An item response theory analysis of DSM-IV criteria for hallucinogen abuse and dependence in adolescents

Li-Tzy Wu[a,*], **Jeng-Jong Pan**[b], **Chongming Yang**[c], **Bryce B. Reeve**[d], and **Dan G. Blazer**[a]

[a]Department of Psychiatry and Behavioral Sciences, School of Medicine, Duke Clinical Research Institute, Duke University Medical Center, Durham, NC 27710, USA

[b]Veterans Health Administration, Washington, DC, 20420, USA

[c]Social Science Research Institute, Duke University, Durham, NC 27705, USA

[d]National Cancer Institute, Bethesda, MD 20814, USA

Abstract

Aim—This study applied both item response theory (IRT) and multiple indicators–multiple causes (MIMIC) methods to evaluate item-level psychometric properties of diagnostic questions for hallucinogen use disorders (HUDs), differential item functioning (DIF), and predictors of latent HUD.

Methods—Data were drawn from 2004–2006 National Surveys on Drug Use and Health. Analyses were based on 1548 past-year hallucinogen users aged 12–17 years. Substance use and symptoms were assessed by audio computer-assisted self-interviewing methods.

Results—Abuse and dependence criteria empirically were arrayed along a single continuum of severity. All abuse criteria indicated middle-to-high severity on the IRT-defined HUD continuum, while dependence criteria captured a wider range from the lowest (*tolerance and time spent*) to the highest (*taking larger amounts and inability to cut down*) severity levels. There was indication of DIF by hallucinogen users' age, gender, race/ethnicity, and ecstasy use status. Adjusting for DIF, ecstasy users (vs. non-ecstasy hallucinogen users), females (vs. males), and whites (vs. Hispanics) exhibited increased odds of HUD.

Conclusions—Symptoms of hallucinogen abuse and dependence empirically do not reflect two discrete conditions in adolescents. Trends and problems related to hallucinogen use among girls and whites should be examined further to inform the designs of effective gender-appropriate and culturally sensitive prevention programs.

Keywords

Differential item functioning; Ecstasy; Hallucinogen use disorder; Item response theory; MDMA

*1**Corresponding author:** Li-Tzy Wu, Sc.D., Department of Psychiatry and Behavioral Sciences, School of Medicine, Duke Clinical Research Institute, Duke University Medical Center, Box 3419, Durham, NC 27710, USA; litzy.wu@duke.edu, litzywu@yahoo.com; tel.: 919-668-6067; fax: 919-668-5418.

Declaration of interest
The authors report no conflict of interest.

1. Introduction

According to Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV, hallucinogens comprise hallucinogenic substances (e.g., lysergic acid diethylamide [LSD], peyote, mescaline, and psilocybin) and ecstasy (MDMA: 3,4-methylenedioxymethamphetamine) (American Psychiatric Association [APA], 2000). Ecstasy, an hallucinogen stimulant, is among the most prevalent emerging drugs used by adolescents (Wu et al., 2006), but because its use became prevalent after publication of DSM-IV (APA, 1994), little information exists concerning its classification (Crowley, 2006).

Generally speaking, psychometric data for hallucinogen use disorder (HUD) criteria in adolescents are lacking. According to DSM-IV (APA, 2000), only individuals not meeting criteria for dependence are assigned an abuse diagnosis, and dependence is generally considered to be more severe than abuse. However, studies applying item response theory (IRT) modeling to alcohol, marijuana, and opioid use disorders have suggested that criteria for dependence and abuse empirically constitute a single dimensional construct, not two discrete conditions (Hartman et al., 2008; Gelhorn et al., 2008; Martin et al., 2006; Wu et al., 2009a).

Another area with critical implications for DSM-V concerns differential item functioning (DIF) in the assessment and analysis of self-reported diagnoses (Wu et al., 2009a; Wu et al., 2009b). DIF occurs when, equating the overall level of a measured condition, drug users respond differentially to diagnostic questions across groups (Wu et al., 2009a). DIF may distort the group comparison of a diagnosis because observed differences may be affected by different interpretation and reporting of symptoms across subgroups (different item threshold), not true differences (Wu et al., 2009a). Hence, evaluating DIF for diagnostic criteria is fundamental to developing less biased estimates for a diagnosis.

Studies typically rely on IRT methods alone to examine psychometric properties of diagnostic questions (Hartman et al., 2008; Gelhorn et al., 2008; Martin et al., 2006). This study applies both IRT and MIMIC (multiple indicators–multiple causes) to enhance the investigation. MIMIC modeling incorporates the measurement component of diagnostic items with a structural equation component into an integrated model (Wu et al., 2009a). It not only detects DIF of diagnostic criteria for demographic variables while controlling for the overall level of HUD, but also examines predictors of HUD. In response to the need for empirical data regarding HUD diagnoses for DSM-V (APA, 2009), we apply IRT analysis to evaluate item-level psychometric properties (unidimensionality, item severity, and item discrimination) of HUD criteria and use MIMIC procedures to identify predictors of HUD while taking into account both measurement errors (DIF) and demographic variables. We also investigate whether ecstasy users have increased odds of HUD as suggested by recent studies (Stone et al., 2006).

2. Methods

2.1. Study sample

This study examined public-use data from adolescents in the 2004–2006 National Surveys on Drug Use and Health (NSDUH) (Substance Abuse and Mental Health Services Administration [SAMHSA], 2005, 2006, 2007). NSDUH is an ongoing survey providing population estimates of substance use and disorders in the U.S. Its sampling frame covers approximately 98% of the total population aged ≥12 years and uses multistage area probability sampling methods to select a representative sample of the civilian non-institutionalized population. Participants were interviewed privately at their places of residence. NSDUH employs a combination of computer-assisted personal interviewing and audio computer-assisted self-interviewing

methodologies to increase the accuracy of participants' reports of substance use behaviors (Turner et al., 1998). Weighted response rates for adolescents ranged from 85–88%.

There was no yearly variation in the distributions of adolescents' age, gender, and race/ethnicity. In this combined sample (N=55,286), 49% were female and 39% were nonwhite. This study focused on past-year hallucinogen users (2.7% of all adolescents), and three years of the data were combined to increase statistical power. Of all hallucinogen users (N=1548), 45% used ecstasy.

2.2. Study variables

Socioeconomic variables included age, gender, race/ethnicity, and total family income (Wu et al., 2006). The survey assessed *hallucinogen use* (ecstasy/MDMA, LSD, phencyclidine, peyote, mescaline, and psilocybin), age of first hallucinogen use, and the *number of days* using hallucinogens in the past year (Wu et al., 2008). Consistent with DSM-IV's definition for hallucinogens (APA, 2000), we excluded users of phencyclidine only (n=37). *Hallucinogen use* was categorized into *ecstasy use* (regardless of use of other hallucinogens) and *non-ecstasy hallucinogen use* (Wu et al., 2008).

Past-year DSM-IV HUDs were assessed by standardized questions (APA, 2000; Wu et al., 2008). *Abuse* criteria included: (A1) serious problems at home, work, or school; (A2) regular consumption that put the user in physical danger; (A3) repeated use that led to trouble with the law; and (A4) problems with family or friends caused by continued use. Six *dependence* criteria were assessed: (D1) tolerance; (D2) more frequent use than intended or inability to maintain limits on use; (D3) inability to reduce or stop use; (D4) spending a great deal of time over a period of a month using or getting over the effects of the hallucinogens; (D5) reduced involvement or participation in important activities because of use; (D6) continued use despite related problems with emotions, nerves, or mental or physical health. "Withdrawal" was not specified as necessary for hallucinogen dependence by DSM-IV (APA, 2000) and was not assessed.

2.3. Data analysis

Factor, IRT, and MIMIC analyses with complex survey procedures (Muthén & Muthén, 2007) were conducted among past-year hallucinogen users (N=1548). Factor analysis of binary data examined IRT's assumption of unidimensionality (Embretson & Reise, 2000). The scree plot of eigenvalues (Cattell, 1996) and the ratio of the first to the second eigenvalue assessed evidence of unidimensionality for IRT modeling (Wu et al., 2009a).

A two-parameter normal ogive IRT model examined the relationship between hallucinogen users' response to each item and their level on the IRT-defined latent HUD severity (measured by the 10 criteria), which is described by a monotonically increasing S-shaped *item characteristic curve* (ICC) (Embretson & Reise, 2000). An ICC is characterized by item severity and discrimination parameters. An item *severity* (threshold) parameter indicates the position of the ICC in relation to the latent continuum. A severity parameter represents the location along the latent HUD continuum, in which a user has a 50% probability for endorsing an item. Higher severity values indicate that the item is associated with a high HUD severity. A *discrimination* parameter measures the degree of precision with which an item distinguishes among hallucinogen users with levels of HUD severity above and below the item's severity levels. A low discrimination value suggests the item is unrelated to the underlying construct or is poorly defined (Baker, 2001).

Finally, MIMIC modeling examined DIF (by gender, age, race/ethnicity, and ecstasy use status) and predictors of HUD. It includes the measurement part of the 10 HUD criteria (the latent

HUD), the regression part of the latent HUD on covariates, and direct effects of covariates on specific items (DIF). Age of first hallucinogen use and number of days using hallucinogens were included as potential predictors of HUD (Wu et al., 2008). Values of Tucker-Lewis index (TLI) and comparative fit index (CFI) >0.95 and values of root mean square error of approximation (RMSEA) <0.06 indicate an excellent fit of the model to the data (Browne & Cudeck, 1993; Hu & Bentler, 1999). All results are weighted estimates except for sample sizes.

3. Results

3.1. Factor analysis

The scree plot and the ratio of the first eigenvalue to the second (6.55/1.11=5.9) indicated a dominant single factor underlying the 10 HUD criteria. The one-factor model of the 10 criteria (CFI=0.971, TLI=0.982, RMSEA=0.006) showed an excellent fit to the data (CFI=0.971, TLI=0.982, RMSEA=0.006), as did the two-factor model (CFI=0.976, TLI=0.985, RMSEA=0.005). The two-factor model indicated that the abuse and dependence factors were highly correlated (correlation coefficient=0.92). Together, they provided evidence for a unidimensionality of the 10 criteria for IRT modeling.

3.2. IRT model (Figure 1)

Of all items (severity parameters: 1.26–3.62), D2 (*taking larger amounts*) and D3 (*inability to cut down*) were rarely endorsed (2% among users) and represented the most severe levels on the HUD continuum (shifted to right end in Figure 1), followed by A3 (*problems with the law*) and A4 (*relationship problems*). D1 (*tolerance*) and D4 (*time spent*) were located among the lowest severity levels. Except for D2 (*taking larger amounts*) and D3 (*inability to cut down*) with low discrimination values (<1.0), all other items exhibited comparatively good-to-high discriminative power (discrimination: 1.17–2.23) in distinguishing between users at a middle-to-high severity level of HUD (severity: 1.32–2.15).

3.3. MIMIC model (Figure 2)

The MIMIC model incorporates a measurement component of the 10 criteria (HUD), regression part of predictors for HUD, and DIF (significant direct effects between demographic variables and HUD items). It fit the data very well (CFI=0.961, TLI=0.964, RMSEA=0.005).

DIF—There was indication of DIF by age, gender, race/ethnicity, and ecstasy use status. Controlling for the level of HUD, adolescents aged 16–17 years (vs. those aged 12–13 years) were less likely to endorse *role interference* (regression coefficient, β=™0.38), females (vs. males) were less likely to endorse *hazardous use* (β= ™0.26), blacks (vs. whites) were less likely to report *time spent* (β= ™0.60), Hispanics were more likely than whites to report *hazardous use* (β=0.67) and *tolerance* (β=0.65), and ecstasy users were more likely than non-ecstasy hallucinogen users to report *tolerance* (β=0.28).

Predictors of HUD—Adjusting for DIF, female gender (vs. male: β=0.22), ecstasy use (vs. use of other hallucinogens: β=0.36), and hallucinogen use on 12+ days (β=0.69) were associated with elevated odds of HUD. Hispanic ethnicity was associated with reduced odds of HUD (vs. whites: β= ™0.55).

4. Discussion

Consistent with studies of adults (Gillespie et al., 2007; Lynskey & Agrawal, 2007) and adolescents for other substance use (Gelhorn et al., 2008; Hartman et al., 2008; Martin et al. 2006; Wu et al., 2009a), this first IRT analysis of HUD criteria in adolescents provides support for a single dimensional construct underlying abuse and dependence criteria. We expand on

prior research by applying MIMIC methods to document elevated odds of HUD among ecstasy users, females, and whites while holding constant the influences of potential measurement errors (DIF) and the pattern of hallucinogen use.

4.1. Item severity and discrimination

IRT results suggest that HUD criteria capture a relatively severe subset of hallucinogen users (severity: 1.3–3.6 standardized units above the mean). Therefore, adolescents are unlikely to report criteria unless their HUD severity reaches a considerable level on the IRT-defined HUD continuum. Lynskey and Agrawal (2007) found a slightly lower range of severity (0.8–2.9) for adult hallucinogen users; however, they examined lifetime HUD symptoms, while we analyzed adolescents' past 12-month HUD symptoms as specified by DSM-IV.

All abuse criteria showed good discrimination (1.2–1.9) in distinguishing hallucinogen users (severity: 1.6–2.2). In contrast, dependence criteria captured a wider range on the continuum (severity: 1.3–3.6), and *taking larger amounts and inability to cut down* were less reliable in distinguishing between hallucinogen users (discrimination: 0.7–0.8). The latter may be explained by the very low rate of adolescents endorsing them (2%). Adolescents also might not use hallucinogens long enough to exhibit impaired control in hallucinogen use (Koob & Le Moal, 2006).

4.2. Differential item functioning

Controlling for HUD severity, ecstasy users were more likely to report *tolerance* than non-ecstasy hallucinogen users, suggesting a lower threshold/severity of this item in ecstasy users. Ecstasy users are often polysubstance users (Wu et al., 2006), and cross-tolerance to multiple substances can occur (Parrott, 2004, 2005). Polysubstance use might enhance ecstasy users' perceived tolerance to hallucinogen use, or they may misattribute tolerance from other substance use to hallucinogen use.

Additionally, controlling for HUD severity, females (vs. males) exhibited reduced odds of endorsing *hazardous use* (high threshold in females). Studies of alcohol and marijuana use disorders also found females showing a high threshold of endorsing hazardous use (Agrawal & Lynskey, 2007; Grant et al., 2007; Martin et al., 2006). *Hazardous use* may be gender-biased in that females might be less likely to admit to hazardous activities than males (Agrawal & Lynskey, 2007; Martin et al., 2006). These findings underscore the need to evaluate whether diagnostic questions or drug users' gender and cultural factors contribute to different thresholds for reporting a given item.

4.3. Predictors of HUD

Stone et al. (2006) reported that ecstasy users had higher odds of hallucinogen dependence than other hallucinogen users. MIMIC modeling reveals a robust association between ecstasy use and HUD by controlling for identified DIF, demographic characteristics, and patterns of hallucinogen use. This finding might be related to ecstasy's combined hallucinogenic and stimulant effects due to elevated levels of dopamine, serotonin, and noradrenaline in the brain (Clemens et al., 2007). Finally, elevated odds of HUD among females and whites indicate that hallucinogen use in these populations should be studied further to inform the design of gender-appropriate and culturally sensitive prevention programs.

4.4. Study limitations

These findings should be interpreted with caution. They cannot be applied to a very small group of institutionalized or homeless adolescents who are not included in the NSDUH. NSDUH also

relies on participants' self-reports, which are subject to memory errors or underreporting (SAMHSA, 2007).

4.5. Conclusions

Study findings suggest that the hierarchical distinction between hallucinogen abuse and dependence in adolescents is not fully justified. The effects of wording and content of diagnostic questions on the assessment of HUDs deserve research to ensure that diagnostic questions are applicable to different gender and racial/ethnic groups.

Acknowledgments

References

Agrawal A, Lynskey MT. Does gender contribute to heterogeneity in criteria for cannabis abuse and dependence? Results from the national epidemiological survey on alcohol and related conditions. Drug and Alcohol Dependence 2007;88(2–3):300–307. [PubMed: 17084563]

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed.. Washington, DC: American Psychiatric Publishing, Inc.; 2000. text rev.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed.. Washington, DC: American Psychiatric Publishing, Inc.; 1994.

American Psychiatric Association. DSM-V: The Future Manual. 2009 [Accessed March 28, 2009]. Available at: http://www.psych.org/dsmv.asp.

Baker, F. The Basics of Item Response Theory. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation, University of Maryland; 2001.

Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. Testing Structural Equation Models. Newbury Park, CA: Sage Publications; 1993. p. 136-162.

Cattell RB. The scree test for the number of factors. Multivariate Behavioral Research 1996;1:245–276.

Clemens KJ, McGregor IS, Hunt GE, Cornish JL. MDMA, methamphetamine and their combination: possible lessons for party drug users from recent preclinical research. Drug and Alcohol Review 2007;26(1):9–15. [PubMed: 17364831]

Crowley TJ. Adolescents and substance-related disorders: research agenda to guide decisions on Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-V). Addiction 2006;101:164–172.

Embretson, SE.; Reise, SP. Item Response Theory for Psychologists. Mahwah, NJ: Lawrence Erlbaum; 2000.

Gelhorn H, Hartman C, Sakai J, Stallings M, Young S, Rhee SH, Corley R, Hewitt J, Hopfer C, Crowley T. Toward DSM-V: an item response theory analysis of the diagnostic process for DSM-IV alcohol abuse and dependence in adolescents. Journal of the American Academy of Child and Adolescent Psychiatry 2008;47(11):1329–1339. [PubMed: 18827724]

Gillespie NA, Neale MC, Prescott CA, Aggen SH, Kendler KS. Factor and item-response analysis DSM-IV criteria for abuse of and dependence on cannabis, cocaine, hallucinogens, sedatives, stimulants, and opioids. Addiction 2007;102(6):920–930. [PubMed: 17523987]

Grant BF, Harford TC, Muthén BO, Yi HY, Hasin DS, Stinson FS. DSM-IV alcohol dependence and abuse: further evidence of validity in the general population. Drug and Alcohol Dependence 2007;86 (2–3):154–166. [PubMed: 16814489]

Hartman CA, Gelhorn H, Crowley TJ, Sakai JT, Stallings M, Young SE, Rhee SH, Corley R, Hewitt JK, Hopfer CJ. Item response theory analysis of DSM-IV cannabis abuse and dependence criteria in adolescents. Journal of the American Academy of Child and Adolescent Psychiatry 2008;47(2):165–173. [PubMed: 18176333]

Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling 1999;6:1–55.

Koob, GF.; Le Moal, M. The Neurobiology of Addiction. Elsevier, MA: Academic Press; 2006.

Lynskey MT, 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). Psychological Medicine 2007;37(9):1345–1355. [PubMed: 17407621]

Martin CS, Chung T, Kirisci L, Langenbucher JW. Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: implications for DSM-V. Journal of Abnormal Psychology 2006;115(4):807–814. [PubMed: 17100538]

Muth´n, BO.; Muth´n, LK. Mplus: Statistical Analysis with Latent Variables. Los Angeles, CA: Muth´n & Muth´n, Inc.; 2007.

Parrott AC. Chronic tolerance to recreational MDMA (3,4–methylenedioxymethamphetamine) or ecstasy. Journal of Psychopharmacology 2005;19(1):71–83. [PubMed: 15671132]

Parrott AC. MDMA (3,4-methylenedioxymethamphetamine) or ecstasy: the neuropsychobiological implications of taking it at dances and raves. Neuropsychobiology 2004;50(4):329–335. [PubMed: 15539865]

Stone AL, Storr CL, Anthony JC. Evidence for a hallucinogen dependence syndrome developing soon after onset of hallucinogen use during adolescence. International Journal of Methods in Psychiatric Research 2006;15(3):116–130. [PubMed: 17019896]

Substance Abuse and Mental Health Services Administration. Results from the 2006 National Survey on Drug Use and Health: National Findings. Rockville, MD: Office of Applied Studies, Substance Abuse and Mental Health Services; 2007.

Substance Abuse and Mental Health Services Administration. Results from the 2005 National Survey on Drug Use and Health: National Findings. Rockville, MD: Office of Applied Studies, Substance Abuse and Mental Health Services; 2006.

Substance Abuse and Mental Health Services Administration. Results from the 2004 National Survey on Drug Use and Health: National Findings. Rockville, MD: Office of Applied Studies, Substance Abuse and Mental Health Services; 2005.

Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. Science 1998;280(5365):867–873. [PubMed: 9572724]

Wu LT, Ringwalt CL, Yang C, Reeve BB, Pan JJ, Blazer DG. Construct and differential item functioning in the assessment of prescription opioid use disorders among American adolescents. Journal of the American Academy of Child and Adolescent Psychiatry 2009a;48(5):563–572. [PubMed: 19318987]

Wu LT, Pan JJ, Blazer DG, Tai B, Stitzer ML, Brooner RK, Woody GE, Patkar AA, Blaine JD. An item response theory analysis of alcohol and marijuana dependences: a National Drug Abuse Treatment Clinical Trials Network study. Journal of Studies on Alcohol and Drugs 2009b;70(3):414–425. [PubMed: 19371493]

Wu LT, Ringwalt CL, Mannelli P, Patkar AA. Hallucinogen use disorders among adult users of MDMA and other hallucinogens. The American Journal on Addictions 2008;17(5):354–363. [PubMed: 18770077]

Wu LT, Schlenger WE, Galvin DM. Concurrent use of methamphetamine, MDMA, LSD, ketamine, GHB, and flunitrazepam among American youths. Drug and Alcohol Dependence 2006;84(1):102–113. [PubMed: 16483730]
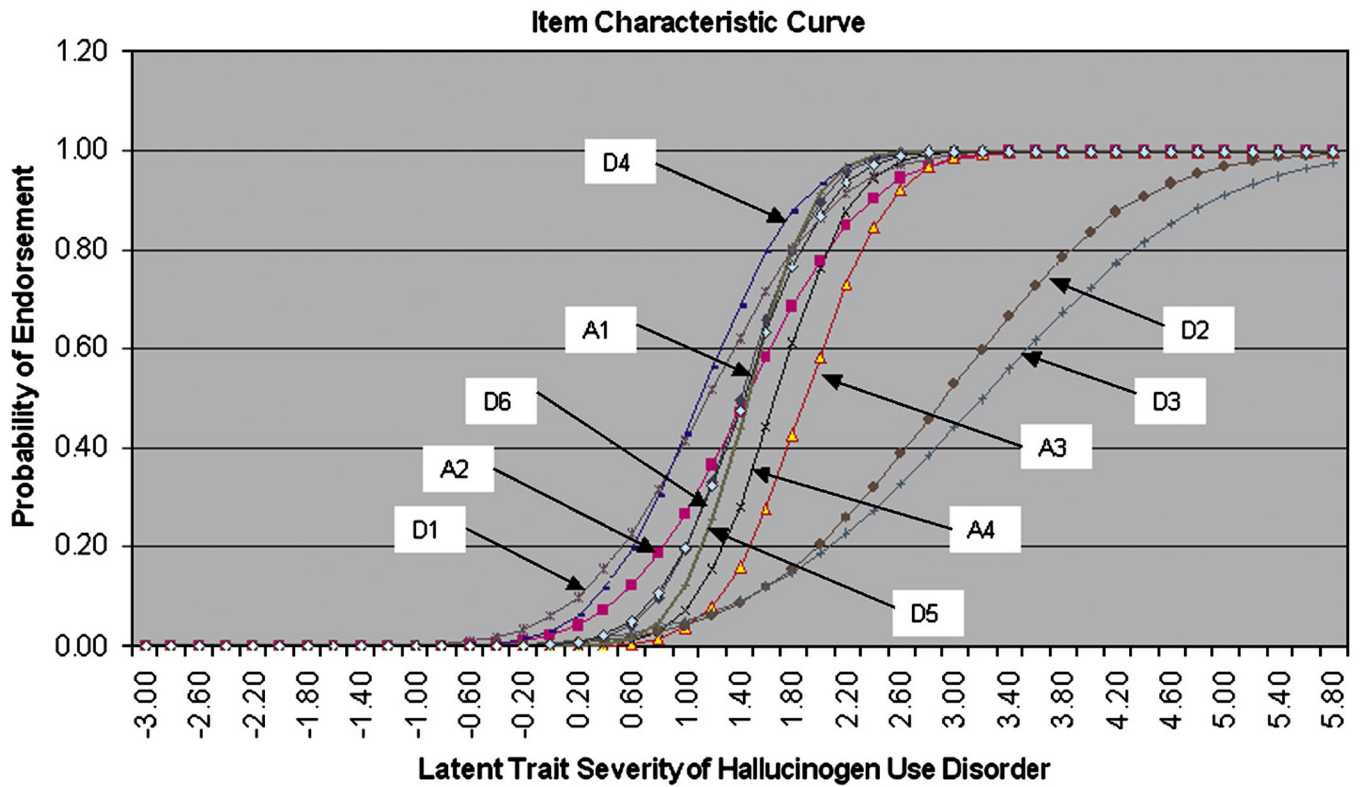
Figure 1.
Item characteristic curves (ICCs) for criterion symptoms of hallucinogen use disorders among
past-year hallucinogen users aged 12–17 years (N = 1548). A1 = role interference; A2 =
hazardous use; A3 = problems with the law; A4 = relationship problems; D1 = tolerance; D2
= taking larger amounts; D3 = inability to cut down; D4 = time spent; D5 = giving up activities;
D6 = continued use despite resulting medical or psychological problems.
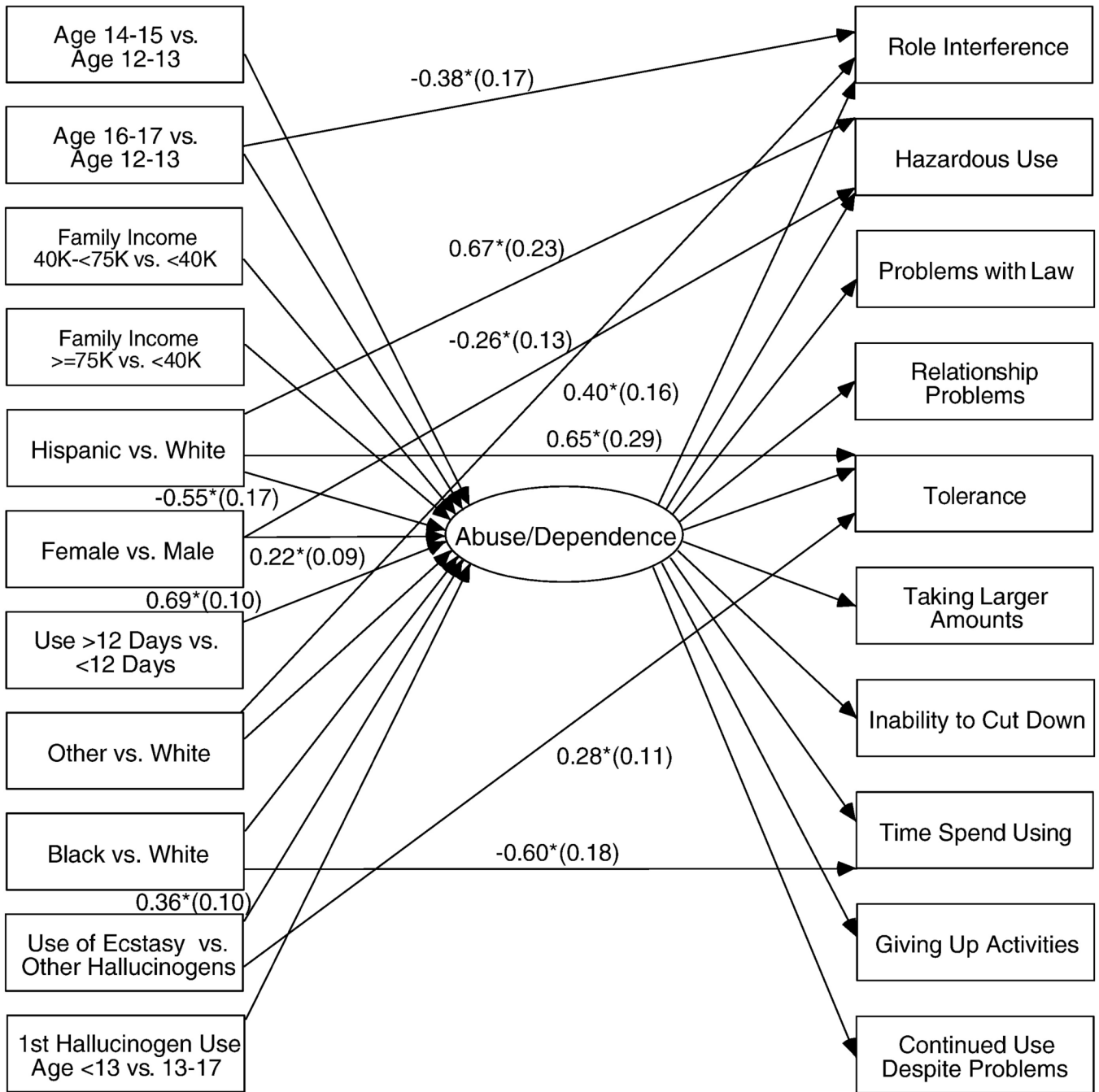
Figure 2.
The MIMIC model of hallucinogen use disorders among past-year hallucinogen users aged
12–17 years (N = 1548). Regression coefficient (standard error); *P value <0.05.