic subgroups among respondents with the same level of underlying trait severity) found with these additions?

Methods and Materials

Study procedures

Adult household residents were selected from the Israeli population register with the following characteristics: Jewish ethnicity, and oversampling for males and for being an immigrant from the Former Soviet Union (FSU) vs. not being in this immigrant group. These characteristics were chosen to achieve the study goals: examining genetic and environmental influences on drinking and smoking. We oversampled males because Israeli women have low drinking rates [36]. FSU immigrants were selected because they have different smoking and drinking behaviors than other Israelis [32,37]. Data collection occurred from 2007-2009. Study procedures, including recruitment, consent and interviewing, have been described in detail previously [12,31,32,34]. In brief, potential respondents received an explanatory letter and follow-up call to schedule an in-person interview. After describing the study to potential participants, interviewers obtained written informed consent as approved by IRBs at New York State Psychiatric Institute, Tel Aviv University, Ness Ziona/Ba'er Yaakov Psychiatric Hospitals and the Israel Ministry of Health. Interviewers administered computer-assisted interviews, as is standard in large epidemiological studies. The interviews were translated into Hebrew or Russian, with

extensive review and back-translation, including work by tri-lingual staff. Among eligible participants, 1,349 were included, for a response rate of 68.9%.

Interviewers were psychiatric nurses, paramedics or survey interviewers who underwent 6 days of structured training via manuals, self-study exercises, didactic PowerPoint presentations, role-plays, and supervisor certification. Ongoing supervision included periodic field observation, structured review of recorded interviews (approximately one-third were reviewed in this manner), and telephone verification of participation and re-administration of six key sets of items (demographics, drinking and smoking) to 15% of randomly designated respondents.

Sample

Of the respondents, 727 smoked ≥100 cigarettes in their lives and comprised the present sample. Of these, 81.7% (N=594) were male, 28.5% (N=207) were FSU immigrants, 19.3% (N=140) were 21-29 years old, 33.6% (N=244) were 30-44, and 47.2% (N=343) were 45 years or older. This sample of lifetime smokers had the following characteristics: 71.7% (N=521) were current smokers; 88.2% (N=641) were daily smokers at some point; mean age of smoking onset was 16.8 (SD=4.1); mean usual cigarettes/day was 15.1 (SD=11.2); mean cigarettes/day during heaviest smoking was 23.4 (SD=15.6); and mean years smoked was 22.1 (SD=12.2). The overall prevalence of lifetime (53.9%) and current smoking (38.6%) was slightly higher than a similar survey of daily smokers in Israel [38,39], likely due to our inclusion of non-daily smokers and oversampling of FSU immigrants, whose smoking prevalence is higher than other Israelis [37]).

Nicotine Use Disorder Criteria

Lifetime nicotine dependence criteria were measured with the Alcohol Use Disorders and Associated Disabilities Interview Schedule (AUDADIS) [40-42]. Test-retest reliability is excellent for number of lifetime nicotine dependence criteria (ICC=0.76) [41]. Nicotine abuse criteria were assessed with questions parallel to the questions for DSM-IV abuse criteria for alcohol and drugs (Table 1). The nicotine questions underwent pre-testing and adjustment for translation prior to data collection. *Post hoc* review by the DSM-5 SUDs workgroup found the questions to represent the proposed DSM-5 nicotine criteria. The craving question

(unbearable/strong desire [Table 1]) was similar to questions from other epidemiologic and genetic studies [19,43-45].

Statistical Analysis

See Figure 1 for definition of key statistical terms used below and examples of their use in this report.

Internal consistency and Dimensionality (defined in Figure 1). Cronbach's α was used to indicate internal consistency [46] of four criterion sets: 1) dependence, 2) dependence and abuse, 3) dependence and craving, and 4) dependence, abuse, and craving. Eigenvalues (using tetrachoric correlation matrices) were ascertained and factor analyses conducted with MPlus 5.1 (www.statmodel.com) for the four criterion sets. Unidimensionality was confirmed when only one eigenvalue was >1 and/or a factor analysis model with one factor showed adequate model fit by Comparative Fit Index (CFI) or Tucker-Lewis Index (TLI)≥0.95, and root mean square error of approximation (RMSEA)≤0.06 [47]. In cases where two eigenvalues were >1, a 2-factor exploratory factor analysis (EFA) model with geomin rotation was performed to determine whether the additional factor provided a coherent and interpretable construct. For factor interpretation, loadings >0.40 indicated that the item and factor were related. Exploratory rather than confirmatory factor analysis was used since the structure of the proposed criterion set was unknown.

Item Response Theory (IRT) and Total Information (defined in Figure 1). For each of the 4 criterion sets, after establishing unidimensionality, we conducted IRT analysis with MPlus 5.1 using a 2-parameter logistic IRT model for dichotomous traits [48-50]. This model estimates the probability of endorsing a criterion (item) at any latent trait value as a function of two item parameters: discrimination and severity (defined in Figure 1). We generated Item Characteristic Curves (ICC; defined in Figure 1) to display these parameters for each item and Total Information Curves (TIC; defined in Figure 1) to display the information provided by the entire criterion set. We used the total information area index (TIA; Figure 1) [51] to quantify the information provided by the entire criterion set across the trait continuum and to test differences in information between criterion sets. TIA was computed by integrating the TIC across the latent trait range, using the R package ltm [52]; TIA can also be calculated by summing the discrimination parameters over all items [51]. We used the bootstrap method

[53] to test whether models with additional criteria had significantly higher TIA than the dependence-only model (code available at http://www.columbia.edu/~mmw2177/irtprog.html). Five hundred bootstrap samples (re-samples of the observed data with replacement) were taken and the TIA calculated from the estimated IRT parameters in each re-sample. The 2.5 and 97.5 percentiles of the TIA bootstrap distribution indicated the 95% confidence interval (CI) for the TIA. Criterion sets with non-overlapping CIs have significant differences in their total information.

Differential Item and Differential Test Functioning (DIF and DTF; defined in Figure 1). We followed Thissen's method [54] of using likelihood ratio tests (LRTs) to identify criteria with DIF [55,56], using MPlus 5.1 to calculate adjusted χ^2-statistics for the LRTs (http://www.statmodel.com/chidiff.shtml). Using the 11 criterion set, we tested for DIF in the proposed criteria (abuse or craving) by gender, FSU group (ethnicity), and age (21-29 or 30+), because smoking behavior differs between these subgroups [37]. Since the dependence criteria as a group showed no DIF [24], they served as "anchor criteria" to set a common metric for cross-group comparisons [57]. For each proposed criterion, we used an LRT to compare a model with parameters for all criteria (anchor criteria plus the proposed criterion) held equal in each subgroup, to a model where discrimination and severity could vary for the proposed criterion. A significant LRT indicated that allowing the parameters to differ by subgroup improved the fit of the model to the data, suggesting that the criterion functions differently by subgroup (exhibits DIF). For each criterion with DIF, additional LRTs were conducted to determine which parameter (discrimination or severity) functioned differently. The reference groups (female, non-FSU, age 30+) were set to mean=0 and variance=1 for latent trait severity; mean and variance were calculated for the focal group (male, FSU, age 21-29). Similar to others [23,26,57], we corrected for multiple testing using the Benjamini-Hochberg procedure [58,59]. For criteria with significant DIF in severity, we calculated the probability of endorsing the criterion within each group across all values of the latent trait to indicate the magnitude of the DIF effect. We also examined DTF, i.e. if the criteria set as a whole (the "test") functioned differently by subgroup, as determined by individuals in different subgroups endorsing different numbers of criteria at the same underlying trait severity [60,61]. We used R code (available on request) to calculate the average difference in the expected number of criteria for individuals with the same trait severity in different subgroups, using the dependence criteria as anchors. A difference of <1 expected number of criteria

by subgroup indicates no DTF, as differences that small should lead to minimal differential diagnosis of NUD by subgroup.

Results

Dependence

We examined the nicotine dependence criteria to confirm results similar to those previously obtained [22-26]. The prevalence of dependence criteria ranged from 80.7% for tolerance to 16.0% for activities given up (Table 1). Internal consistency, α=0.66, was slightly below the level (0.70) indicating good internal consistency [62]. The first two eigenvalues (3.371, 0.900) supported a 1-factor model, as did model fit indices and factor loadings (Table 2).

IRT parameters (Table 2) indicate that the order of discrimination, from low to high, was: activities given up, withdrawal, quit/control, time spent, tolerance, larger/longer, and physical/psychological. The order of severity, from low to high, was: tolerance, physical/psychological, quit/control, larger/longer, time spent, withdrawal, and activities given up. Test information (Figure 2) was highest at mild-to-moderate severity levels.

Dependence and Abuse

Neglect roles (7.7%) had low prevalence while hazardous use (36.3%) and social/interpersonal (41.3%) had intermediate prevalence (Table 1). Together, the dependence and abuse criteria showed good internal consistency (α=0.71). While two eigenvalues (4.340, 1.045) exceeded 1.0, fit indices indicated that the 1-factor model fit the data (Table 2). Furthermore, the EFA 2-factor model indicated correlation between the factors (0.53, s.e.=0.08) and a second factor consisting of only one criterion (hazardous use), further supporting a 1-factor model.

IRT parameters (Table 2) show that relative to each other, the order of the dependence item severity was unchanged by adding abuse items, with only minor differences in discrimination order. The abuse criteria showed low-to-intermediate discrimination, and medium-to-high severity. TIA for dependence and abuse criteria (13.6, 95%CI=12.6-14.9) significantly exceeded TIA for dependence criteria only (10.0, 95%CI=9.0-

10.7), as shown by non-overlapping CIs generated by the bootstrap method. This information was added mainly at the moderate-to-severe trait levels (Figure 2).

Dependence and Craving

Craving (unbearable/strong desire) had high prevalence (52.4%, Table 1). Adding craving to the dependence criteria produced α=0.69. The first two eigenvalues (3.774, 0.909) indicated a 1-factor model, as did model fit indices (Table 3). Craving showed significant factor loading (0.626), while other loadings differed only slightly (≤0.03) for this model compared to the dependence-only model.

IRT parameters (Table 3) indicate that discriminations and severities differed only slightly (≤0.15) for this model compared to the dependence-only model. The TIC for dependence and craving is shown in Figure 2. The TIA for dependence and craving (11.3, 95%CI=10.4-12.5) did not differ significantly from TIA for dependence-only (10.0, 95% CI 9.0-10.7).

Dependence, Abuse, and Craving

The 11 dependence, abuse, and craving criteria showed good internal consistency (α=0.73). A large first eigenvalue (4.735), a second eigenvalue near 1 (1.094), and model fit indices (Table 3) supported a 1-factor model, with factor loadings differing only slightly (≤0.03) in this model compared to the dependence and abuse model. Similar to results from dependence and abuse, a 2-factor EFA found correlation between the factors (0.59, s.e.=0.11) and only hazardous use on the second factor, which did not support the 2-factor model.

IRT parameters (Table 3, Figure 3) show that discriminations and severities differed only slightly (≤0.08) for this model compared to the dependence and abuse model. TIA for this model (15.0, 95%CI=13.8-16.4) was significantly greater than TIA for the dependence-only model (10.0, 95%CI=9.0-10.7) and for the dependence and craving model (11.3, 95%CI=10.4-12.5), with increased information mostly in the moderate range (Figure 2). However, TIA was not increased significantly compared to the dependence and abuse model (13.6, 95%CI=12.6-14.9).

Differential Functioning

The final model included the eleven proposed DSM-5 NUD criteria (dependence, abuse, and craving). We tested the nicotine abuse and craving criteria for differential item functioning (DIF) using the dependence criteria as anchors, and tested the entire criterion set for differential test functioning (DTF). DIF in the severity parameter was found by FSU group and age, but not gender. Craving was more severe (endorsement less likely at the same latent trait level) in the FSU group (χ^2=23.3, p<0.0001); e.g., at the mean NUD latent trait level, 32.4% of the FSU group endorsed craving vs. 57.6% of the non-FSU group. Hazardous use was less severe in the age 21-29 group (χ^2=21.0, p<0.0001); e.g., at the mean NUD level, 49.5% of the 21-29 group endorsed hazardous use vs. 26.9% of the 30+ group. Results were consistent with different age cutoffs (e.g., 21-44 vs. 45+). For the total criterion set, the average expected difference in number of criteria endorsed was <1 in all subgroups (0.17 for gender; 0.46 for FSU; 0.77 for age).

Supplementary analysis: Current smokers

Results on current nicotine dependence and abuse criteria in current smokers were very similar to lifetime results and are available online (Table S1, Figures S1, S2). The dependence and abuse criteria showed good internal consistency (α=0.70) and unidimensionality (see Table S1 for model fit indices and factor loadings). The order of criterion discrimination and severity (Table S1, Figure S1) was similar to the order for lifetime criteria. TIA for current dependence and abuse criteria (13.5, 95% CI 12.5-15.3) was significantly greater than for dependence only (10.0, 95% CI 9.0-11.3) (Figure S2).

Discussion

This study in an Israeli sample evaluated proposed changes to DSM-5 NUD criteria intended to improve them and align them with DSM-5 criteria for other SUDs. We replicated previous findings that DSM-IV nicotine dependence criteria formed a unidimensional latent trait [22-26]. Importantly, we showed that DSM-IV nicotine dependence, abuse and craving criteria form a unidimensional latent trait, with the criteria intermixed across the latent severity continuum. Adding the proposed abuse criteria increased α to an acceptable level and significantly increased the total information, providing strong support for adding these criteria to the DSM-5 NUD criteria. Craving fit well with the latent trait formed by the nicotine dependence and abuse criteria, but its

addition only slightly increased α and did not significantly increase the total information over the dependence-only or dependence and abuse models. These findings are similar to those on other substances [2-13,19,20] and provide empirical evidence for aligning NUD criteria with other SUD criteria.

Although these results for nicotine dependence, abuse, and craving are consistent with alcohol and drug disorder criteria in terms of unidimensionality, there were differences in discrimination and severity: activities given up showed lower discrimination for nicotine than illicit drugs, tolerance and physical/psychological problems were less severe for nicotine than illicit drugs, and nicotine abuse criteria generally showed higher severity [7,8]. Also, nicotine craving showed lower severity than alcohol craving [19,20]. These differences may point to inherent differences between nicotine and other substances. Alternatively, lower discrimination for activities given up and higher severity for abuse criteria may be due to Israeli tobacco control policies and social norms that, while changing, remain somewhat more permissive than in the U.S. and some European countries [63-65]. Additional research should determine the relationship between tobacco control and criteria discrimination or severity. Nevertheless, differences in specific parameters are less important than the strong overall psychometric similarities between criteria for nicotine and other substances, which support the proposed changes to the NUD criterion set.

While our IRT results support a unidimensional NUD diagnosis that includes nicotine abuse and craving, two criteria exhibited DIF. Hazardous use showed DIF by age, similar to other studies on alcohol and cannabis [6,9,11,66]. Nicotine craving showed DIF by ethnicity, consistent with research documenting DIF for alcohol craving by race [19]. However, the overall NUD diagnosis should be unaffected, since, in all subgroups, the expected number of criteria endorsed should not differ by 1 or more criteria for the same underlying trait severity. Nevertheless, other questions indicating these criteria should be assessed across diverse populations to identify ones that operate more similarly across demographic groups.

We note limitations. Data were collected by self-report, similar to other epidemiologic studies. However, we used a measure (AUDADIS) with good test–retest reliability that is well validated and used in numerous large studies in the U.S. and elsewhere. While lifetime criteria may be affected by memory and recall bias, the

similarity of the dependence results to those using current criteria [22-24,26] suggests that recall problems did not affect the lifetime findings. Further, results for current nicotine dependence and abuse criteria in current smokers were similar to lifetime results. The craving criterion was assessed for the lifetime timeframe only and current craving should be included in future studies. Lastly, this paper provides a psychometric overview of the proposed NUD diagnostic criteria but does not address content, construct, and predictive validity. Validation analyses are underway and will be reported separately.

Study strengths are noted. The standardized measures were used in many previous analyses of proposed DSM-5 criteria. Data collection involved stringent quality assurance procedures. We used state-of-the-art statistical methodology to analyze the data, including a novel method to test whether adding criteria significantly increased total information. Previous findings on the latent structure of alcohol disorder criteria from this Israeli sample [12] were consistent with those found elsewhere, suggesting that this sample is a source of generalizable information. Finally, the sample contributes unique results to the existing literature, as it included data on nicotine abuse and craving criteria as well as nicotine dependence.

In summary, the results support addition of abuse and craving criteria to NUD in DSM-5, standardizing the diagnostic criteria across all substances. Adding abuse criteria with medium-to-high severity addresses the criticism that DSM-IV nicotine dependence provides little information at the severe end of the spectrum [26]. Adding craving addresses what some consider a core aspect of the disorder [28,29]. Evidence for unidimensionality supports development of a NUD severity scale based on the number of criteria endorsed, overcoming concerns about the dichotomous nature of DSM-IV [26,28]. As heavy nicotine use is among the most preventable causes of death worldwide [67], greater attention to nicotine disorders is warranted. A more informative NUD diagnosis that standardizes the criteria with other addictive substances and facilitates severity measurement should help facilitate improved diagnosis and treatment of NUDs.

Acknowledgements

This research was funded by National Institutes of Health grants R01AA013654, R01DA018652, K05AA014223 (Hasin), F31DA026689 (Keyes), K23DA016743 (Aharonovich), and the New York State

Psychiatric Institute (Hasin). We would like to acknowledge the helpful consultations of Rachel Bar-Hamburger, Ph.D., Rina Meyer and Zalman Shoval in collecting the data in Israel, and extensive input from the DSM-5 Substance Disorders Workgroup, who reviewed and approved the questions included in the analyses, and discussed earlier versions of the analyses and text in a series of conference calls.

References

1. American Psychiatric Association. DSM-5 Development, Proposed revision: Substance-use disorder. http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=431, accessed Oct 24 2010 (Archived by WebCite®, http://www.webcitation.org/5tiZ8UbMB).

2. Borges G., Ye Y., Bond J., Cherpitel C. J., Cremonte M., Moskalewicz J. *et al.* The dimensionality of alcohol use disorders and alcohol consumption in a cross-national perspective. *Addiction* 2010; 105: 240-54.

3. Compton W. M., Saha T. D., Conway K. P., Grant B. F. The role of cannabis use within a dimensional approach to cannabis use disorders. *Drug Alcohol Depend* 2009; 100: 221-7.

4. Gelhorn H., Hartman C., Sakai J., Stallings M., Young S., Rhee S. H. *et al.* Toward DSM-V: an item response theory analysis of the diagnostic process for DSM-IV alcohol abuse and dependence in adolescents. *J Am Acad Child Adolesc Psychiatry* 2008; 47: 1329-39.

5. Gillespie N. A., Neale M. C., Prescott C. A., Aggen S. H., Kendler K. S. Factor and item-response analysis DSM-IV criteria for abuse of and dependence on cannabis, cocaine, hallucinogens, sedatives, stimulants and opioids. *Addiction* 2007; 102: 920-30.

6. Kahler C. W., Strong D. R. A Rasch model analysis of DSM-IV Alcohol abuse and dependence items in the National Epidemiological Survey on Alcohol and Related Conditions. *Alcohol Clin Exp Res* 2006; 30: 1165-75.

7. Langenbucher J. W., Labouvie E., Martin C. S., Sanjuan P. M., Bavly L., Kirisci L. *et al.* An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *J Abnorm Psychol* 2004; 113: 72-80.

8. Lynskey M. T., Agrawal A. Psychometric properties of DSM assessments of illicit drug abuse and dependence: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Psychol Med* 2007; 37: 1345-55.

9. Martin C. S., Chung T., Kirisci L., Langenbucher J. W. Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: implications for DSM-V. *J Abnorm Psychol* 2006; 115: 807-14.

10. Proudfoot H., Baillie A. J., Teesson M. The structure of alcohol dependence in the community. *Drug Alcohol Depend* 2006; 81: 21-6.

11. Saha T. D., Chou P. S., Grant B. F. Toward an alcohol use disorder continuum using item response theory: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychol Med* 2006; 36: 931-41.

12. Shmulewitz D., Keyes K., Beseler C., Aharonovich E., Aivadyan C., Spivak B. *et al.* The dimensionality of alcohol use disorders: results from Israel. *Drug Alcohol Depend* 2010; 111: 146-54.

13. Teesson M., Lynskey M., Manor B., Baillie A. The structure of cannabis dependence in the community. *Drug Alcohol Depend* 2002; 68: 255-62.

14. Beseler C. L., Hasin D. S. Cannabis dimensionality: dependence, abuse and consumption. *Addict Behav* 2010; 35: 961-9.

15. Hasin D. S., Liu X., Alderson D., Grant B. F. DSM-IV alcohol dependence: a categorical or dimensional phenotype? *Psychol Med* 2006; 36: 1695-705.

16. World Health Organization. *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research*. Geneva: World Health Organization; 1993.

17. Goldstein R. Z., Volkow N. D. Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. *Am J Psychiatry* 2002; 159: 1642-52.

18. O'Brien C. P., Childress A. R., Ehrman R., Robbins S. J. Conditioning factors in drug abuse: can they explain compulsion? *J Psychopharmacol* 1998; 12: 15-22.

19. Keyes K. M., Krueger R. F., Grant B. F., Hasin D. S. Alcohol craving and the dimensionality of alcohol disorders. *Psychol Med* 2010; in press.

20. Cherpitel C. J., Borges G., Ye Y., Bond J., Cremonte M., Moskalewicz J. *et al.* Performance of a craving criterion in DSM alcohol use disorders. *J Stud Alcohol Drugs* 2010; 71: 674-84.

21. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*. Washington, D.C.: American Psychiatric Association; 2000.

22. McBride O., Strong D. R., Kahler C. W. Exploring the role of a nicotine quantity-frequency use criterion in the classification of nicotine dependence and the stability of a nicotine dependence continuum over time. *Nicotine Tob Res* 2010; 12: 207-16.

23. Rose J. S., Dierker L. C. DSM-IV nicotine dependence symptom characteristics for recent-onset smokers. *Nicotine Tob Res* 2010; 12: 278-86.

24. Saha T. D., Compton W. M., Pulay A. J., Stinson F. S., Ruan W. J., Smith S. M. *et al.* Dimensionality of DSM-IV nicotine dependence in a national sample: An item response theory application. *Drug Alcohol Depend* 2010; 108: 21-8.

25. Strong D. R., Kahler C. W., Ramsey S. E., Brown R. A. Finding order in the DSM-IV nicotine dependence syndrome: a Rasch analysis. *Drug Alcohol Depend* 2003; 72: 151-62.

26. Strong D. R., Kahler C. W., Colby S. M., Griesler P. C., Kandel D. Linking measures of adolescent nicotine dependence to a common latent continuum. *Drug Alcohol Depend* 2009; 99: 296-308.

27. Hughes J. R. Nicotine Withdrawal, Dependence, and Abuse. In: Widiger T. A., American Psychiatric Association Taskforce on DSM-IV, editors. *DSM-IV Sourcebook*. Washington, D.C.: American Psychiatric Association; 1994.

28. Shiffman S., Waters A. J., Hickcox M. The nicotine dependence syndrome scale: a multidimensional measure of nicotine dependence. *Nicotine Tob Res* 2004; 6: 327-48.

29. DiFranza J., Ursprung W. W., Lauzon B., Bancej C., Wellman R. J., Ziedonis D. *et al.* A systematic review of the Diagnostic and Statistical Manual diagnostic criteria for nicotine dependence. *Addict Behav* 2010; 35: 373-82.

30. Hasin D., Rahav G., Meydan J., Neumark Y. The drinking of earlier and more recent Russian immigrants to Israel: comparison to other Israelis. *J Subst Abuse* 1998; 10: 341-53.

31. Hasin D., Aharonovich E., Liu X., Mamman Z., Matseoane K., Carr L.G. *et al.* Alcohol dependence symptoms and alcohol dehydrogenase 2 polymorphism: Israeli Ashkenazis, Sephardics, and recent Russian immigrants. *Alcohol Clin Exp Res* 2002; 26: 1315-21.

32. Hasin D., Aharonovich E., Liu X., Mamman Z., Matseoane K., Carr L. *et al.* Alcohol and ADH2 in Israel: Ashkenazis, Sephardics, and recent Russian immigrants. *Am J Psychiatry* 2002; 159: 1432-4.

33. Rahav G., Hasin D., Paykin A. Drinking patterns of recent Russian immigrants and other Israelis: 1995 national survey results. *Am J Public Health* 1999; 89: 1212-6.

34. Spivak B., Frisch A., Maman Z., Aharonovich E., Alderson D., Carr L. G.*et al.* Effect of ADH1B genotype on alcohol consumption in young Israeli Jews. *Alcohol Clin Exp Res* 2007; 31: 1297-301.

35. Saha T. D., Stinson F. S., Grant B. F. The role of alcohol consumption in future classifications of alcohol use disorders. *Drug Alcohol Depend* 2007; 89: 82-92.

36. Neumark Y. D., Rahav G., Teichman M., Hasin D. Alcohol drinking patterns among Jewish and Arab men and women in Israel. *J Stud Alcohol* 2001; 62: 443-7.

37. Baron-Epel O., Haviv-Messika A., Tamir D., Nitzan-Kaluski D., Green M. Multiethnic differences in smoking in Israel: pooled analysis from three national surveys. *Eur J Public Health* 2004; 14: 384-9.

38. Degenhardt L., Chiu W. T., Sampson N., Kessler R. C., Anthony J. C., Angermeyer M. *et al.* Toward a global view of alcohol, tobacco, cannabis, and cocaine use: findings from the WHO World Mental Health Surveys. *PLoS Med* 2008; 5: e141.

39. Storr C. L., Cheng H., Alonso J., Angermeyer M., Bruffaerts R., de Giralomo. G. *et al.* Smoking estimates from around the world: data from the first 17 participating countries in the World Mental Health Survey Consortium. *Tob Control* 2010; 19: 65-74.

40. Grant B. F., Harford T. C., Dawson D. A., Chou P. S., Pickering R. P. The Alcohol Use Disorder and Associated Disabilities Interview schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. *Drug Alcohol Depend* 1995; 39: 37-44.

41. Grant B. F., Dawson D. A., Stinson F. S., Chou P. S., Kay W., Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug Alcohol Depend* 2003; 71: 7-16.

42. Grant B. F., Stinson F. S., Dawson D. A., Chou S. P., Ruan W. J., Pickering R. P. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Arch Gen Psychiatry* 2004; 61: 361-8.

43. Foroud T., Wetherill L. F., Liang T., Dick D. M., Hesselbrock V., Kramer J. *et al.* Association of alcohol craving with alpha-synuclein (SNCA). *Alcohol Clin Exp Res* 2007; 31: 537-45.

44. Grant B. F. Prevalence and correlates of alcohol use and DSM-IV alcohol dependence in the United States: results of the National Longitudinal Alcohol Epidemiologic Survey. *J Stud Alcohol* 1997; 58: 464-73.

45. Hasin D., McCloud S., Li Q., Endicott J. Cross-system agreement among demographic subgroups: DSM-III, DSM-III-R, DSM-IV and ICD-10 diagnoses of alcohol use disorders. *Drug Alcohol Depend* 1996; 41: 127-35.

46. Cronbach L. J. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16: 297-334.

47. Hu L., Bentler P. M. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* 1999; 6: 1-55.

48. Birnbaum A. Some latent trait models. In: Lord F. M., Novick M. R., editors. *Statistical Theory of Mental Test Scores*. Reading, M.A.: Addison-Wesley; 1968. p. 397-472.

49. Lord F. M., Novick M. R. *Statistical Theories of Mental Test Scores*. Reading, M.A.: Addison-Wesley; 1968.

50. Muthén L. K., Muthén B. O. *Mplus User's Guide*. Fifth edn. Los Angeles, C.A.: Muthén & Muthén; 1998-2007.

51. de Ayala R. J. *The Theory and Practice of Item Response Theory*. New York: Guilford Press; 2009. p. 115.

52. Rizopoulos D. ltm: An R package for latent variable modelling and item response theory analyses. *J Stat Softw* 2006; 17: 1-25.

53. Efron B. The jackknife, the bootstrap and other resampling plans. *CBMS–NSF Regional Conference Series in Applied Mathematics 38*. Philadelphia, PA: SIAM; 1982.

54. Thissen D. Irtlrdif v.20b: software for the computation of the statistics involved in item response theory likelihood ratio tests for differential item functioning. Chapel Hill, NC; University of North Carolina; 2001. Unpublished manuscript.

55. Thissen D., Steinberg L., Wainer H. Use of item response theory in the study of group differences in trace lines. In: Wainer H., Braun H., editors. *Test Validity*. Hillsboro, NJ: Erlbaum; 1988. p. 147-69.

56. Thissen D., Steinberg L., Wainer H. Detection of differential item functioning using the parameters of item response models. In: Holland P. W., Wainer H., editors. *Differential item functioning*. Hillsboro, NJ: Erlbaum; 1993. p. 67-113.

57. Rose J. S., Dierker L. C. An item response theory analysis of nicotine dependence symptoms in recent onset adolescent smokers. *Drug Alcohol Depend* 2010; 110: 70-9.

58. Benjamini Y., Hochberg Y. Controlling false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995; 57: 289-300.

59. Thissen D., Steinberg L., Kuang D. Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. *J Educ Behav Stat* 2002; 27: 77-83.

60. Morales L. S., Flowers C., Gutierrez P., Kleinman M., Teresi J. A. Item and scale differential functioning of the Mini-Mental State Exam assessed using the Differential Item and Test Functioning (DFIT) Framework. *Med Care* 2006; 44: S143-S151.

61. Raju N., van der Linden W., Fleer P. IRT-Based Internal Measures of Differential Functioning of Items and Tests. *Appl Psychol Meas* 1995; 19: 353-68.

62. Nunnally J. O. *Psychometric Theory*. New York: McGraw-Hill; 1978.

63. Hausner A. Israel: empowering the public to enforce smoke-free law. *Tob Control* 2007; 16: 219-20.

64. Rosen L. J., Zucker D. M., Rosen B. J., Connolly G. N. Second-hand smoke levels in Israeli bars, pubs and cafes before and after implementation of smoke-free legislation. *Eur J Public Health* 2011; 21: 15-20.

65. Shafey O., Dolwick S., Guindon G. E. Tobacco Control Country Profiles. http://www.who.int/tobacco/global_data/en/, accessed Feb 8 2011 (Archived by WebCite®, http://www webcitation org/5woGhee1W).

66. Harford T. C., Yi H. Y., Faden V. B., Chen C. M. The dimensionality of DSM-IV alcohol use disorders among adolescent and adult drinkers and symptom patterns by age, gender, and race/ethnicity. *Alcohol Clin Exp Res* 2009; 33: 868-78.

67. Lopez A. D., Mathers C. D., Ezzati M., Jamison D. T., Murray C. J. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006; 367: 1747-57.

Table 1. Prevalence of and questions used to assess lifetime nicotine criteria

Criterion	%	N	In your ENTIRE LIFE did you EVER…
Dependence[a]			
Tolerance	80.7	587	Find that you had to use much more tobacco than you once did to get the effect you wanted? OR Increase your smoking by at least 50 percent? OR Find that the first cigarette of the day had a much stronger effect than it used to? OR Find that you no longer got dizzy or nauseous from smoking?[b]
Withdrawal	24.5	178	Withdrawal syndrome[c] (4 or more symptoms that cause distress or dysfunction) OR smoke to avoid having any of these symptoms?
Larger/ longer	55.3	402	Have a period when you often smoked more than you intended to?
Quit/control	61.4	446	Want to stop or cut down on your smoking, regardless of whether or not you actually tried, more than once? OR Find that you were unable to stop or cut down on your smoking, more than once?
Time spent	38.5	280	Find yourself chain smoking?
Activities given up	16.0	116	Give up or cut down on activities that you were interested in or that gave you pleasure or that were important to you – like associating with friends or relatives or attending social activities because smoking was not permitted at the activity?
Physical/ psychological	79.0	574	Continue to smoke even though you knew it was causing you a health problem or making a health problem worse? OR Continue to smoke even though made you jittery, anxious, or depressed?
Abuse[d]			
Neglect roles	07.7	56	Find that your smoking interfered with taking care of your work, school work, or work at home?
Hazardous use	36.3	264	Smoke in a situation that increased your chances of getting hurt – like smoking in bed or smoking around flammable chemicals?
Social/ interpersonal	41.3	300	Continue to smoke even though it made other people like family members angry or unhappy?
Craving			
Unbearable/ strong desire	52.4	381	When you have run out of cigarettes, do you always or often find it almost unbearable until you can get them? OR Do you always or often get a strong desire to smoke when you haven't smoked for a while?

[a] DSM-IV Text Revision [21], p. 197, [b] p. 264, [c] p. 266, [d] p. 199

Table 2. Factor Analysis and Item Response Theory Analysis of Nicotine Dependence and Abuse Criteria in Lifetime Smokers, N=727

	Factor Loadings				Item Response Theory parameters			
Lifetime Criteria	Dependence	Dependence and Abuse			Dependence		Dependence and Abuse	
	1-factor	1-factor	2-factors		Discrimination (s.e.)	Severity (s.e.)	Discrimination (s.e.)	Severity (s.e.)
Nicotine Dependence								
Tolerance	0.658	0.657	0.650	0.024	1.499 (0.22)	-1.319 (0.14)	1.505 (0.20)	-1.315 (0.13)
Withdrawal	0.596	0.598	0.595	0.019	1.326 (0.19)	1.121 (0.13)	1.329 (0.17)	1.117 (0.12)
Larger/longer	0.705	0.648	0.683	-0.031	1.680 (0.20)	-0.185 (0.07)	1.452 (0.17)	-0.200 (0.07)
Quit/control	0.622	0.600	0.717	-0.138	1.328 (0.17)	-0.462 (0.08)	1.250(0.15)	-0.479 (0.09)
Time spent	0.631	0.673	0.480	0.274	1.392 (0.17)	0.461 (0.08)	1.567 (0.19)	0.431 (0.08)
Activities given up	0.458	0.471	0.443	0.049	0.928 (0.15)	2.073 (0.30)	0.976 (0.15)	1.997 (0.26)
Physical/psychological	0.726	0.745	0.713	0.059	1.792 (0.26)	-1.116 (0.11)	1.909 (0.26)	-1.083 (0.10)
Nicotine Abuse								
Neglect roles		0.567	0.699	-0.158			1.335 (0.14)	2.354 (0.29)
Hazardous use		0.600	0.002	0.974			1.263 (0.09)	0.582 (0.09)
Social/interpersonal		0.531	0.329	0.288			1.037 (0.08)	0.417 (0.10)
Model Fit Indices								
Comparative Fit Index (CFI)	0.995	0.970	0.991					
Tucker-Lewis Index (TLI)	0.994	0.972	0.989					
Root mean square error of approximation (RMSEA)	0.019	0.038	0.024					
Akaike Information Criterion (AIC)					5388.903		7483.770	
Bayesian Information Criterion (BIC)					5453.148		7575.549	
Sample-size corrected BIC (SS-BIC)					5408.693		7512.043	

Table 3. Factor Analysis and Item Response Theory Analysis of Nicotine Dependence, Abuse, and Craving Criteria in Lifetime Smokers, N=727

	Factor Loadings		Item Response Theory parameters			
Lifetime Criteria	**Dependence and Craving**	**Dependence, Abuse, and Craving**	**Dependence and Craving**		**Dependence, Abuse, and Craving**	
Nicotine Dependence	**1-factor model**	**1-factor model**	**Discrimination (s.e.)**	**Severity (s.e.)**	**Discrimination (s.e.)**	**Severity (s.e.)**
Tolerance	0.662	0.657	1.494 (0.21)	-1.321 (0.13)	1.465 (0.19)	-1.335 (0.13)
Withdrawal	0.612	0.607	1.414 (0.19)	1.081 (0.12)	1.394 (0.17)	1.088 (0.12)
Larger/longer	0.678	0.634	1.535 (0.18)	-0.190 (0.07)	1.372 (0.15)	-0.202 (0.08)
Quit/control	0.627	0.603	1.350 (0.16)	-0.454 (0.08)	1.256 (0.15)	-0.474 (0.09)
Time spent	0.630	0.664	1.389 (0.17)	0.464 (0.08)	1.544 (0.19)	0.439 (0.08)
Activities given up	0.472	0.477	0.995 (0.16)	1.972 (0.26)	1.022 (0.15)	1.930 (0.24)
Physical/psychological	0.729	0.740	1.780 (0.23)	-1.118 (0.10)	1.822 (0.22)	-1.104 (0.10)
Nicotine Abuse						
Neglect roles		0.534			1.261 (0.22)	2.437 (0.32)
Hazardous use		0.615			1.346 (0.15)	0.564 (0.09)
Social/interpersonal		0.547			1.095 (0.14)	0.405 (0.09)
Nicotine Craving						
Unbearable/strong desire	0.626	0.643	1.335 (0.16)	-0.091 (0.08)	1.426 (0.17)	-0.085 (0.07)
Model Fit Indices						
CFI	0.997	0.973				
TLI	0.997	0.977				
RMSEA	0.013	0.035				
AIC			6265.929		8338.017	
BIC			6339.352		8438.974	
SS-BIC			6288.547		8369.117	

Figure 1. Definition of Key Statistical Terms

Figure 2. Total Information Curves for Nicotine Dependence, Abuse, and Craving Criteria in Lifetime Smokers, N=727

Figure 3. Item Characteristic Curves for Nicotine Dependence, Abuse, and Craving Criteria in Lifetime Smokers, N=727

Figure 1.

Term	Definition	Use in the present analysis
Item	An observed variable for each respondent; here, a diagnostic criterion.	11 nicotine items: 7 nicotine dependence criteria, 3 abuse criteria, and craving.
Internal consistency	A measure indicating the correlation of all items in a set.	Used to determine if the nicotine items are correlated with each other.
Latent variable	A trait that is not directly observable but that can be inferred from a pattern of relationships among observed variables.	Nicotine use disorder (NUD) is inferred by the observed relationships between dependence, abuse and craving items.
Dimensionality	The number of latent variables measured by a set of items. Dimensionality is assessed by eigenvalues and factor analysis. An item set that measures one latent variable is unidimensional.	Examination of whether the 7 dependence criteria and 3 abuse criteria represent two dimensions (nicotine dependence, nicotine "abuse") or if the 10 criteria plus craving measure a single dimension (NUD).
Eigenvalues	Eigenvalues quantify how much variability in the items is associated with each dimension (latent trait). The number of dimensions underlying an item set is equal to the number of eigenvalues >1.	The dependence criteria resulted in 1 eigenvalue >1, suggesting a 1-factor model for factor analysis. The dependence and abuse criteria resulted in a second eigenvalue ~ 1; a 2-factor model was also considered.
Factor analysis (FA)	FA examines the relationship between each item and the underlying latent variables (reported as factor loadings) and determines how well the model fits the data. FA is often conducted for models with different number of factors to see which model fits best and provides interpretable factors.	For the dependence and abuse criteria, we tested both 1-factor and 2-factor models, and used model fit indices, factor correlation, and the lack of a meaningful second factor to select the 1-factor model.
Model fit indices	Different indices quantify model fit. These include comparative fit indices (e.g., CFI, TLI), which quantify fit of the hypothesized model to the null model (no relationship between the items), and absolute fit indices, such as RMSEA.	For the dependence and abuse items, both the 1-factor and 2-factor model resulted in model fit indices indicating good model fit. The more parsimonious 1-factor model is preferred.
Item response theory (IRT) analysis	For unidimensional item sets, IRT analysis estimates the probability of endorsing each item along the latent variable continuum as a function of two parameters, severity and discrimination	IRT analysis was used to learn more about the proposed DSM-5 nicotine criteria.
Item Characteristic Curves (ICC)	ICCs visually represent the probability of endorsing each item (y-axis) across the latent variable severity continuum (x-axis)**.**	Figure 3 shows ICCs for the 11 nicotine criteria. On the x-axis, the left side is mild and the right side is severe.
Severity	Severity represents the location along the underlying latent trait continuum (x-axis) where the item has a 50% probability of being endorsed (y-axis). A rare item indicates a more severe trait, as only those with a severe trait will manifest a rare item.	Respondents with all NUD severity values will manifest criteria of mild severity (high prevalence, e.g., trying to quit/cut down), while only respondents with severe NUD will manifest criteria of high severity (low prevalence, e.g., important activities given up to smoke).
Discrimination	Discrimination, the slope of the ICC at the item's severity, shows how the probability of endorsing an item changes across the severity continuum. Item information is directly proportional to discrimination; steeper slopes indicate higher information while flatter slopes indicate lower information.	When respondents endorse high-discrimination items, information is provided about NUD severity since there is a specific NUD severity range at which this item is likely to be endorsed. Endorsing low-discrimination items does not provide this information as the item's endorsement probability is similar across NUD severity.
Total information curve (TIC)	A TIC shows how precisely the item set measures the underlying trait at *each* value along the severity continuum. A "flat" TIC indicates equal information across all severities; a "peaked" TIC indicates more information about the trait at a specific severity.	Figure 2 suggests that the dependence, abuse, and craving criteria are most informative for intermediate severity NUDs because the TIC peaks in the middle of the latent trait severity.
Total information area index (TIA)	The total information provided by the item set across the *entire* trait continuum. TIAs from different item sets can be tested to indicate if the amount of total information in one item set differs from another.	Tests of differences in the TIAs showed that adding craving to the dependence criteria did not significantly increase the total information, but adding the abuse criteria did.
Differential item functioning (DIF)	DIF indicates that the item parameters (discrimination or severity) vary across population subgroups. Ideally, items should function the same across subgroups.	Craving functions differently by FSU group; at the same levels of NUD severity, individuals from the FSU group were less likely to endorse craving than from the non-FSU group.
Differential test functioning (DTF)	DTF occurs when respondents in population subgroups, at the same latent trait severity, endorse a different number of criteria in total. DTF could lead to differential diagnosis in subgroups.	At each NUD severity, population subgroups were expected to endorse close to the same number of total criteria, suggesting that the criteria set measures NUD equivalently in all the subgroups (no DTF).

Figure 2.

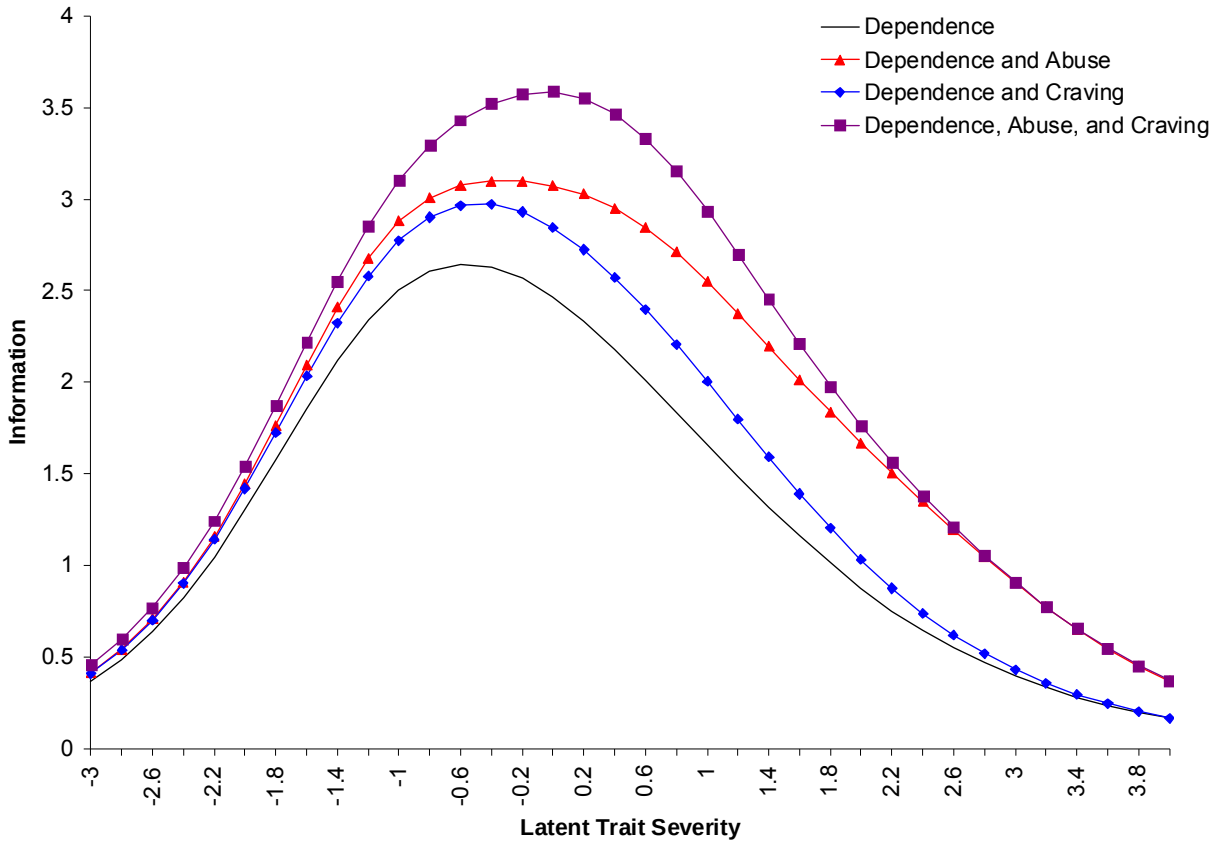

Dependence
Dependence and Abuse
Dependence and Craving
Dependence, Abuse, and Craving
Information
Latent Trait Severity
4
3.5
3
2.5
2
1.5
1
0.5
0
-3
-2.6
-2.2
-1.8
-1.4
-1
-0.6
-0.2
0.2
0.6
1
1.4
1.8
2.2
2.6
3
3.4
3.8

Figure 3

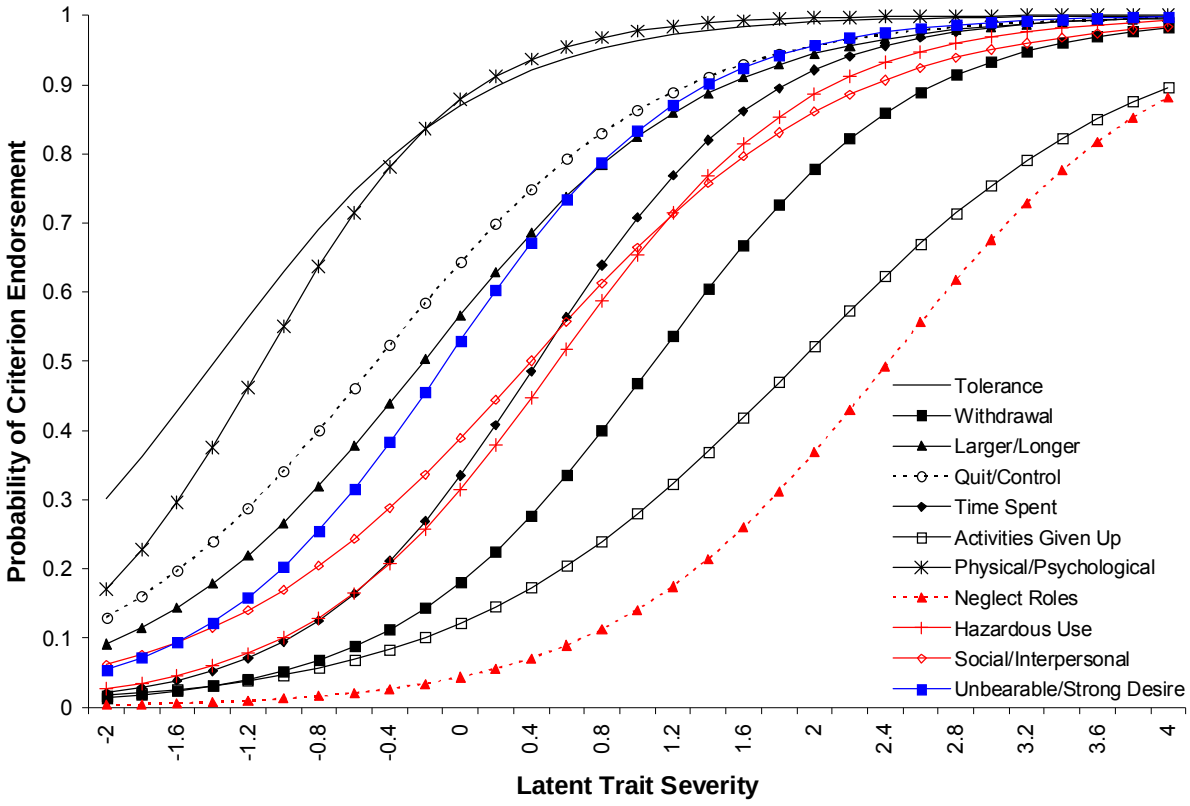

Probability of Criterion Endorsement
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-2
-1.6
-1.2
-0.8
-0.4
0
0.4
0.8
1.2
1.6
2
2.4
2.8
3.2
3.6
4
Latent Trait Severity
Tolerance
Withdrawal
Larger/Longer
Quit/Control
Time Spent
Activities Given Up
Physical/Psychological
Neglect Roles
Hazardous Use
Social/Interpersonal
Unbearable/Strong Desire

Table S1. Prevalence, Factor Analysis, and Item Response Theory Analysis of Nicotine Dependence and Abuse Criteria in Current Smokers, N=521

		Factor Loadings				Item Response Theory parameters			
Lifetime Criteria		**Dependence**	**Dependence and Abuse**			**Dependence**		**Dependence and Abuse**	
Nicotine Dependence	**% (N)**	**1-factor**	**1-factor**	**2-factors**		**Discrimination (s.e.)**	**Severity (s.e.)**	**Discrimination (s.e.)**	**Severity (s.e.)**
Tolerance	62.6 (326)	0.712	0.711	0.253	0.527	1.731 (0.26)	-0.450 (0.09)	1.693 (0.24)	-0.454 (0.09)
Withdrawal	23.8 (124)	0.517	0.547	0.507	0.124	1.013 (0.18)	1.379 (0.21)	1.105 (0.18)	1.302 (0.18)
Larger/longer	38.2 (199)	0.790	0.768	0.353	0.496	2.208 (0.37)	0.385 (0.08)	2.025 (0.29)	0.402 (0.08)
Quit/control	30.1 (157)	0.506	0.499	0.716	-0.102	0.979 (0.17)	1.028 (0.17)	0.955 (0.16)	1.047 (0.17)
Time spent	24.4 (127)	0.655	0.675	0.003	0.726	1.520 (0.24)	1.046 (0.13)	1.603 (0.25)	1.021 (0.12)
Activities given up	10.0 (52)	0.464	0.453	0.168	0.329	0.940 (0.20)	2.686 (0.42)	0.947 (0.20)	2.669 (0.46)
Physical/ psychological	68.9 (359)	0.689	0.672	0.421	0.335	1.584 (0.24)	-0.724 (0.10)	1.528 (0.22)	-0. 739 (0.11)
Nicotine Abuse									
Neglect roles	06.5 (34)		0.606	0.692	0.017			1.525 (0.34)	2.312 (0.34)
Hazardous use	24.0 (125)		0.555	-0.205	0.795			1.144 (0.19)	1.260 (0.17)
Social/ interpersonal	32.4 (169)		0.497	0.050	0.486			0.947 (0.15)	0.921 (0.16)
Model Fit Indices									
Comparative Fit Index (CFI)		0.981	0.974	0.992					
Tucker-Lewis Index (TLI)		0.976	0.973	0.989					
Root mean square error of approximation (RMSEA)		0.039	0.036	0.023					
Akaike Information Criterion (AIC)						3829.351		5175.137	
Bayesian Information Criterion (BIC)						3888.931		5260.252	
Sample-size corrected BIC (SS-BIC)						3844.492		5196.767	

Figure S1: Item Characteristic Curves for Nicotine Dependence and Abuse Criteria in Current Smokers, N=521

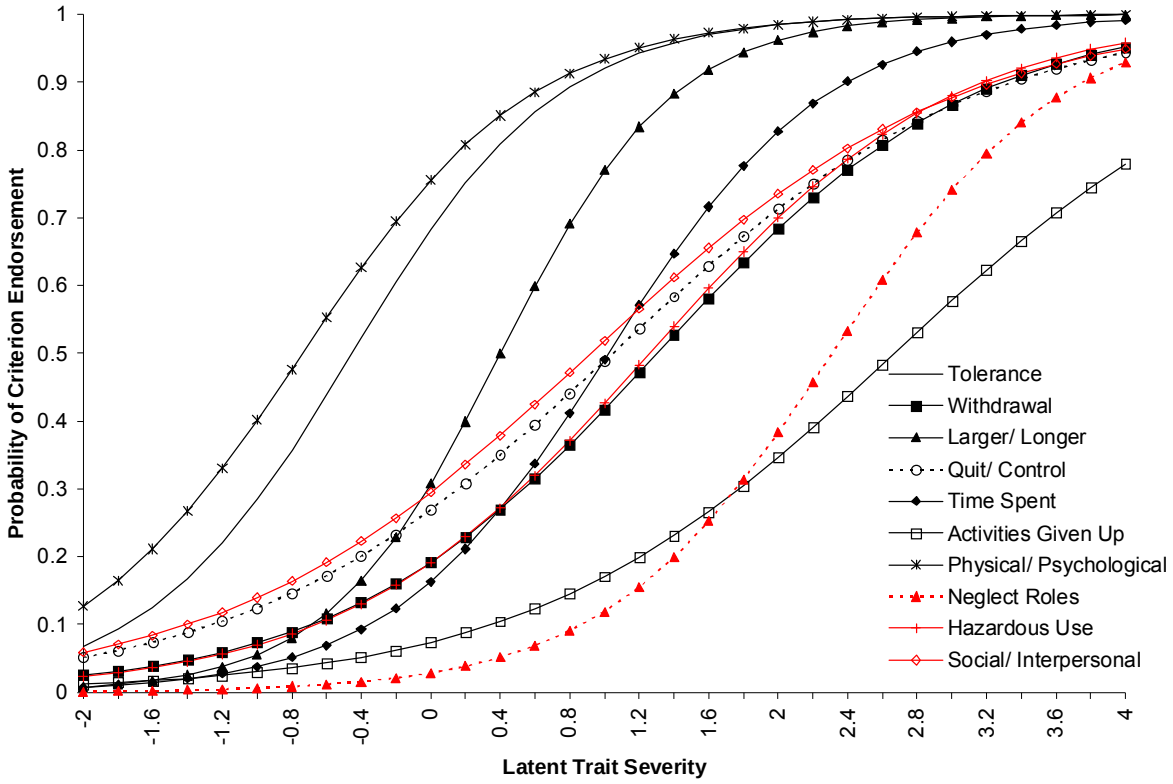
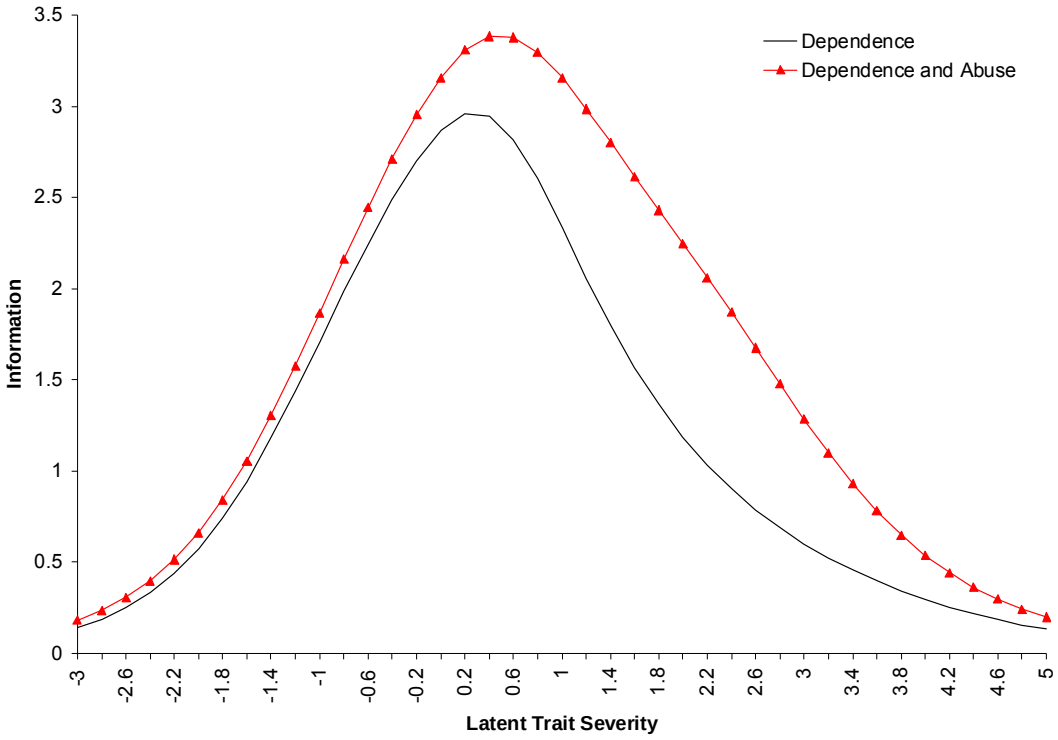
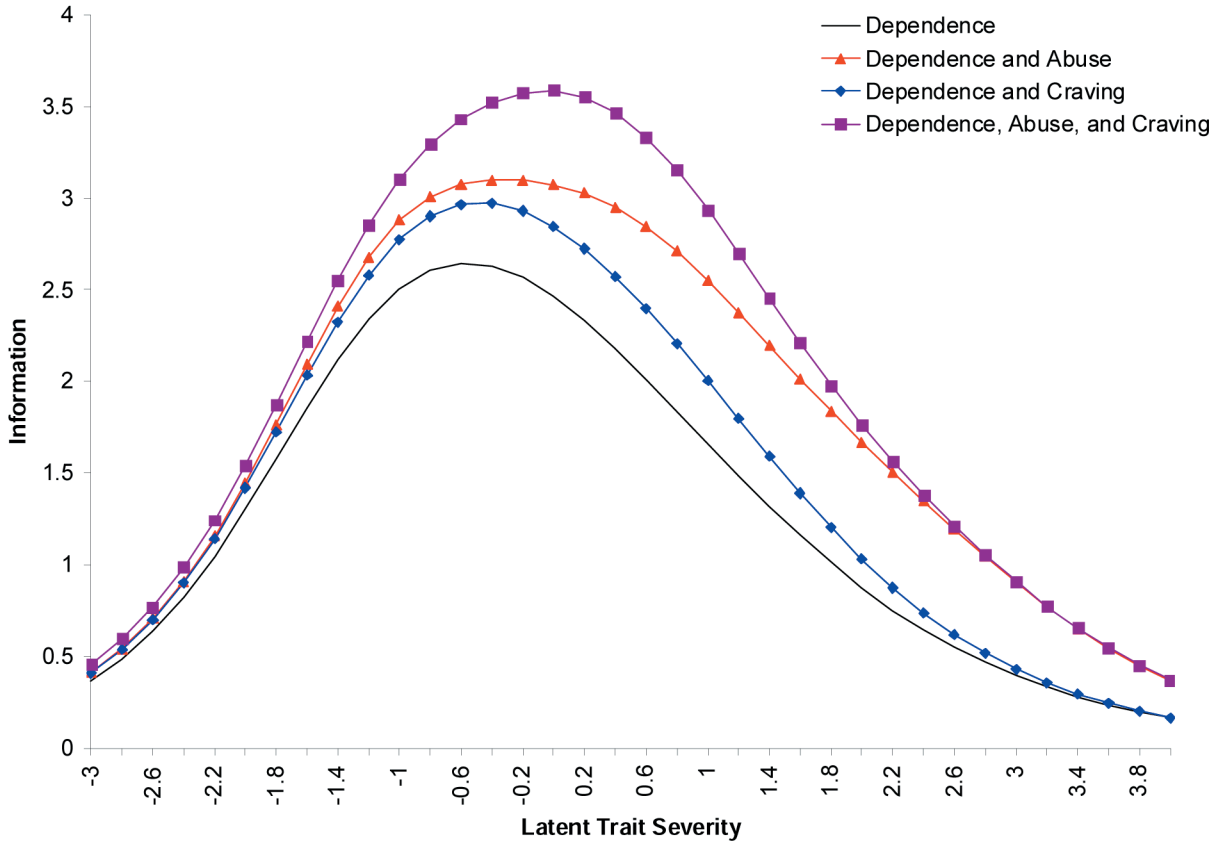


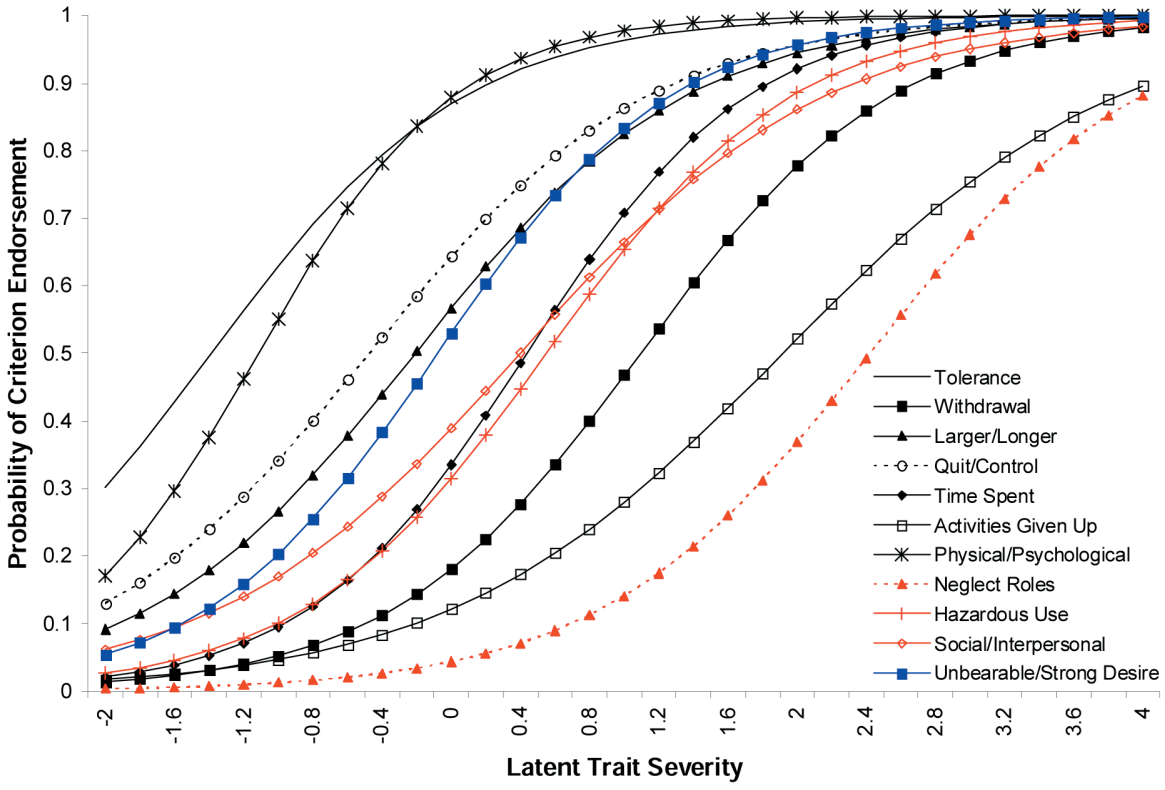
Figure S2: Total Information Curves for Nicotine Dependence and Abuse Criteria in Current Smokers, N=521



Term	Definition	Use in the present analysis
Item	An observed variable for each respondent; here, a diagnostic criterion.	11 nicotine items: 7 nicotine dependence criteria, 3 abuse criteria, and craving.
Internal consistency	A measure indicating the correlation of all items in a set.	Used to determine if the nicotine items are correlated with each other.
Latent variable	A trait that is not directly observable but that can be inferred from a pattern of relationships among observed variables.	Nicotine use disorder (NUD) is inferred by the observed relationships between dependence, abuse and craving items.
Dimensionality	The number of latent variables measured by a set of items. Dimensionality is assessed by eigenvalues and factor analysis. An item set that measures one latent variable is unidimensional.	Examination of whether the 7 dependence criteria and 3 abuse criteria represent two dimensions (nicotine dependence, nicotine "abuse") or if the 10 criteria plus craving measure a single dimension (NUD).
Eigenvalues	Eigenvalues quantify how much variability in the items is associated with each dimension (latent trait). The number of dimensions underlying an item set is equal to the number of eigenvalues >1.	The dependence criteria resulted in 1 eigenvalue >1, suggesting a 1-factor model for factor analysis. The dependence and abuse criteria resulted in a second eigenvalue ~ 1; a 2-factor model was also considered.
Factor analysis (FA)	FA examines the relationship between each item and the underlying latent variables (reported as factor loadings) and determines how well the model fits the data. FA is often conducted for models with different number of factors to see which model fits best and provides interpretable factors.	For the dependence and abuse criteria, we tested both 1-factor and 2-factor models, and used model fit indices, factor correlation, and the lack of a meaningful second factor to select the 1-factor model.
Model fit indices	Different indices quantify model fit. These include comparative fit indices (e.g., CFI, TLI), which quantify fit of the hypothesized model to the null model (no relationship between the items), and absolute fit indices, such as RMSEA.	For the dependence and abuse items, both the 1-factor and 2-factor model resulted in model fit indices indicating good model fit. The more parsimonious 1-factor model is preferred.
Item response theory (IRT) analysis	For unidimensional item sets, IRT analysis estimates the probability of endorsing each item along the latent variable continuum as a function of two parameters, severity and discrimination	IRT analysis was used to learn more about the proposed DSM-5 nicotine criteria.
Item Characteristic Curves (ICC)	ICCs visually represent the probability of endorsing each item (y-axis) across the latent variable severity continuum (x-axis).	Figure 3 shows ICCs for the 11 nicotine criteria. On the x-axis, the left side is mild and the right side is severe.
Severity	Severity represents the location along the underlying latent trait continuum (x-axis) where the item has a 50% probability of being endorsed (y-axis). A rare item indicates a more severe trait, as only those with a severe trait will manifest a rare item.	Respondents with all NUD severity values will manifest criteria of mild severity (high prevalence, e.g., trying to quit/cut down), while only respondents with severe NUD will manifest criteria of high severity (low prevalence, e.g., important activities given up to smoke).
Discrimination	Discrimination, the slope of the ICC at the item's severity, shows how the probability of endorsing an item changes across the severity continuum. Item information is directly proportional to discrimination; steeper slopes indicate higher information while flatter slopes indicate lower information.	When respondents endorse high-discrimination items, information is provided about NUD severity since there is a specific NUD severity range at which this item is likely to be endorsed. Endorsing low-discrimination items does not provide this information as the item's endorsement probability is similar across NUD severity.
Total information curve (TIC)	A TIC shows how precisely the item set measures the underlying trait at *each* value along the severity continuum. A "flat" TIC indicates equal information across all severities; a "peaked" TIC indicates more information about the trait at a specific severity.	Figure 2 suggests that the dependence, abuse, and craving criteria are most informative for intermediate severity NUDs because the TIC peaks in the middle of the latent trait severity.
Total information area index (TIA)	The total information provided by the item set across the *entire* trait continuum. TIAs from different item sets can be tested to indicate if the amount of total information in one item set differs from another.	Tests of differences in the TIAs showed that adding craving to the dependence criteria did not significantly increase the total information, but adding the abuse criteria did.
Differential item functioning (DIF)	DIF indicates that the item parameters (discrimination or severity) vary across population subgroups. Ideally, items should function the same across subgroups.	Craving functions differently by FSU group; at the same levels of NUD severity, individuals from the FSU group were less likely to endorse craving than from the non-FSU group.
Differential test functioning (DTF)	DTF occurs when respondents in population subgroups, at the same latent trait severity, endorse a different number of criteria in total. DTF could lead to differential diagnosis in subgroups.	At each NUD severity, population subgroups were expected to endorse close to the same number of total criteria, suggesting that the criteria set measures NUD equivalently in all the subgroups (no DTF).


Dependence
Dependence and Abuse
Dependence and Craving
Dependence, Abuse, and Craving
Information
Latent Trait Severity
0
0.5
1
1.5
2
2.5
3
3.5
4
-3
-2.6
-2.2
-1.8
-1.4
-1
-0.6
-0.2
0.2
0.6
1
1.4
1.8
2.2
2.6
3
3.4
3.8

f2


Probability of Criterion Endorsement
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-2
-1.6
-1.2
-0.8
-0.4
0
0.4
0.8
1.2
1.6
2
2.4
2.8
3.2
3.6
4
Latent Trait Severity
Tolerance
Withdrawal
Larger/Longer
Quit/Control
Time Spent
Activities Given Up
Physical/Psychological
Neglect Roles
Hazardous Use
Social/Interpersonal
Unbearable/Strong Desire

f3