Do DSM-5 substance use disorder criteria differ by user care settings? An Item Response Theory analysis approach.

C. Kervran[# 1,2], D. Shmulewitz[3,4], F. Serre[1,2], C. Denis[6], P. Roux[7], M. Jauffret-Roustide[8,9,10], L. Lalanne[11], D. Hasin[3,4,5], M. Auriacombe[* 1,2,6]

Affiliations

[1]University of Bordeaux, Addiction Team, Sleep Addiction and Neuropsychiatry Laboratory (SANPSY), USR CNRS 3413, Bordeaux, France
[2] Pôle Addictologie, CH Ch. Perrens and CHU de Bordeaux, Bordeaux, France
[3] New York State Psychiatric Institute, New York, NY, USA
[4] Department of Psychiatry, Columbia University Irving Medical Center, New York, NY, USA
[5] Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA
[6] Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia,
[7] Aix Marseille University, INSERM, IRD, SESSTIM, Sciences Economiques and Sociales de la Santé and Traitement de l'Information Médicale, Marseille, France
[8] Cermes3, Inserm U988, CNRS UMR 8236, Université de Paris, EHESS, Paris, France
[9] British Columbia Centre on Substance Use (BCCSU), Vancouver, Canada
[10] Baldy Center for Law and Social Policy, Buffalo University of Social Sciences, New York, USA
[11] INSERM 1114, Department of Psychiatry and Addictology, University Hospital of Strasbourg, Fédération de Médecine Translationnelle de Strasbourg (FMTS), 67000 Strasbourg, France.

***Corresponding author**
Marc Auriacombe
e-mail: marc.auriacombe@u-bordeaux.fr
Pôle Addictologie, CHCP, 121 rue de la Béchade, 33076 Bordeaux Cedex, France
Tel: +33 556 561 738
Fax: +33 556 561 727

#Note
At time of publication, Charlotte Kervran PhD is now at Team MéRISP, Inserm U1219, Bordeaux Population Health Research Center (BPH), University of Bordeaux, Bordeaux, France.
,

Declarations of interest:
No author reports conflicts of interest related to this work. Dr. Marc Auriacombe, reports grants from French Health Ministry Research Grant Program (PHRC), grants from Aquitaine Regional Council Research Program (CRA); Dr. Marc Auriacombe, Dr. Perrine Roux, Dr. Laurence Lalanne and Dr. Marie Jauffret-Roustide report grants from French Government Addiction Agency (MILDT/MILDECA). Dr Hasin reports grants from NIAAA, NIDA and NYSPI during the conduct of the study. Ms. Kervran reports grant from French Rotary club of Saint Medard-En-Jalles and Libourne-Saint Emilion.

ABSTRACT

Aim: To examine differences in the psychometric characteristics of diagnostic criteria for Substance Use Disorders (SUD) between substance users in harm reduction settings (HR) and substance users seeking treatment (Tx).

Methods: Differential Item and Test Functioning (DIF & DTF) analysis were performed to examine differences in the difficulty of endorsement and in discrimination of the 11 diagnostic criteria and to test if the criteria set as a whole (the "test") functioned differently by care settings (Tx vs. HR) for alcohol, cocaine, cannabis, opiates and tobacco. To test uniform and nonuniform DIF, multiple indicator multiple cause (MIMIC) structural equation models were used.

Results: Regardless of the substance, the DSM-5 criteria "craving", "large amount", "time spent", "tolerance" and "activities given up" had similar functioning by care settings. Little evidence for DIF was found for other criteria. The criteria set as a whole did not function differently by care settings for alcohol, cocaine and tobacco. At the same trait severity, compared to HR, the Tx subgroup had a greater number of endorsed criteria for cannabis and a smaller number of endorsed criteria for opioids.

Conclusion: The unidimensionality of the 11 DSM-5 criteria and applicability of all criteria and diagnosis was confirmed in this large sample of problematic substance users. While the majority of the criteria related to loss of control of substance use, functioned well in both care settings, the criteria related to consequences of substance use had several differential functioning.

Keywords: Differential item functioning, Substance use disorder, DSM-5, Harm Reduction setting, patient seeking treatment.

Introduction

Substance use disorders have a variety of negative consequences, and are responsible for substantial morbidity and mortality worldwide (Degenhardt et al., 2018). Recent changes to Substance Use Disorder (SUD) diagnosis in the Diagnostic and Statistical Manual of Mental Disorders version 5 (DSM-5) included deleting the legal problem criterion, merging the remaining abuse criteria with dependence criteria and adding craving as a new criterion, for a set of 11 criteria to diagnose SUD (American Psychiatric Association, 2013; Auriacombe, Serre, Denis, & Fatseas, 2018; Hasin et al., 2013). Studies across substances in general population and clinical samples highlighted that the 11 criteria were distributed over a continuum of severity (latent trait) supporting the unidimensionality and the psychometric validity of these 11 criteria (Castaldelli-Maia et al., 2015; Chung, Martin, Maisto, Cornelius, & Clark, 2012; Hagman, 2017; Hasin et al., 2013; Serier, Venner, & Sarafin, 2019; Shmulewitz, Greene, & Hasin, 2015). The current validity of the 11 DSM-5 alcohol, cannabis, cocaine, opioids and tobacco use disorder criteria as latent constructs was observed in clinical samples from various types of addiction treatment settings, where the clinical outcomes were substance use reduction or abstinence (Hasin, Fenton, Beseler, Park, & Wall, 2012; Kervran et al., 2020; Serier et al., 2019).

However, substance users attempting to abstain in addiction treatment settings may have different characteristics and expectations than substance users in harm reduction (HR) settings. For example, needle exchange program participants are substance users seeking support to reduce health consequences of their substance use, while not abstaining from use (Des Jarlais, 1995; Hawk et al., 2017). The prevalence of SUD is high among substance users in HR settings (Kidorf et al., 2004), but they may have less loss of control of substance use and may be diagnosed more on the consequences of their use. No studies have reported on the prevalence and distribution of each diagnostic criterion among HR participants (Kervran et al., 2018). Inversely, some criteria could be specific to substance users in treatment settings (Kessler, Molnar, Feurer, & Appelbaum, 2001), as strong desires to quit use (Serier et al., 2019), or craving are often described as particularly prevalent among people trying to abstain (Cummings, Jaen, & Giovino, 1985; Shiffman et al., 1986).

An important element for the validity of the DSM-5 SUD diagnostic criteria is to determine whether criteria or criteria sets function differently across population subgroups. Differential item functioning (DIF) occurs if a criterion (an “item”) does not have the same relationship to a latent variable across subgroups (Derringer et al., 2013; Embretson & Reise, 2000; Hasin et al., 2013; Shmulewitz et al., 2015). Criteria with DIF may not work the same way in specific groups (Shmulewitz et al., 2015). There are two types of DIF. Uniform indicates a difference in the difficulty of the criterion by subgroups and Nonuniform indicates a difference in discrimination by subgroup (Woods & Grimm, 2011). If enough items function differently in term of difficulty (Uniform DIF) in a set and do so consistently, this can lead to differential test functioning (DTF), which means that specific groups are more likely to get a diagnosis because they are more likely to endorse some criteria regardless of the underlying severity (Shmulewitz et al., 2011). If all criteria show DIF in the same direction, this could have implications for the validity of the SUD construct, but if they have DIF in different directions, this could balance out and not threaten the overall validity. Previous studies in clinical samples have found miscellaneous although inconsistent Uniform DIF and DTF by sociodemographic and psychiatric disorders (Hasin et al., 2012; Kervran et al., 2020). These instances of DIF and DTF were not consistent enough to indicate overall invalidity of the 11 criteria. However, no studies have compared DIF and DTF between samples of substance users from different care settings, notably users who sought treatment to stop their substance use entirely as distinct from users in harm-reduction settings.

For the same SUD latent trait severity, diagnostic criteria that are related to consequences of substance use may be more frequent and diagnostic criteria related to the loss of control of substance use (i.e., craving, strong desires/difficulty to quit use) may be less frequent among substance users in HR settings than among substance users seeking addiction treatment. If this were the case, such criteria might function differently, and have implications for the validity of the SUD construct. Thus, this study examined the psychometric properties of the 11 DSM-5 SUD diagnostic criteria for alcohol, opioids, cocaine, cannabis and tobacco in a large community sample

of substance users, comparing substance users in harm reduction (HR) settings and substance users in outpatient addiction treatment (Tx) settings.

Methods

Participants and study design

Participants were selected from two cohort studies, the Addiction Aquitaine cohort (ADDICTAQUI) (Auriacombe, Accessed September 2019) and the COhort to identify Structural and INdividual factors associated with drug USe (COSINUS) (Auriacombe et al., 2019). Participants seeking addiction treatment (Tx) (n=1,359) and therefore attempting to abstain from use of substances from the ADDICTAQUI cohort were recruited at their entry in outpatient addiction clinics if they met criteria for a DSM-5 substance use disorder for at least one substance (Auriacombe, Accessed September 2019). Participants seeking support in harm reduction settings (HR) (n=130) from the COSINUS cohort were users of psychoactive substances who had injected at least once during the past month (Auriacombe et al., 2019). The common inclusion criteria for the two samples were: over 18 years old, French-speaking, problematic regular users of at least one substance (alcohol, cocaine, tobacco, opioids, cannabis) and provided documented informed consent. Participants were considered problematic regular users if they used the substance at least 2 times per week for the past 12 months. Both cohorts were approved by French Regulation and ethical committee (CNIL, CPP, CEEI/IRB) (Auriacombe, Accessed September 2019; Auriacombe et al., 2019).

Measures

Each subject completed a baseline interview that included a face to face standardized and structured interview, administered by trained interviewers. Sociodemographic variables (age, gender, level of education) and substance use-related variables were collected with the Addiction Severity Index (ASI) for the Tx participants (Denis et al., 2016; McLellan et al., 1992) and a standardized questionnaires based on the ASI, specifically created for the COSINUS cohort (Auriacombe et al., 2019) for the HR participants. Among regular users of both groups and for each substance, the 11 DSM-5 SUD criteria were evaluated over the past 12 months (Hasin et al., 2013; Kervran et al., 2020). Polyaddiction was defined as qualifying for more than one current SUD.

Analyses

For each substance, analyses were conducted separately among regular users: 876 alcohol users (Tx n= 787, HR n=89); 233 opioid users (Tx n= 131, HR n=102); 223 cocaine users (Tx n= 141, HR n=82); 599 cannabis users (Tx n= 504, HR n=95); and 1,142 tobacco users (Tx n= 1014, HR n=128). The majority of the sample was from the ADDICTAQUI cohort, for which the unidimensionality and psychometric validity of the 11 SUD DSM-5 criteria were previously reported (Kervran et al., 2020). In this paper, the HR subgroup (COSINUS) was added to the previously studied Tx subgroup.

Unidimensionality and Item Response Theory (IRT) models

For each substance, a one-factor confirmatory factor analysis (CFA) model of the 11 SUD criteria confirmed unidimensionality when the model showed adequate fit, based on comparative fit index (CFI) ≥0.95 or Tucker-Lewis index (TLI) ≥0.95 and RMSEA (smaller values indicate better fit) ≤.06 (Hu & Bentler, 1999). Then, a 2-parameter logistic IRT model was estimated, to examine each criterion's difficulty to be endorsed (inversely related to frequency) and discrimination (how well the criterion differentiated between respondents with high and low severity of the latent trait) (Embretson & Reise, 2000; Hasin et al., 2012; Shmulewitz et al., 2011).

Differential Item Functioning (DIF)

For each substance, Differential Item (criterion) Functioning analysis was carried out using multiple indicator multiple cause (MIMIC) structural equation models (Jones, 2006; Kline, 2011; C. M. Woods, Oltmanns, & Turkheimer, 2009), allowing examination of uniform DIF (differences in item difficulty). A strength of the MIMIC model is its ability to examine and to adjust the impact of

DIF during analyses, as well as the ability to examine multiple group variables simultaneously, allowing control for several covariables (Teresi & Fleishman, 2007). The relatively small sample size for the HR subgroup led to the choice of MIMIC modeling (Jones, 2006; Kline, 2011; C. M. Woods et al., 2009). Each criteria set was modeled as a latent variable indicated by the 11 criteria and regressed on care settings (i.e., Tx vs HR), while adjusting for age, gender, years of regular substance use (dichotomized by the median in each group), level of education (>12 years vs. <12 years) and poly-addiction (Figure 1). A second model was performed, testing DIF by age and gender. DIF is indicated when a criterion shows a statistically significant association with care setting, after accounting for the association with the latent variable. This independent association (between the criterion and the care settings) is indicated by statistically significant modification indices (MI), which suggest that adding that association to the model would improve model fit. If DIF was indicated, we determined which care settings (Tx or HR subgroup) showed higher probability of endorsement of the criterion at the mean of the latent trait.

Nonuniform DIF (differences in item discrimination) analysis was also carried out using another MIMIC structural equation model (Figure 1) (Woods & Grimm, 2011). Each criteria set was modeled as a latent variable indicated by the 11 criteria and the latent variable was regressed on the covariates, adjusting for age, gender, level of education and poly-addiction. Each criterion was regressed on care settings and on the interaction between care setting and the latent trait. (Woods & Grimm, 2011). DIF is indicated when a criterion shows a statistically significant association with interaction between care setting and latent trait. Nonuniform DIF were also tested for age and gender for each criterion.

To adjust for multiple testing of the 11 criteria, we used Bonferroni adjustment, considering a result significant at $p=0.05/11=0.0045$ (≤0.005). All IRT analysis were conducted in Mplus 8 (Muthen & Muthen, 1998-2017).

Differential Test Functioning (DTF)

For each substance, we examined DTF by care settings (Tx vs. HR) (Morales, Flowers, Gutierrez, Kleinman, & Teresi, 2006; Raju, Van der Linden, & Fleer, 1995). We used R code (R Core Team, 2018) to calculate the average difference in the expected number of criteria for individuals with the same trait severity in each treatment environment (Shmulewitz et al., 2011). A difference of <1 expected number of criteria by subgroup indicates no DTF, as differences that small should lead to minimal differential diagnosis of SUD by care settings.

Results

Sample Description

The majority of the individuals were male (68%), the mean age was 38 years, and 56% had a high level of education (≥12 years) (Table 1). The large majority of regular users met SUD diagnosis (alcohol 93%, opioids 98%, cocaine 92%, cannabis 92% and tobacco 89%). Overall more than half (66%) qualified for polyaddiction, with more polyaddiction in the HR subgroup. The endorsement rate to each of the DSM-5 criterion was high for alcohol, opioids, cocaine, cannabis, and more variable for tobacco (10.9% to 80.0%) (Table 2).

Dimensionality & IRT

For alcohol, opioids, cocaine and cannabis, unidimensionality was confirmed by the model fit indices (Table 2). For tobacco, unidimensionality was confirmed by RMSEA, but TLI and CFI were slightly below the recommended 0.95. Across all substances, discrimination parameters ranged from 0.57 to 2.38, indicating that criteria had relatively high ability to delineate individuals who were higher vs. lower on the latent trait, and the difficulty estimates of ≤0.0 indicated that even subjects below average on the latent trait had at least a 50% probability of endorsing the criterion. IRT analysis indicated that for each substance, the 11 items were spread across the severity continuum, with "craving" showing high discrimination and "hazardous use" showing low discrimination (Figure 2).

Differential Item Functioning

Care settings

DIF analyses (Table 3) indicated that for all substances, the "craving", "large amount", "time spent", "tolerance" and "activities given up" criteria showed no Uniform DIF, and, for alcohol and opioids, all criteria showed no Uniform-DIF.

For cocaine, the "withdrawal" criterion was significantly (p=.001) more likely to be endorsed in Tx subgroup than HR subgroup. For cannabis, the "quit/control" criterion was significantly (p=.001) more likely to be endorsed in Tx subgroup than HR subgroup, while the "social problems" criterion was significantly (p=.001) less likely to be endorsed in Tx subgroup than HR. For tobacco, the "quit/control" and "psychological/physiological" problems criteria were significantly (p<.001 and p=.001 respectively) more likely to be endorsed in Tx subgroup than HR subgroup, while the "neglect role" and "hazardous use" criteria were significantly (p<.001) less likely to be endorsed in Tx subgroup than HR. No Nonuniform DIF was found for the care setting for the 5 substances.

Age

For cannabis, the "hazardous use" criterion was significantly (p=.001) more likely to be endorsed in older subjects than younger subjects. For alcohol, the "tolerance" criterion was significantly more likely to be endorsed in younger subjects than older (p<.0001) and "physical/psychological problems" criterion was significantly more likely to be endorsed in older than younger subjects (p=.004). For tobacco, "difficulty to quit or control" was more likely to be endorsed in older than younger subjects (p<.0001) and "hazardous use" criterion was more likely to be endorsed in younger than older subjects (p=.002). No Nonuniform DIF was found for the age for the 5 substances.

Gender

For alcohol, cocaine, cannabis and tobacco, the "hazardous use" criterion was significantly more endorsed by males than by females (all p's ≤0.0001). No Nonuniform DIF was found for the gender for the 5 substances.

Differential Test Functioning (DTF)(Figure 3)

For the total criterion set, the average expected difference in number of criteria endorsed between the two care setting subgroups was <1 for alcohol (0.39), cocaine (0.39) and tobacco (0.51). For cannabis, the difference was 1.16, with the Tx subgroup endorsing a greater number of criteria than the HR subgroup at the same trait severity. For opioids, the difference was 1.09, with the HR subgroup endorsing more criteria than the Tx subgroup at the same trait severity.

Discussion

To our knowledge, this is the first study to examine differential item and test functioning of DSM-5 SUD criteria between substance users accessing treatment (Tx) and substance users in harm reduction settings (HR). Across all substances (alcohol, cannabis, cocaine, opioids, and tobacco), the majority of the criteria had similar functioning regardless of the care setting. For cannabis and opioids, the criteria set as a whole functioned differently (DTF) by care settings.

Consistent with prior studies, unidimensionality was confirmed for alcohol, opioids, cocaine and cannabis, further supporting the DSM-5 changes (Hasin et al., 2012; Hasin et al., 2013; Kervran et al., 2020). For tobacco, unidimensionality of the 11 DSM-5 criteria was equivocal, which could be explained by the low factor loadings and DIF of several criteria (Kervran et al., 2020).

From the loss of control of use of substances dimension of SUD (Martin, Langenbucher et al. 2014, Auriacombe, Serre et al. 2018), the "craving", "using in larger amount than intended" and "time spent to use" criteria had similar functioning across care settings. For cannabis and tobacco, the "difficulty to quit or control" criterion was more frequently endorsed in the past 12 months, in the treatment setting subgroup, which may explain why they sought treatment for addiction. Substance users in harm reduction settings could be either more interested in reducing harm than

controlling cannabis or tobacco use, or they could have been more successful in controlling the use of those substances since such individuals were likely to also use other substances (e.g., alcohol, opioids and cocaine) to compensate (Copersino et al., 2006; Peters & Hughes, 2010). Substitution between substances has been reported to occur to relieve craving or withdrawal (Copersino et al., 2006), related to common conditioned cues (Midanik, Tam, & Weisner, 2007), or similar pharmacological pathways, helping to control the use of one substance with another substance. Furthermore, some substance users in harm reduction settings could have other priorities than their use of tobacco and cannabis, due to their precarious situation.

Of the pharmacological adaptation criteria (withdrawal and tolerance), cocaine users seeking addiction treatment endorsed "withdrawal" significantly more frequently than cocaine users in harm reduction settings. This could be because subjects entering treatment settings may attempt to reduce use before treatment access, and may experience more withdrawal as a consequence than substance users in HR settings who are not necessarily attempting to reduce use and are in an environment where access to substances is easier (Kervran et al., 2019; Moracchini et al., 2012). Another hypothesis is that users seeking addiction treatment are more prone to experience aversive withdrawal than the others and thus are more likely to look for help and addiction care.

Of the consequences of use criteria, those related to social impairment and risky use (physical or psychological), which tend to be contextually and culturally bound (Auriacombe et al., 2018; Martin, Langenbucher, Chung, & Sher, 2014) , functioned differently by care settings. The HR subgroup was more likely to endorse “social/interpersonal problems due to use” for cannabis and “neglect roles because of use” and “hazardous use” for tobacco, perhaps due to socio-economic difficulties, often related to substance use or dealing activities that are more frequent among substance users in HR than for those in treatment (Kervran et al., 2019; Moracchini et al., 2012). For tobacco, “psychological/physiological problems because of use” criterion was more frequently endorsed by substance users in treatment settings, possibly because this group was more aware of such damage caused by tobacco use.

There were no criteria with consistent differential functioning by care settings across all substances. No Nonuniform DIF was found for the care setting for the 5 substances.

Overall, the expected number of criteria endorsed was quite similar for those in harm reduction and treatment settings. For opioids, no individual criteria had a DIF, but insignificant functional differences of several criteria, when taken together, were added to give a significant DTF. These results assume that opioid users in HR will present more criteria at each level of latent trait severity than those in treatment. This suggests that opioid users in HR are more likely to be diagnosed with an opioid use disorder than opioid users in treatment and thus may be over-diagnosed in this population. Substance users in treatment setting endorsed a greater number of cannabis criteria, perhaps driven by the increased likelihood of endorsing the “difficulty to quit or control” criterion. For cannabis the DTF indicates that cannabis users in Tx are more likely to be diagnosed or diagnosed more severely for a cannabis use disorder than those in HR. Future studies should determine exact causes of the DTF.

Our DIF analyses, by age and by gender indicated that the majority of differences of functioning on criterion difficulty were associated with age, consistent with previous findings (Kervran et al., 2020). Such findings are more likely to be linked to the total duration of lifetime use. It seems of interest to use the duration of use or the age of onset of SUD as a covariate in future studies. In all the substance use disorders explored, males endorsed more frequently (lower difficulty) the “hazardous use” criterion than females, except opioids, suggesting that these criteria may not work well in female subgroups (Derringer et al., 2013). As shown in a study, males were significantly higher than females on all of the sensation seekers scale scores (Shulman, Harden, Chein, & Steinberg, 2015; Zuckerman, Eysenck, & Eysenck, 1978), explaining that males reported more situations physically at risk under the influence of alcohol, in addition to the disinhibition induced by alcohol use. This will need to be monitored in future studies.

Study limitations are acknowledged. The sample size for the HR subgroup was small for some substances. A larger sample would give more precise estimates and greater ability to detect differences. While the MIMIC model is the recommended method for small sample sizes (as in HR group) and is convenient because it allows every item to be evaluated, it often leads to high Type I error rates (Chun, Stark, Kim, & Chernyshenko, 2016). But in our study no Nonuniform DIF was found for the care setting. In a previous study among the substance user treatment seeking subgroup, DIF and DTF by psychiatric comorbidities were found (Kervran et al., 2020). Since substance users in treatment settings showed more severe mental health disability than substance users in HR settings (Compton, Thomas, Stinson, & Grant, 2007; Kidorf et al., 2004; Kidorf et al., 2010; Metzger et al., 1990; Ross et al., 2002), future studies should explore if psychiatric comorbidities in the treatment subgroup could explain the DIF and DTF observed. These populations are different in terms of precariousness (Moracchini et al., 2012; Ross et al., 2005), which has not been fully explored in this study, and which could explain the DIF of criteria related to consequences of use that are contextually and culturally bound. Finally, we cannot exclude that some users attending HR settings may also have been attending treatment settings and conversely, some subjects seeking treatment may also have been attending HR settings. Nonetheless, our data show that HR participants were different from Tx participants, especially on substance use, supporting that we did not study the same substance users recruited from different locations, but different subjects in different care settings.

In conclusion, the present study in a large sample of substance users in two different types of care settings supported the unidimensionality of the 11 DSM-5 criteria and applicability of all criteria and diagnosis. It is important to distinguish « peripheral » characteristics from the « core » characteristics of the disorder (Martin, Langenbucher et al. 2014, Auriacombe, Serre et al. 2018). Central features can be defined as symptoms and constructs that directly reveal the underlying internal dysfunction of the disorder, such as loss of control of use for substance use disorders (Martin, Langenbucher et al. 2014). Peripheral features do not directly index these internal dysfunctions (loss of control), but its consequences (Martin, Langenbucher et al. 2014). While the majority of the loss of control of substance use criteria, including craving, larger amount/longer, time spent and quit/control criteria, considered by some to be the core feature of addiction (American Psychiatric Association, 2013; Auriacombe et al., 2018; Martin et al., 2014), functioned well in the different care settings, the criteria related to consequences of use had several instances of differential functioning, calling into question their inter-context stability (DIF). This is why we suggest that they may be considered as peripheral rather than central characteristics of use disorders, notwithstanding that they should be a cause for concern by treatment providers and are a cause for impairment and suffering for users with addiction (Auriacombe et al., 2018; Martin et al., 2014; Pickard & Ahmed, 2016).

Several studies suggest that the substance use disorders should be diagnosed using the central dimension of the disease which is loss of control of use (Martin, Langenbucher et al. 2014, Auriacombe, Serre et al. 2018). Future studies should explore if the current set of 11 criteria could be simplified with less criteria more focused on those related to the direct expression of the loss of control of substance use (including craving, larger amount/longer, time spent and quit/control criteria) without losing any of the current psychometric quality.

Acknowledgements
The authors express their thanks to all participants for their contribution. The authors are grateful to medical staff of the Bordeaux recruitment sites of ADDICTAQUI and COSINUS for facilitating recruitment of participants (Saman Sarram, Jacques Dubernet, Bérengère Gelot, Cécile Gazel, Marie Chivé, Benedicte Apert, Victor Harly, Nathalie Ramdjee, Louise Jakubiec, Jean-Michel Delile, Véronique Latour, Jean-Pierre Daulouède). A very special thanks to all the interviewers of the ADDICTAQUI Bordeaux site team 2009-2019 (Floriane Sevin, Fuschia Serre, Xavier Balducci, Juliette Leboucher, Marion Malivert, Laure Childer, Jean-Marc Alexandre, Romain Debrabant, Cécile Denis, Hortense Hurmic, Charlotte Kervran, Manon Chevalier, Sarah Moriceau, Lucie Fournet, Marion Desset, Anais Saillard) and the interviewers of the COSINUS Bordeaux site team 2016-2018 (Charlotte Kervran, Sarah Moriceau, Manon Chevalier).

Figure 1: MIMIC models for Uniform and Nonuniform Differential Item Functioning (DIF) testing; the solid lines indicate the tested DIF and the dashed lines the adjustment variables links.

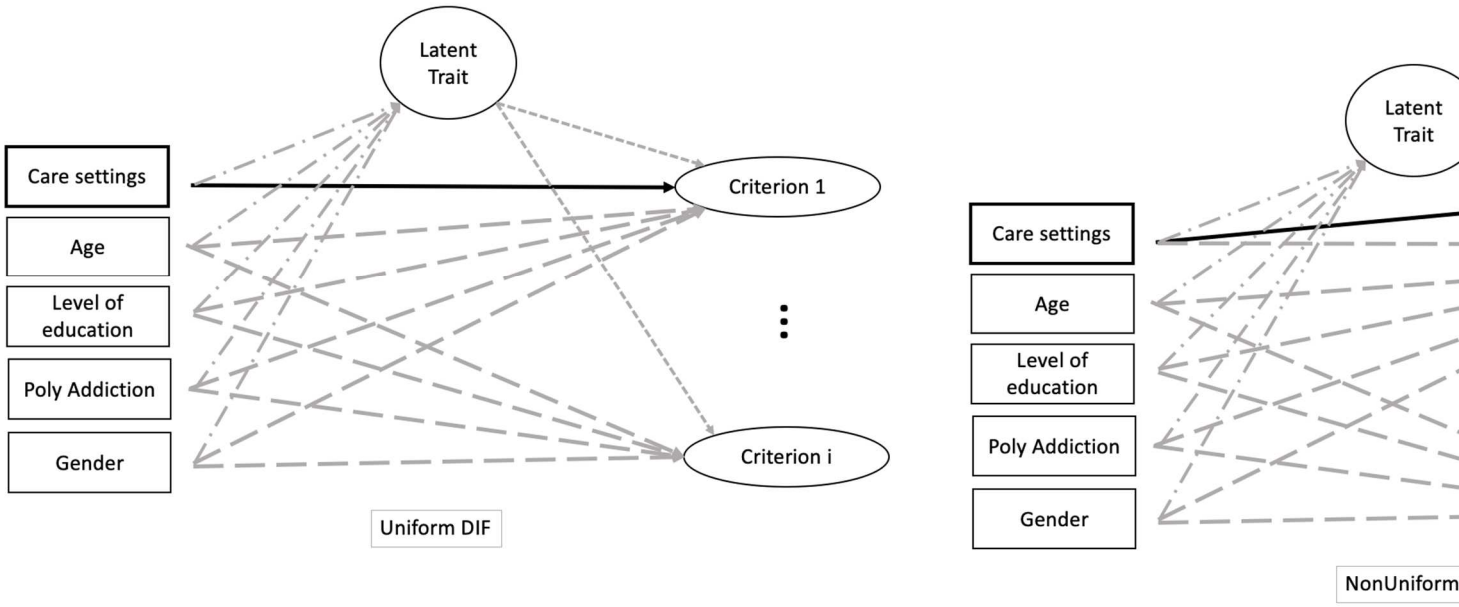


Table 1: Sociodemographic and addiction severity variables in a substance using sample, by substance and care setting.

	Total sample		Alcohol (n=876)		Opioids (n=233)		Cocaine (n=223)		Cannabis (n=599)		Tobacco (n=1142)	
	HR	**Tx**	**HR**	**Tx**	**HR**	**Tx**	**HR**	**Tx**	**HR**	**Tx**	**HR**	**Tx**
Age – Mean (SD)	34.2(8)	38.6 (12)	34.9(8.1)	40.3(11.6)	33.6(7.9)	34.3(9.7)	34.2(7.8)	32.8(7.8)	35.2(8.1)	33.5(9.6)	34.1(7.8)	38.3(11.3)
Males – n (%)	99 (76.2)	914 (67.3)	74(83.2)	569(72,3)	79(77.5)	89(67.9)	63(76.8)	115(81.6)	75(79.0)	398(79.0)	97(75.8)	660(65.1)
Level of education ≥ 12 years – n (%)	34 (26.6)	799 (59.3)	21(23.9)	471(60.5)	32(32.0)	58(44.6)	16(19.5)	69(48.9)	22(23.7)	242(48.2)	34(27.0)	609(60.4)
Number of years of regular use – Mean (SD)	-		14.7(7.9)	14.3(10.6)	6.3(4.8)	5.1(4.5)	6.3 (6.6)	4.3(4.6)	18.0(8.3)	13.8(8.9)	19.5(8.0)	20.0(10.5)
Mean number of use days in the last 30 days	-		26.3(7.1)	15.4(11.7)	25.5(8.3)	23.5(10.9)	14.0(10.7)	5.3(7.7)	23.9(8.8)	20.2(12.1)	29.4(3.7)	27.0(8.3)
Poly-addiction – n (%)	129 (99.2)	769 (56.6)	89(100)	554(70.4)	102(100)	100(76.3)	81(98.8)	127(90.1)	94(99.0)	410(81.4)	127(99.2)	701(69.1)
DSM-5 criteria												
No. endorsed SUD criteria – Mean (SD)	-		6.6 (2.8)	6.9 (3.0)	8.3 (2.2)	7.4 (2.6)	7.2 (3.1)	7.2 (3.2)	5.6 (2.8)	6.7 (2.9)	5.5 (2.3)	5.0 (2.5)
SUD diagnosis – n (%)			87 (98)	728 (93)	102 (100)	128 (98)	76 (93)	133 (93)	87 (91)	474 (95)	119 (94)	928 (92)
Mild			15 (17)	60 (8)	1 (1)	7 (5)	8 (10)	17 (12)	14 (15)	56 (11)	14 (11)	157 (16)
Moderate			15 (17)	122 (16)	15 (15)	23 (18)	9 (11)	15 (9)	27 (28)	78 (16)	44 (35)	334 (33)
Severe			57 (64)	546 (69)	86 (84)	98 (75)	59 (72)	101 (72)	46 (48)	340 (68)	61 (48)	437 (43)

S.D.: Standard Deviation; HR: substance users in harm reduction setting; Tx: substance users in treatment setting.

Table 2: Parameter estimates from Confirmatory factor analysis 1-factor model, by substance.

	Alcohol (n=876)		Opioids (N=233)		Cocaine (n=223)		Cannabis (n=599)		Tobacco (n=1142)	
	factor loading	Prevalence (%)	factor loading	Prevalence (%)	factor loading	Prevalence (%)	factor model loading	Prevalence (%)	factor loading	Prevalence (%)
Large amount/longer	0.469	81.3	0.586	73.8	0.736	77.6	0.647	58.4	0.568	73.9
Quit/control	0.707	65.1	0.566	81.1	0.667	57.0	0.627	55.9	0.498	67.6
Time spent	0.629	52.5	0.662	67.4	0.793	68.6	0.615	54.4	0.593	39.0
Craving	0.825	70.9	0.699	87.1	0.826	78.9	0.789	77.8	0.780	80.0
Neglect roles	0.603	51.9	0.739	35.6	0.680	43.5	0.529	38.2	0.484	10.9
Social/interpersonal problems	0.735	65.3	0.573	60.9	0.689	64.1	0.589	57.4	0.410	26.9
Activities given up	0.782	51.6	0.778	67.0	0.794	58.7	0.646	53.4	0.579	15.1
Hazardous use	0.380	72.1	0.421	59.7	0.478	73.1	0.328	71.3	0.354	22.2
Psychological/physical problems	0.673	59.1	0.553	69.5	0.815	64.1	0.623	54.8	0.448	58.2
Tolerance	0.613	66.2	0.601	85.8	0.690	72.2	0.636	64.6	0.645	45.6
Withdrawal	0.731	49.5	0.678	89.7	0.673	50.7	0.741	65.8	0.697	70.1
Model fit indices										
CFI	0.967		0.954		0.980		0.970		0.919	
TLI	0.958		0.942		0.975		0.962		0.899	
RMSEA	0.053		0.048		0.048		0.045		0.055	

CFI: Comparative Fit Index, TLI: Tucker Lewis Index, RMSEA: Root Mean Square error of approximation

Figure 2: Item characteristic curves from IRT models of alcohol, opioids, cocaine, cannabis and tobacco DSM-5 use disorder.

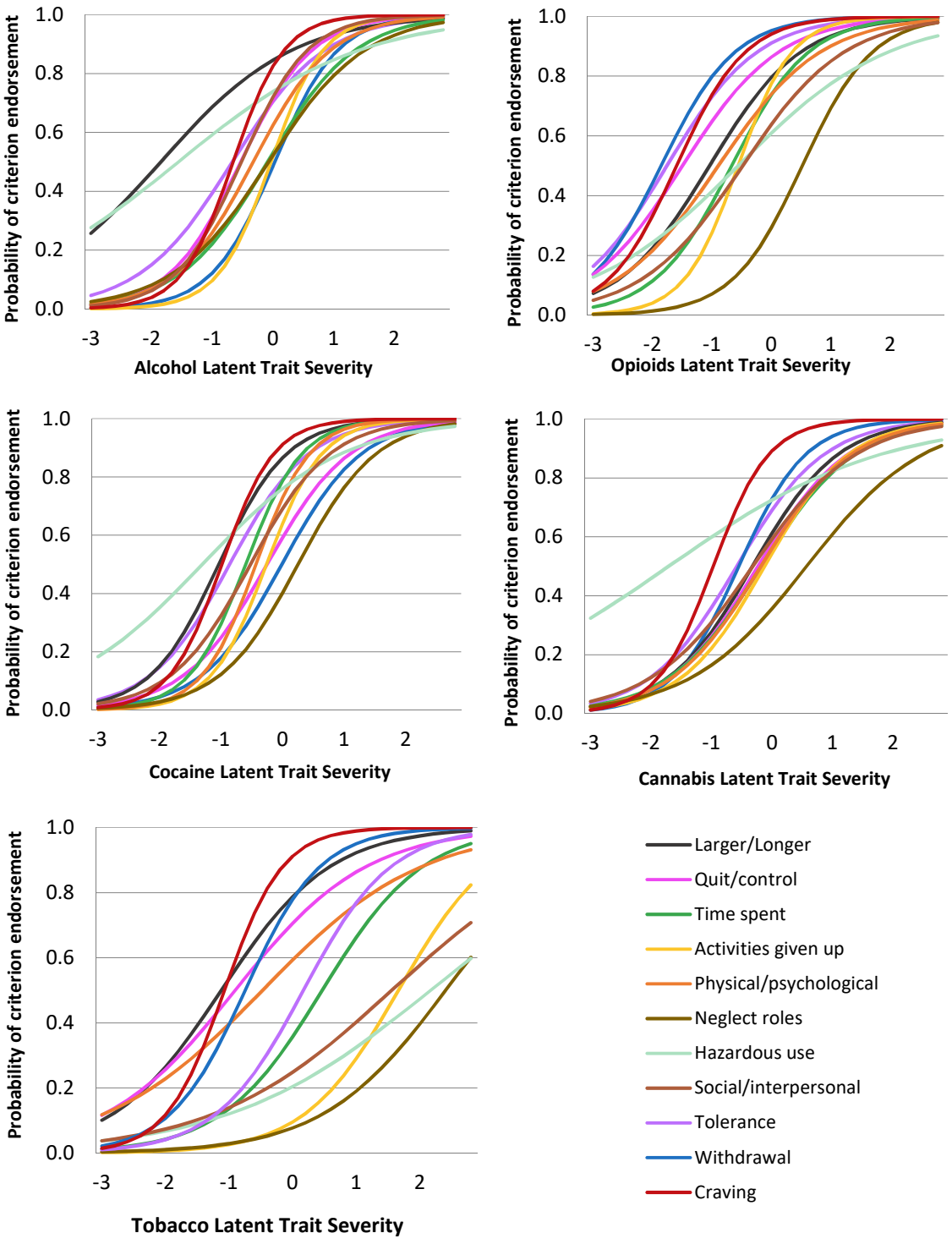
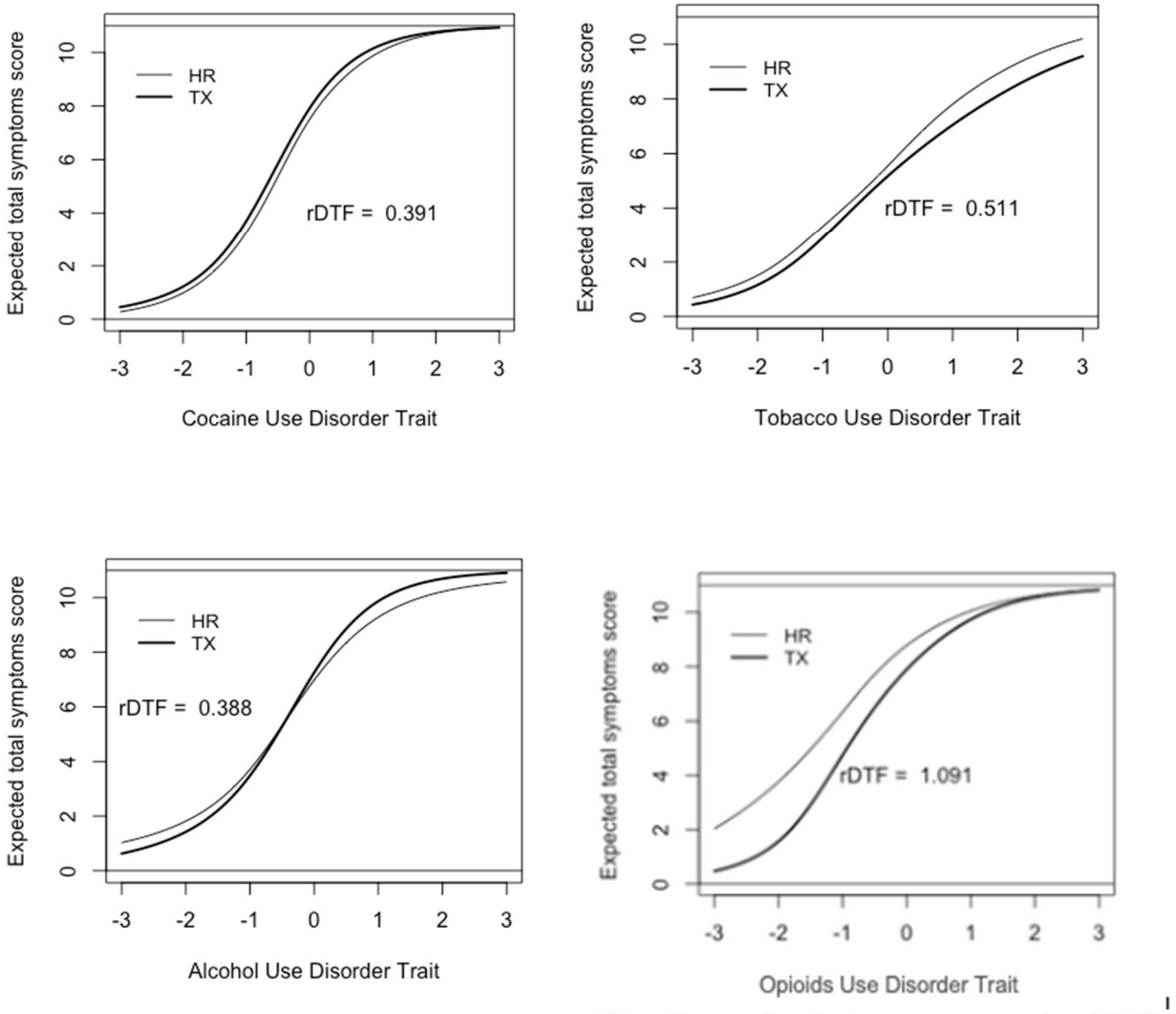


Table 3: Uniform Differential Item Functioning (DIF): modification indices (MI) values of the independent association test (between the criterion and the care setting), adjusted for age, gender, poly-addiction and level of education.

DSM classification of criteria	Criteria	**Alcohol (n=876)**	**Opioids (N=233)**	**Cocaine (n=223)**	**Cannabis (n=599)**	**Tobacco (n=1142)**
Loss of control	Large amount/longer	7.052	1.298	0.675	0.097	0.326
	Quit/Control	4.685	2.742	5.378	**12.390**	**29.252**
	Time spent	1.067	0.308	0.511	0.158	7.090
	Craving	0.092	1.068	0.719	0.391	1.901
Social Impairment	Activities given up	0.121	0.385	0.640	0.005	1.017
	Social/interpersonal problems	1.427	4.754	1.435	**13.739**	0.006
	Neglect role	0.082	0.076	0.268	0.430	**31.972**
Risky use	Hazardous use	6.920	2.083	0.081	1.941	**24.448**
	Psychological/Physical problems	1.112	0.082	2.524	0.059	**10.835**
Pharmacological adaptation	Tolerance	3.447	0.083	2.128	0.092	1.960
	Withdrawal	0.856	1.999	**8.303**	1.160	3.648

Note: Uniform DIF or independent association (between the criterion and the care settings) is indicated by statistically significant modification indices (MI) (in bold)

Figure 3: Differential Test functioning by care settings.



Note: For each substance, we examined DTF by care settings (Tx vs. HR). The rDTF was the average difference in the expected number of criteria for individuals with the same trait severity in each subgroup.

Reference

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*. Arlington, VA: American Psychiatric Association,.

Auriacombe, M. (Accessed September 2019). *ADDICTAQUI - Aquitaine Addiction Cohort: Trajectories of people with addiction (substances or behaviour) in contact with health-care system. Medical, neurobiological, sociological and psychological characteristics. Prospective multicentric, multidisciplinary study. https://epidemiologie-france.aviesan.fr/en/epidemiology/records/cohorte-addiction-aquitaine-trajectoires-de-personnes-presentant-une-addiction-aux-substances-ou-une-addiction-comportementale-en-contact-avec-le-dispositif-de-soins.-caracteristiques-medicales-neurobiologiques-sociologiques-et-psychologiques.-etude.*

Auriacombe, M., Roux, P., Vilotitch, A., Krichherr, S., Kervran, C., Chauvin, C., . . . COSINUS group. (2019). The impact of drug consumption rooms on risk practices and access to care in people who inject drugs in France: the COSINUS cohort study protocol *BMJ Open,* 9(2).

Auriacombe, M., Serre, F., Denis, C., & Fatseas, M. (2018). *Diagnosis of addictions*. In H. Pickard & S. Ahmed (Eds.), *The Routledge Handbook of the Philosophy and Science of Addiction* (pp. 132–144). London and New York

Castaldelli-Maia, J. M., Wang, Y. P., Borges, G., Silveira, C. M., Siu, E. R., Viana, M. C., . . . Andrade, L. H. (2015). Investigating dimensionality and measurement bias of DSM-5 alcohol use disorder in a representative sample of the largest metropolitan area in South America. *Drug Alcohol Depend,* 152, 123–130.

Chun, S., Stark, S., Kim, E. S., & Chernyshenko, O. S. (2016). MIMIC Methods for Detecting DIF Among Multiple Groups:Exploring a New Sequential-Free Baseline Procedure. *Applied Psychological Measurement,* 40(7), 486-499.

Chung, T., Martin, C. S., Maisto, S. A., Cornelius, J. R., & Clark, D. B. (2012). Greater prevalence of proposed DSM-5 nicotine use disorder compared to DSM-IV nicotine dependence in treated adolescents and young adults. *Addiction,* 107(4), 810–818.

Compton, W. M., Thomas, Y. F., Stinson, F. S., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Arch Gen Psychiatry,* 64(5), 566–576.

Copersino, M. L., Boyd, S. J., Tashkin, D. P., Huestis, M. A., Heishman, S. J., Dermand, J. C., . . . Gorelick, D. A. (2006). Quitting among non-treatment-seeking marijuana users: reasons and changes in other substance use. *Am J Addict,* 15(4), 297–302.

Cummings, K. M., Jaen, C. R., & Giovino, G. (1985). Circumstances surrounding relapse in a group of recent exsmokers. *Prev Med,* 14(2), 195–202.

Degenhardt, L., Charlson, F., Ferrari, A., Santomauro, D., Erskine, H., & Mantilla‐Herrara, A. (2018). The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Psychiatry,* 5(12), 987–1012.

Denis, C., Fatseas, M., Beltran, V., Serre, F., Alexandre, J. M., Debrabant, R., . . . Auriacombe, M. (2016). Usefulness and validity of the modified Addiction Severity Index: A focus on alcohol, drugs, tobacco, and gambling. *Subst Abus,* 37(1), 168–175.

Derringer, J., Krueger, R. F., Dick, D. M., Agrawal, A., Bucholz, K. K., Foroud, T., . . . McGue, M. (2013). Measurement invariance of DSM-IV alcohol, marijuana and cocaine

dependence between community-sampled and clinically overselected studies. *Addiction,* 108(10), 1767–1776.
Des Jarlais, D. C. (1995). Harm reduction--a framework for incorporating science into drug policy. *Am J Public Health,* 85(1), 10–12.
Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists.* Mahwah, NJ: Lawrence Earlbaum Associates.
Hagman, B. T. (2017). Development and psychometric analysis of the Brief DSM-5 Alcohol Use Disorder Diagnostic Assessment: Towards effective diagnosis in college students. *Psychol Addict Behav,* 31(7), 797–806.
Hasin, D. S., Fenton, M. C., Beseler, C., Park, J. Y., & Wall, M. M. (2012). Analyses related to the development of DSM-5 criteria for substance use related disorders: 2. Proposed DSM-5 criteria for alcohol, cannabis, cocaine and heroin disorders in 663 substance abuse patients. *Drug Alcohol Depend,* 122(1-2), 28–37.
Hasin, D. S., O'Brien, C. P., Auriacombe, M., Borges, G., Bucholz, K., Budney, A., . . . Grant, B. F. (2013). DSM-5 criteria for substance use disorders: recommendations and rationale. *Am J Psychiatry,* 170(8), 834–851.
Hawk, M., Coulter, R. W. S., Egan, J. E., Fisk, S., Reuel Friedman, M., Tula, M., & Kinsky, S. (2017). Harm reduction principles for healthcare settings. *Harm Reduct J,* 14(1), 70.
Hu, L., & Bentler, P. (1999). Cutoff criteriafor fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation modeling,* 6, 1–55.
Jones, R. N. (2006). Identification of measurement differences between English and Spanish language versions of the Mini-Mental State Examination. Detecting differential item functioning using MIMIC modeling. *Med Care,* 44(11 Suppl 3), S124–133.
Kervran, C., Serre, F., Denis, C., Moriceau, S., Fatseas, M., Roux, P., . . . Auriacombe, M. (2018). Diagnostic characteristics of addiction and craving among active substance users in Needle exchange programs and Supervised injection facilities: a review of the literature. *College on Problems of Drug Dependence*. San Diego, USA.
Kervran, C., Serre, F., Denis, C., Roux, P., Jauffret-Roustide, M., Lalanne, L., & Auriacombe, M. (2019). Substance Use Disorder diagnosis and craving: comparison of patient seeking treatment and active substance users in harm reduction settings. *College on Problems of Drug Dependence*. San Antonio, USA.
Kervran, C., Shmulewitz, D., Serre, F., Stohl, M., Denis, C., Hasin, D., & Auriacombe, M. (2020). Item Response Theory analyses of DSM-5 substance use disorder criteria in French outpatient addiction clinic participants. How much is craving special? *Drug Alcohol Depend,* 212, 108036.
Kessler, R. C., Molnar, B. E., Feurer, I. D., & Appelbaum, M. (2001). Patterns and mental health predictors of domestic violence in the United States: results from the National Comorbidity Survey. *Int J Law Psychiatry,* 24(4-5), 487-508.
Kidorf, M., Disney, E. R., King, V. L., Neufeld, K., Beilenson, P. L., & Brooner, R. K. (2004). Prevalence of psychiatric and substance use disorders in opioid abusers in a community syringe exchange program. *Drug Alcohol Depend,* 74(2), 115–122.
Kidorf, M., King, V. L., Peirce, J., Burke, C., Kolodner, K., & Brooner, R. K. (2010). Psychiatric distress, risk behavior, and treatment enrollment among syringe exchange participants. *Addict Behav,* 35(5), 499–503.
Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling. 3rd ed: Guilford*.
Martin, C. S., Langenbucher, J. W., Chung, T., & Sher, K. J. (2014). Truth or consequences in the diagnosis of substance use disorders. *Addiction,* 109(11), 1773–1778.

McLellan, A. T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., . . . Argeriou, M. (1992). The Fifth Edition of the Addiction Severity Index. *J Subst Abuse Treat,* 9(3), 199–213.

Metzger, D., Woody, G. E., Druley, P., DePhillips, D., Navaline, H., McLellan, A. T., & O'Brien, C. P. (1990). Psychiatric symptoms, high risk behaviors and HIV positivity among methadone patients. *NIDA Res Monogr,* 105, 490–491.

Midanik, L. T., Tam, T. W., & Weisner, C. (2007). Concurrent and simultaneous drug and alcohol use: results of the 2000 National Alcohol Survey. *Drug Alcohol Depend,* 90(1), 72-80.

Moracchini, C., Frauger, E., Pauly, V., Nordmann, S., Thirion, X., Micallef, J., & Reseau Francais des Centres d'Addictovigilance. (2012). Les CAARUD, lieux privilégiés d'émergence de signaux pour l'addictovigilance. *Therapie,* 67(5), 437–445.

Morales, L. S., Flowers, C., Gutierrez, P., Kleinman, M., & Teresi, J. A. (2006). Item and scale differential functioning of the Mini-Mental State Exam assessed using the Differential Item and Test Functioning (DFIT) Framework. *Med Care,* 44(11 Suppl 3), S143–151.

Muthen, L. K., & Muthen, B. O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Peters, E. N., & Hughes, J. R. (2010). Daily marijuana users with past alcohol problems increase alcohol consumption during marijuana abstinence. *Drug Alcohol Depend,* 106(2-3), 111–118.

Pickard, H., & Ahmed, S. (2016). *How do you know you have a drug problem? The role of knowledge of negative consequences in explaining drug choice in humans and rats.* In N. Heather & G. Segal (Eds.), *Addiction and Choice, rethinking the relationship.* (pp. 29 – 48). Oxlford, UK: Oxford university press

R Core Team. (2018). R: A Language and Environment for Statistical Computing. In R. F. f. S. Computing (Ed.). Vienna, Austria.

Raju, N., Van der Linden, W., & Fleer, P. (1995). IRT-Based internal measures of differential functionning of items and test. *Appl Pqychol Meas,* 19, 353–368.

Ross, J., Teesson, M., Darke, S., Lynskey, M., Ali, R., Ritter, A., & Cooke, R. (2005). The characteristics of heroin users entering treatment: findings from the Australian treatment outcome study (ATOS). *Drug Alcohol Rev,* 24(5), 411–418.

Ross, J., Teesson, M., Darke, S., Lynskey, M., Hetherington, K., Mills, K., . . . Fairbairn, S. (2002). *Characteristics of heroin users entering three treatment modalities in new south wales: baseline findings from the Australian Treatment Outcome Study (ATOS)* NDARC Technical Report

Serier, K. N., Venner, K. L., & Sarafin, R. E. (2019). Evaluating the Validity of the DSM-5 Alcohol Use Disorder Diagnostic Criteria in a Sample of Treatment-seeking Native Americans. *J Addict Med,* 13(1), 35–40.

Shiffman, S., Shumaker, S. A., Abrams, D. B., Cohen, S., Garvey, A., Grunberg, N. E., & Swan, G. E. (1986). Models of smoking relapse. *Health Psychol,* 5 Suppl, 13–27.

Shmulewitz, D., Greene, E. R., & Hasin, D. (2015). Commonalities and Differences Across Substance Use Disorders: Phenomenological and Epidemiological Aspects. *Alcohol Clin Exp Res,* 39(10), 1878–1900.

Shmulewitz, D., Keyes, K. M., Wall, M. M., Aharonovich, E., Aivadyan, C., Greenstein, E., . . . Hasin, D. (2011). Nicotine dependence, abuse and craving: dimensionality in an Israeli sample. *Addiction,* 106(9), 1675–1686.

Shulman, E. P., Harden, K. P., Chein, J. M., & Steinberg, L. (2015). Sex differences in the developmental trajectories of impulse control and sensation-seeking from early adolescence to early adulthood. *J Youth Adolesc,* 44(1), 1–17.

Teresi, J. A., & Fleishman, J. A. (2007). Differential item functioning and health assessment. *Qual Life Res,* 16 Suppl 1, 33-42.
Woods, C. M., & Grimm, K. J. (2011). Testing for Nonuniform Differential Item Functioning With Multiple Indicator Multiple Cause Models. *Applied Psychological Measurement,* 35(5), 339-361.
Woods, C. M., Oltmanns, T. F., & Turkheimer, E. (2009). Illustration of MIMIC-Model DIF Testing with the Schedule for Nonadaptive and Adaptive Personality. *J Psychopathol Behav Assess,* 31(4), 320–330.
Zuckerman, M., Eysenck, S., & Eysenck, H. J. (1978). Sensation seeking in England and America: cross-cultural, age, and sex comparisons. *J Consult Clin Psychol,* 46(1), 139–149.