



The Psychometric Properties of the Drug Use Disorders Identification Test (DUDIT): A Review of Recent Research[☆]



Martin Hildebrand, Ph.D., LL.M.^{*}

Private Practice, De Bilt, The Netherlands

ARTICLE INFO

Article history:

Received 15 September 2014

Received in revised form 15 December 2014

Accepted 19 January 2015

Keywords:

Assessment

Drug use

DUDIT

Psychometric properties

Screening

Review

ABSTRACT

This article reviews the reliability (internal consistency, interrater reliability), factor structure, and validity (i.e., sensitivity and specificity) of the 11-item Drug Use Disorders Identification Test (DUDIT). An extensive literature review was conducted using the Medline, Psychinfo and PubMed databases from January 2005 to October 2014. Search terms were “Drug Use Disorders Identification Test” and “DUDIT”. All articles that addressed reliability, factor structure, sensitivity and specificity of the DUDIT were examined. Publications in which the DUDIT was used as a screening tool and/or studies that used the DUDIT to compare use and abuse/dependence of drugs in participants at baseline and at follow-up but had no data on its psychometric properties, were not included. In total, 18 usable publications were identified and included in the review. In general, the DUDIT yields satisfactory measures of reliability and validity for use as a clinical or research tool. Internal consistency reliability estimates (Cronbach's α) are generally $>.90$. Most studies also revealed favorable sensitivity (ranging from .85 to 1.00) and specificity (ranging from .75 to .92) in a variety of populations. Results on the factor structure, however, are somewhat equivocal. Future directions for research on the DUDIT are suggested.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

In the last decades, several brief screening instruments have been developed to detect problematic drug use. Scores on these so-called screeners can assist professionals in a variety of settings (e.g., psychiatric hospitals, addiction treatment facilities, criminal justice, forensic psychiatric institutions) to determine if individuals are at risk for problems (health or other) related to their drug use and if further assessment and treatment is warranted. Responding to the need for a simple self-report screening instrument focusing on *current* drug-related problems and facilitating comparison between different groups, Berman, Bergman, Palmstierna, and Schlyter (2003, 2005) constructed the Drug Use Disorders Identification Test (DUDIT) as a brief screening scale to identify individuals with (recent) drug problems. The purpose of the DUDIT items is to identify use patterns and various drug-related problems. The authors chose items that would help identify individuals who appear to have a drug problem or drug dependence, as well as screening out those who do not have such problems.

Although a variety of drug use screening instruments have been developed, including the self-report 10-item Drug Abuse Screening Test (DAST; Skinner, 1982), the six-item UNCOPE (Use, Neglect, Cut down, Objection, Preoccupied, Emotional discomfort; Hoffmann, Hunt, Rhodes, & Riley, 2003), as well as brief screening interview schedules such as the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST; WHO ASSIST Working Group, 2002), the 11-item DUDIT

has gained some popularity, especially in Northern Europe. The instrument has several advantages over other instruments: its attempt to identify *at-risk* drug use, not only drug dependence; the brevity of the instrument (11 items); the continuous interval scaling instead of dichotomous (e.g., yes or no); the focus on current (i.e., past year) drug use and related consequences, thus identifying possible drug use problems; the inclusion of a drug list with illegal substances or commonly abused prescription medications for clients' and counselors' easy reference; its suitability for use in diverse settings and in public health surveys of drug use (Berman et al., 2003, 2005).

The DUDIT has been in use for approximately one decade. To the best of our knowledge, in this period, no review of its psychometric properties has been published. In this article, a concise review of the literature that addresses issues on the reliability (internal consistency, test-retest, interrater reliability), factor structure and validity (i.e., sensitivity and specificity) of the DUDIT is provided in order to (a) help professionals (clinicians, policy makers, researchers) to better understand the strengths and weaknesses of the instrument, (b) to provide them objective and comprehensive information on the DUDIT to assist in making decisions on whether or not to incorporate the DUDIT into their professional activities or settings, and (c) to point out areas in which future research should focus.

2. Literature search

The Medline, PsychINFO, and PubMed databases were searched from January 2005 to October 2014 for the terms *Drug Use Disorders Identification Test* and *DUDIT* (in title or abstract) to find potential

[☆] Author Note: Martin Hildebrand works as a consultant forensic psychologist and researcher in private practice, www.martinhildebrand.nl.

^{*} Hessenweg 123-B, 3731 JG De Bilt, The Netherlands. Tel.: + 31 30 2400465.

E-mail address: info@martinhildebrand.nl.

articles. In this way, a total of 55 potential usable articles were found. Articles that included measures of reliability, factor structure or validity (i.e., sensitivity and specificity) of the DUDIT were included in the review.^{1,2} The articles would be included if they have investigated at least one of the psychometric characteristics of the DUDIT with adults or adolescent samples, i.e., (interrater) reliability, test–retest reliability, internal consistency, factor structure, sensitivity/specificity. Articles were excluded if the DUDIT was used as a screening tool to identify problematic substance use but no data on reliability, factor structure or validity were presented (e.g., Cruce, Nordström, & Öjehagen, 2007; Wüsthoff, Waal, & Gräwe, 2014). Articles were also excluded if the DUDIT was used to only compare use and abuse/dependence of drugs in participants at baseline and at follow-up (e.g., Lange et al., 2014). In this way, a total of 15 usable publications were obtained (and 40 articles were excluded). Examination of the reference sections of these 15 studies did not lead to inclusion of additional relevant studies. However, due to our work on other studies, two more relevant studies were identified. Also, one study that was accepted for publication at the time of preparing this review was included. Thus, 18 usable publications were identified and included in the review.

3. Administration and scoring of the DUDIT

The DUDIT is an 11-item screening instrument developed to identify non-alcohol drug use patterns and various drug-related problems in the general public as well as in individuals in clinical settings who are likely to meet criteria for a substance dependence diagnosis (Berman et al., 2005). The first nine items are scored on a 5-point scale ranging from 0 to 4, and the last two are scored on a 3-point scale with values 0, 2, and 4. Thus, total scores range from 0 to 44, with higher scores suggestive of a more severe drug problem. The DUDIT cutoff score for any type of problematic use (i.e., harmful use, substance abuse, or dependency) is generally recommended as 6 for men and 2 for women (Berman et al., 2005; also Cruce & Öjehagen, 2007), with different scores considered optimal depending on the population studied. The following *risk levels* are suggested for the DUDIT scores: no drug related problems (total scores 0–5/1); possible drug related problems, i.e., risky or harmful drug habits that might be diagnosed as substance abuse/harmful use or dependence (6/2–24); probably heavily dependent on drugs (scores ≥ 25) (Berman et al., 2005). In addition to items covering amount and frequency of use and signs of hazardous use or dependence, the participant is asked about which drugs he or she is using, including illicit drugs and prescription drugs of abuse. The DUDIT was initially evaluated in a general Swedish population sample and validated in a sample of drug users in prison and in a detoxification setting (Berman et al., 2005). The DUDIT may be administered either as a self-report questionnaire or as an oral interview.

¹ Berman, Palmstierna, Källmén, and Bergman (2007) also created an extended version of the DUDIT, named DUDIT-E ('E' stands for 'Extended'). This 54-item instrument was developed to provide detailed information about drug use, including the individuals' perceptions about use. The DUDIT-E contains additional questions about treatment readiness and drug-related consequences (positive and negative) for individuals who have been identified by the DUDIT as possibly having a drug problem. In this way, ambivalent cognition towards change, which is considered important in enhancing problem awareness and in relation to motivational interventions, is acknowledged (Berman et al., 2007). The few studies that examined the psychometric properties of the DUDIT-E were not included in this review.

² We like to point out that another research group (Stuart, Moore, Kahler, & Ramsey, 2003; Stuart, Moore, Kahler, Ramsey, & Strong, 2004; Stuart, Moore, Ramsey, & Kahler, 2003, 2004; see also Stuart et al., 2008) developed a 14-item brief drug abuse screening instrument that they called the Drug Use Disorders Identification Test (DUDIT) – as did Berman and colleagues. To the authors' knowledge there are no peer reviewed published reports describing this version's development and/or psychometric properties (i.e., factor structure, validity), although we did find a few studies that reported on the internal consistency of this 14-item DUDIT. These studies were also excluded from this review.

4. Results

4.1. Reliability

In Table 1, data on the internal consistency and test–retest reliability is summarized. Internal consistency reliability estimates have been reported in 16 studies including 20 (sub)samples, and covered seven countries: Sweden (Berman et al., 2005; Durbeej et al., 2010; Sinadinovic, Berman, Hasson, & Wennberg, 2010; Sinadinovic, Berman & Wennberg, 2011), Norway (Gundersen, Mordal, Berman, & Bramness, 2013; Landheim, Bakken, & Vaglum, 2006; Lobmaier, Berman, Gossop, & Ravndal, 2013; Nesvåg et al., 2010), Turkey (Evren, Ogel, Evren, & Bozkurt, 2014; Evren, Ovali, Karabulut, & Cetingok, 2014), the Netherlands (Hildebrand & Noteborn, 2015; Hillege, Das, & de Rooter, 2010), USA (Reddy, Dick, Gerber, & Mitchell, 2014; Voluse et al., 2012), Hungary (Matuszka et al., 2014), and South Africa (Martin, Viljoen, Kidd, & Seedat, 2014). The study populations included psychotic patients (Gundersen et al., 2013; Nesvåg et al., 2010), prison inmates (Berman et al., 2005, Evren, Ogel et al., 2014; Lobmaier et al., 2013), probationers (Hildebrand & Noteborn, 2015), youth/adolescents (Hillege et al., 2010; Martin et al., 2014), drug addicts treated in a hospital (Berman et al., 2005), drug abusers who participated in outpatient and/or residential treatment programs (Evren, Ovali et al., 2014; Landheim et al., 2006; Matuszka et al., 2014; Voluse et al., 2012), suspected offenders detained at a forensic psychiatric assessment unit (Durbeej et al., 2010), women with posttraumatic stress disorder symptoms randomly assigned to an intervention on alcohol and drug abuse risk (Reddy et al., 2014), and individuals from the general population (Berman et al., 2005; Sinadinovic et al., 2010, 2011). Internal consistency reliability (Cronbach's α) ranged from .74 to .97, with most studies reporting a Cronbach's $\alpha > .90$ (Table 1). According to Nunnally (1978), a Cronbach's $\alpha > .70$ is one of the standard definitions of adequate reliability. In addition, Bland and Altman (1997) considered Cronbach's α of .70–.80 satisfactory for a reliable comparison between groups although for clinical purposes, they advise a minimum of .90.

Five studies (Berman et al., 2005; Evren, Ogel et al., 2014; Evren, Ovali et al., 2014; Hildebrand & Noteborn, 2015; Matuszka et al., 2014) investigated the item–total correlation, i.e., correlations of individual items on the DUDIT with the DUDIT total score. In a reliable scale all items should correlate with the total score. The results are encouraging: all data have item–total correlations above .30. In their initial investigation of the psychometric properties of the DUDIT, Berman et al. (2005) reported the lowest item–total correlation for item 3 “How many times a day do you take drugs on a typical day when you use drugs?” ($r = .38$) in their drug user sample. Others (Evren, Ogel et al., 2014; Matuszka et al., 2014) reported the lowest item–total correlation for item 9 “How often over the past year have you had guilt feelings or a bad conscience because you used drugs?” ($r = .66$ and $r = .51$, respectively) in a sample of prisoners (Evren, Ogel et al., 2014) and a mixed sample of outpatient treatment program participants, mandatory drug treatment program participants, young adults at risk of drug use, and students (Matuszka et al., 2014). Evren, Ovali et al. (2014) reported the lowest item–total correlation for item 2 “Do you use more than one type of drug on the same occasion?” in a sample of inpatients with drug use disorders ($r = .47$), whereas Hildebrand and Noteborn (2015) reported the lowest item–total correlation ($r = .45$) for item 10 “Have you or anyone else been hurt (mentally or physically) because you used drugs?” in a sample of probationers. Three studies (Berman et al., 2005; Evren, Ogel et al., 2014; Evren, Ovali et al., 2014) reported the highest item–total correlation for item 8 “How often over the past year have you needed to take a drug the morning after heavy use the day before?” (r between .73 and .83). The other two studies reporting on the item–total correlation (Hildebrand & Noteborn, 2015; Matuszka et al., 2014) both reported the highest item–total correlation for item 5 “Over the past year, have you felt that your longing for drugs was so strong that you could not resist it?” ($r = .83$) and item 6 “Has it

Table 1
Reliability (internal consistency and test–retest reliability) of the Drug Use Disorders Identification Test.

Year	Authors	Subjects	Internal consistency (Cronbach's α)	Test–retest reliability
2005	Berman et al.	154 (heavy) drug users from prison, probation, and inpatient detoxification settings, age range 18–74 years (17% women)	.80	
		1109 randomly selected individuals (16 years and up) from the general population (+/- 52% women)	.93	
2006	Landheim et al.	160 patients who received treatment for substance abuse (29% women)	.95	
2010	Durbbeej et al.	181 suspected offenders with signs of mental health problems, age 18–60 years (9% women)	.94	
2010	Hillege et al.	740 adolescents in the upper grades of two secondary schools (53% female)	.92	
2010	Nesvåg et al.	205 first-episode psychosis patients, age between 18 and 65 years (36% women)	.93 (M) .96 (F)	
2010	Sinadinovic et al.	1211 individuals (15 years and up) who created a valid account on a web-based service offering self-assessment and self-monitoring of alcohol and drug use via on-line screening (49% women)	.97 (retest .96, $n = 97$)	.05 ($n = 74$)
2011	Sinadinovic et al.	1833 individuals (16 years and up) randomly selected from the general population (53% women)	.86 (internet version) .85 (IVR)	
2012	Voluse et al.	153 substance abusers in outpatient and residential treatment (22% women)	.94	
2013	Gundersen et al.	161 emergency psychiatric patients (55% women)	.94/.97 (M) ^a .96/.90 (F) ^a	
2013	Lobmaier et al.	110 male drug-involved prison inmates, age 18–61 years	.94	
2014	Evren, Ogel et al.	202 male prisoners	.93	
2014	Evren, Ovali et al.	258 inpatients with drug use disorder, including 100 male adolescents, 35 male alcohol dependent adults, 123 adult heroin dependents (9% women)	.93	.77 ($n = 123$)
2014	Martin et al.	1149 randomly selected school-attending youths ($M = 16.2$ years) from secondary schools (60% women)	.89	
2014	Matuszka et al.	98 outpatient treatment program participants (16% women)	.90 (retest .88, $n = 53$) ^b	.91 ($n = 53$) ^b
		85 mandatory drug treatment program participants (12% women)	.85	
		76 young adults at risk of drug use (34% women)	.88	
		109 students (40% women)	.77	
		Total sample ($N = 368$)	.92	
2014	Reddy et al.	38 women with PTSD symptoms randomly assigned to a yoga intervention on alcohol and drug abuse risk	.74	
2015	Hildebrand and Noteborn	358 probationers, 18 years and up (12% women)	.94	

Note. ICC = Intraclass correlation coefficient. M = male. F = female. IVR = interactive voice response. Test–retest reliability was calculated using Pearson's r (Evren, Ovali et al., 2014; Sinadinovic et al., 2010) or ICC (Matuszka et al., 2014).

^a Internal consistency reliabilities for patients with current psychosis and no current psychosis, respectively.

^b Cronbach's α at retest and test–retest reliability investigation (at a two-week interval) were performed in the outpatient treatment program participants and mandatory drug treatment program participants subsamples, and comprised 53 patients.

happened, over the past year, that you have not been able to stop taking drugs once you started?" ($r = .83$; Hildebrand & Noteborn, 2015).

Three studies were identified that investigated the test–retest reliability of the DUDIT total score (Evren, Ovali et al., 2014; Matuszka et al., 2014; Sinadinovic et al., 2010). Sinadinovic et al. (2010) developed a web-based service, called eScreen, which offers, among other things, a web-based version of the DUDIT. Test–retest reliability (Pearson's r) was calculated for those individuals from the general population who chose to retest themselves within a week ($n = 74$). The result was very disappointing, $r = .05$. The most plausible explanation for these disappointing results, according to the authors, is that different individuals used the same user account within a brief time interval. The authors also considered the possibility that with the first test users tried the site (to see how it works and what kinds of questions were asked) and that the second test reflects their genuine responses. This second explanation is considered to be less likely, because it would have generated lower internal consistency reliability coefficients (Sinadinovic et al., 2010). In contrast, Matuszka et al. (2014) reported a test–retest Intraclass Correlation coefficient (ICC) of .91 for the DUDIT in a subsample of outpatient treatment program participants and mandatory drug treatment program participants ($n = 53$) who were retested two weeks after the initial administration of the test. Evren, Ovali et al. (2014) reported a test–retest coefficient (Pearson's r) of .77 for the DUDIT in a subsample of heroin dependent inpatients ($n = 123$). Unfortunately, these authors did not report the interval between test and retest.

4.1.1. Interrater reliability

So far, interrater reliability of DUDIT scores is examined in only one study. Hildebrand and Noteborn (2015) reported initial evidence for the interrater reliability of the DUDIT scores when administered as an oral

interview using a joint interview approach. In a sample of 56 probationers, interrater reliabilities of most of the DUDIT items demonstrated to be acceptable to good, with Krippendorff's alpha ($K\alpha$) $> .67$ for eight of the 11 individual items. The interrater reliability of the DUDIT total score, in terms of risk level (0–5 = no drug related problems; 6–24 = possible drug related problems; ≥ 25 = probable drug dependence) was found to be acceptable as well ($K\alpha = .76$).

4.2. Factor structure

To date, factor analyses of the DUDIT have been reported in six studies (Berman et al., 2005; Evren, Ogel et al., 2014; Evren, Ovali et al., 2014; Hildebrand & Noteborn, 2015; Matuszka et al., 2014; Voluse et al., 2012), covering five countries: Sweden, Turkey, Hungary, USA, and the Netherlands. Table 2 summarizes the results of these factor analytic studies. Berman et al. (2005) used a general population sample ($N = 1109$) and a drug user sample ($N = 154$) in their initial investigation of the psychometric properties of the DUDIT. For the drug user sample (either inpatients at an addiction detoxification unit, prison detainees, prison inmates, or probation clients), principal axis factoring (PAF) with oblique rotation yielded three factors labelled *Dependence* (items 1, 4–8; factor loadings 0.48, 0.56, 0.82, 0.60, 0.74, and 0.90, respectively), *Drug-related problems* (items 2, 10, 11; factor loadings 0.76, 0.70, and 0.67) and *Intensity of use* (items 3 and 9; factor loadings 0.66 and -0.54). The authors did not report eigenvalues nor did they report the percentage of explained variance for this sample. For the general population sample, PAF with oblique rotation yielded two factors explaining 73.8% of the variance; the first factor explained 58.9% of the variance and reflects *Dependence on drugs* (items 2, 4–8), the second factor explained 14.9% of the variance and reflects

Table 2
Factor structure of the Drug Use Disorders Identification Test.

Year	Authors	Subjects	Methods	Number of components/factors
2005	Berman et al.	154 (heavy) drug users from prison, probation, and inpatient detoxification settings, between 18–74 years (17% women) 1109 randomly selected individuals (16 years and up) from the general population (+/– 52% women)	PAF/Oblique	3 labeled <i>Dependence</i> (items 1, 4–8), <i>Drug-related problems</i> (items 2, 10–11), <i>Intensity of use</i> (items 3, 9)
			PAF/Oblique	2 (58.9% and 14.9%) ^a labeled <i>Dependence</i> (items 2, 4–8) and <i>Drug-related problems</i> (items 1, 3, 9–11)
2012	Voluse et al.	153 substance abusers in outpatient and residential treatment (22% women)	PCA	1 (64.9%) ^a
2014	Evren, Ogel et al.	202 male prisoners	PCA/Varimax	1 (58.3%) ^a
2014	Evren, Ovali et al.	258 inpatients with drug use disorder (100 male adolescents, 35 male alcohol dependent adults, 123 adult heroin dependents (9% women))	CFA 1 F	Support for unidimensional construct
			PCA/Varimax	2 (58.7% and 10.2%) ^a consisting of items 4–11 and items 1–3, respectively
2014	Matuszka et al.	259 participants (98 treatment program patients; 85 mandatory drug treatment program patients; 76 young adults at risk of drug use) (20% women)	CFA 1 F	Support for unidimensional construct
			PCA/Varimax	2 (51% and 10.2%) ^a labeled <i>Drug use</i> (items 4–11) and <i>Frequency of drug use</i> (items 1–3)
2015	Hildebrand and Noteborn	358 probationers, 18 years and up (12% women)	CFA 1 vs. 2 vs. 3 Fs	No convergence for 3, 2 slightly better 1 but no good fit
			PAF/Oblique	1 (62.6%) ^a

Note. PAF = Principal axis factoring. PCA = principal component analysis. CFA = confirmatory factor analysis. F = factor. Fs = factors.

^a Explained variance.

Drug-related problems (items 1, 3, 9–11), according to the authors. However, according to the authors, the factor analysis in the general population sample is questionable because of highly skewed positive distribution and probable selective dropout of individuals with higher DUDIT scores (Berman et al., 2005). Eigenvalues and factor loadings were not reported.

Matuszka et al. (2014) also examined the DUDIT using exploratory techniques. They employed data from 259 participants (98 treatment program patients, 85 mandatory drug treatment program patients, and 76 young adults at risk of drug use) and found that principal component analyses (PCA) with varimax rotation yielded a two-factor solution that was similar (to an extent) to that produced by Berman et al. (2005) in their general population sample. The two factors, taken together, accounted for 61.2% of the total variance. The first factor explained 51% of the variance and reflects *Problems associated with drug use* (items 4–11; factor loadings 0.68, 0.81, 0.82, 0.70, 0.79, 0.68, 0.43, and 0.50, respectively); the second factor explained 10.2% of the variance and resembles *Frequency of drug use* (items 1–3; factor loadings 0.73, 0.79, 0.83). Eigenvalues were not reported by the authors. The factor solution identified by Matuszka et al. (2014) seems to more clearly separate the items reflecting substance use problems from items indicating the frequency of drug use than the two-factor (general population sample) or three-factor (drug user sample) solutions identified by Berman et al. (2005). According to Matuszka et al. (2014), this may be due to the different characteristics of the study populations.

Voluse et al. (2012) also performed a PCA for their sample of 153 substance abusers in outpatient and residential treatment. They found that only one component reached the criterion of an eigenvalue > 1 (7.14) and the variance accounted for by this component was 64.9%. The next largest eigenvalue (0.81) only accounted for 7.23% of the total variance. This led the authors to conclude that the DUDIT produced a unidimensional construct in their sample. Item–component loadings were all in the “very good” to “excellent” range (according to the suggested cutoff for size of loading by Comrey and Lee, 1992; see also Tabachnik and Fidell, 2007, p. 649) varying from 0.63 (item 2) to 0.90 (item 7 “How often over the past year have you taken drugs and then neglected to do something you should have done?”). Furthermore, nine of the 11 items had item–component loadings > 0.71, the suggested cutoff for excellent loading (Comrey & Lee, 1992; Tabachnik & Fidell, 2007).

Recently, Evren, Ogel et al. (2014), using PCA with varimax rotation for their sample of 202 prisoners, also found that only one component reached the criterion of an eigenvalue > 1 (6.41) and the variance accounted for by this component was 58.3%. Subsequently, the authors examined the unidimensionality of the DUDIT using confirmatory factor analysis (CFA), which

provided further support for the unidimensional structure of the DUDIT ($\chi^2/df = 1.13$; root mean square error of approximation [RMSEA] = 0.026; goodness of fit index [GFI] = 0.965, adjusted GFI = 0.933; parsimony GFI = 0.511; normed fit index [NFI] = 0.972; comparative fit index [CFI] = 0.997; incremental fit index [IFI] = 0.997). Item–component loadings were all in the “very good” to “excellent” range, varying from 0.65 (item 9) to 0.81 (item 8), and nine of the 11 items had loadings > .71. In another recent study by Evren, Ovali et al. (2014), the factor structure of the DUDIT was investigated in a sample of 258 adolescent and adult inpatients with a drug use disorder ($n = 223$) or alcohol dependence ($n = 35$). PCA with varimax rotation yielded two factors explaining 68.85% of the variance; the first factor explained 58.65% of the variance (items 4–11; factor loadings 0.64, 0.71, 0.76, 0.72, 0.73, 0.67, 0.81, and 0.85, respectively) and the second factor explained 10.2% of the variance (items 1–3; factor loadings 0.81, 0.85, and 0.59). Since the eigenvalue of the first factor (6.45) was more than three times larger than the eigenvalue of the second factor (1.12), the authors stated that “the output for the two-factor solution indicates a unidimensional construct for the measure” (p. 42). The assumed unidimensionality was further assessed using CFA which produced a good fit (RMSEA = 0.066; GFI = 0.941; adjusted GFI = 0.945; parsimony GFI = 0.585; NFI = 0.954; CFI = 0.975; IFI = 0.975). Item–component loadings were all in the “fair” to “excellent” range, varying from 0.53 (item 2) to 0.87 (item 8), with eight of the 11 items having item–component loadings > 0.71.

Finally, Hildebrand and Noteborn (2015) used CFA to test *a priori* one, two, and three-factor models using DUDIT data gathered from 358 probationers. In Model 1 (unidimensional model), implied by the scoring procedure of the DUDIT and supported by Voluse et al. (2012), all items were specified to load on a single factor. Model 2 (two-factor) included a *Dependence on drugs* factor and a *Drug-related problems* factor, and was supported by the results of Berman et al. (2005) in their general population sample. Model 3 (three-factor), supported by the results of the drug user sample of Berman et al. (2005) specified the following factors: *Drug dependence*, *Drug related problems*, and *Intensity of use*. The results showed that the data did not support the structure specified of the three-factor model, and the one- and two-factor solutions produced somewhat equivocal results. Both models were rejected based on the chi-square test statistic (one-factor model: $\chi^2_{44} = 384.18$, $p < .0001$; two-factor model: $\chi^2_{43} = 323.73$, $p < .0001$), the χ^2/df ratio (8.73 and 7.53, respectively), and RMSEA values (0.147 and 0.135, respectively), but all other indicators of model fit (IFI, CFI, Tucker–Lewis Index [TLI], standardized root mean square residual [SRMR]) (almost) reached the prespecified cutoff points (CFI, IFI, and TLI ≥ 0.95 ; SRMR $\leq .08$) for both

Table 3
Validity of the DUDIT: sensitivity, specificity, positive predictive value, negative predictive value, and area under the curve (AUC).

Year	Authors	Subjects	Diagnostic measure	Sensitivity (cutoff)	Specificity (cutoff)	PPV	NPV	AUC
2005	Berman et al.	154 (heavy) drug users from prison, probation, and inpatient detoxification settings (17% women)	DSM-IV/ICD-10 DUD ^a : harmful use/abuse dependence	- .90/.90 ^b (25)	- .78/.88 ^b (25)	-	-	.50/.56 ^b .94/.95 ^b
2010	Durbeej et al.	181 suspected offenders with signs of mental health problems (9% women)	DSM-IV DUD ^c : abuse dependence	- .85 (12)	- .85 (12)	-	-	.61 .93
2010	Nesvåg et al.	205 first-episode psychosis patients (36% women)	Any DSM-IV DUD ^d	.93/1.00 (3 M/1 F)	.75/.77 (3 M/1 F)	.65/ .50 ^e	.96/ .98 ^e	.97 (M) .97 (F)
2011	Berg and Asskildt	74 acute patients in a psychiatric intensive care unit (% women NR)	ICD-10 DUD diagnoses from medical records	.78	.13	-	-	-
2012	Kader et al.	30 HIV positive patients attending HIV community health clinic (% women NR)	Biomarkers: hair urine	.00 (6 M/2 F) 1.00 (6 M/2 F)	.70 (6 M/2 F) .66 (6 M/2 F)	.00 .10	1.00 1.00	- -
2012	Voluse et al.	153 substance abusers in outpatient and residential treatment (22% women)	DAST-10 ^f	.90 (8)	.85 (8)	-	-	.95
2013	Gundersen et al.	161 emergency psychiatric patients (55% women)	Any DSM-IV DUD ^g	1.00/.92 ^h (M) (5) 1.00/1.00 ⁱ (F) (1)	.88/.85 ^h (M) (5) .92/.85 ⁱ (F) (1)	-	-	≥.93 (M) ≥.96 (F)
2014	Evren, Ogel et al.	202 male prisoners	Any drug use disorder ^j	.95 (10)	.79 (10)	.88	.80	.95
2014	Evren, Ovali et al.	258 inpatients with drug use disorder (9% women)	Any drug use disorder ^j	.96 (10)	.94 (10)	.99	.81	.975
2014	Matuszka et al.	98 outpatient treatment program participants (16% women) and 109 students (60% women)	ICD-10 substance dependence diagnoses from medical staff ^k	.95 (2.1)	.81 (2.1)	-	-	.96

Note. DUDIT = Drug Use Disorders Identification Test. PPV = positive predictive value. NPV = negative predictive value. DUD = drug use disorder. NR = Not reported. M = male. F = female.

^a According to SCAN diagnostic interviews.

^b Against DSM-IV and ICD-10 criteria, respectively.

^c According to SCAN ($n = 16$) or Mental disorder, Substance Abuse and Crime Schedules (MSAC; $n = 165$) diagnostic assessment tools.

^d DSM-IV DUD diagnoses were based on the SCID-I (Structured Clinical Interview for DSM-IV, patient version, modules A–E). The validation analysis included patients with symptoms of abuse or dependence within the last year only.

^e Values for optimal cut-off scores three or above for men and one or above for women. For PPV and NPV values for each cut-off score of the DUDIT, see Nesvåg et al. (2010).

^f Participant's scores on the 10-item version of the DAST were dichotomously classified, with scores ≥ 3 being indicative of the presence of a drug problem, and scores of 2 or less being indicative of the absence of a drug problem.

^g DSM-IV DUD diagnoses were determined based on the MINI.

^h Sensitivity and specificity for psychotic and nonpsychotic men, respectively.

ⁱ Sensitivity and specificity for psychotic and nonpsychotic women, respectively.

^j According to SCID-I interview

^k Not reported in the article, information obtained from B. Matuszka (personal communication, December 12, 2014)

models, suggesting adequate fit. To identify a solution that may be more consistent with the data, the authors conducted a PAF analysis (oblique rotation) and found that only one factor on the DUDIT reached the criterion of an eigenvalue > 1 (6.89), and the variance accounted for by this factor was 62.61%. All factor loadings were in the “fair” to “excellent” range, and varied from 0.46 (item 10) to 0.87 (item 6), with nine of the 11 items having factor loadings > 0.71 .

4.3. The DUDIT against diagnostic criteria: sensitivity and specificity

Screening instruments like the DUDIT do not provide a formal (DSM or ICD) diagnosis, but may be useful to indicate the presence of drug abuse or drug dependence. Indeed, one of the goals of using the DUDIT is to better enable professionals (e.g., clinicians) to separate individuals into groups. These groups comprise those who are identified as individuals with drug abuse or drug dependence, as well as those who do not have such problems. Since the DUDIT has a noncategorical structure, analyses of sensitivity and specificity are needed to decide on optimal diagnostic cutoff scores. However, these cutoff scores may vary between settings due to differences in prevalence and severity of the drug-related problems (see, for example, Berman et al., 2005; Nesvåg et al., 2010) and different cutoff scores may be considered optimal depending on the population studied (e.g., Gundersen et al., 2013). The sensitivity (true positive rate) of the DUDIT reflects its accuracy in identifying diagnosed drug abusers. Specificity (true negative rate) is the rate at which the test correctly identifies nondrug abusers. Based on sensitivity and specificity proportions, an optimal cutoff score can be chosen for the outcome variable. Table 3 summarizes the results of the 10 studies (Berg &

Asskildt, 2011; Berman et al., 2005; Durbeej et al., 2010; Evren, Ogel et al., 2014; Evren, Ovali et al., 2014; Gundersen et al., 2013; Kader, Seedat, Koch, & Parry, 2012; Matuszka et al., 2014; Nesvåg et al., 2010; Voluse et al., 2012), covering six countries (Sweden, Turkey, Hungary, Norway, South Africa, USA), that examined the sensitivity and the specificity of the DUDIT using various external validating measures.

Using different cutoff scores, as well as different methods to establish diagnostic criteria (Table 3), eight of the 10 studies revealed favorable sensitivity (range from .85 to 1.00) and specificity (range from .75 to .92). Berman et al. (2005) used the Schedules for Clinical Assessment in Neuropsychiatry (SCAN; World Health Organization, 1999) for validity diagnosis of harmful use/substance abuse and dependence according to both ICD-10 and DSM-IV diagnoses in their sample of 154 heavy drug users. Receiver Operating Characteristic (ROC)³ curves

³ ROC analysis computes an area under the curve (AUC) statistic by plotting the sensitivity of a tool (e.g., the DUDIT) against its specificity (Mossman, 1994). The AUC score can range from 0 (perfect negative prediction) to 1 (perfect positive prediction); an AUC of .5 indicates that the tool is not able to predict any better than chance. The most common interpretation of the AUC is the probability that a randomly chosen individual from the positive group (here: individuals with a drug abuse or dependence diagnosis) has a test score (DUDIT score) higher than a randomly chosen individual from the negative group (participants with no diagnosis). The ROC analysis also produces the sensitivity and the specificity over all possible cutoff scores of a screening tool. When choosing a cutoff score, adequate sensitivity must be ensured to exclude false positives (e.g., the proportion of participants that are falsely recognized by the DUDIT in fulfilling criteria for a diagnosis) since screening is intended to capture as many affected individuals as possible. However, the sensitivity must be balanced against the specificity for minimizing the false negatives (e.g., the proportion that are falsely recognized by the DUDIT as not fulfilling criteria for a diagnosis).

were calculated in order to identify cutoff values for maximum sensitivity and specificity. For the ICD-10 dependence syndrome, the AUC value was .95. At a DUDIT cutoff score of 25 points, sensitivity was .90 and specificity was .88. For the DSM-IV dependence diagnosis, the AUC value was .94 and a cutoff score of 25 points yielded .90 sensitivity and (less than optimal) specificity of .78. For the diagnoses of harmful use and substance abuse the AUC value was .56 and .50 respectively, indicating that the DUDIT did not predict these diagnoses better than chance.

The sensitivity and specificity of the DUDIT has also been tested among 153 substance users in outpatient and residential settings (Voluse et al., 2012). Results revealed that the AUC (.95) was in the “excellent” range, and, using the optimal cutoff score of 8 in their sample, sensitivity and specificity scores of .90 and .85, respectively, were found against the 10-item DAST. The authors did not differentiate between men and women. In fact, only two studies (Gundersen et al., 2013; Nesvåg et al., 2010) reported separate sensitivity and specificity scores for men and women. Another study, conducted by Evren, Ovali et al. (2014), also used a sample of inpatients with drug use disorders ($N = 258$) to examine the sensitivity and specificity of the DUDIT. The ROC analysis revealed that the DUDIT could screen for any drug use disorder in this population, with $AUC = .975$. At the optimal cut-point of 10, the DUDIT had a sensitivity of .96 and a specificity of .94, positive predictive power of .99, and negative predictive power of .81, against the SCID-I criterion of any drug use disorder. Matuszka et al. (2014) investigated the predictive validity of the DUDIT in a sample of outpatient treatment seekers with a high severity of drug use ($n = 98$; positive cases) and students ($n = 109$; control group, negative cases). The AUC was .96 for the separation of these groups. Results indicated that a (low) score of 2.1 points was optimal in this sample for identifying drug users, with sensitivity and specificity scores of .95 and .81, respectively.

Another subgroup receiving research attention is the group of patients with mental disorders, a group that tends to have a high prevalence of drug use disorders (e.g., Addington & Addington, 2007; Margolese, Malchy, Negrete, Tempier, & Gill, 2004; Regier et al., 1990). Among 205 first-episode psychosis patients, Nesvåg et al. (2010) found that, at the suggested cut-point of 6 for men and 2 for women, the DUDIT had a sensitivity of .84 for men and 1.00 for women, and a specificity of .86 for men and .80 for women, against the Structured Clinical Interview for DSM-IV (SCID-I; First, Spitzer, Gibbon, & Williams, 1995) criterion of any drug use disorder. With the cut-point lowered to 3 for men, sensitivity improved to .93 but sensitivity decreased to .75. A lower cutoff score of 1 for women implied a sensitivity of 1.00 and a specificity of .77. DUDIT scores of 3 or above for men and 1 or above for women met the author’s preset criteria (i.e., sensitivity above .80 and specificity above .70) for optimal cutoff scores. The DUDIT also performed favorably among 161 emergency psychiatric patients, including patients with current psychosis symptoms (Gundersen et al., 2013). The authors differentiated between patients with and without current psychotic symptoms, and between men and women. Drug abuse and dependence diagnoses according to the DSM-IV were determined based on the Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). For the DUDIT in relation to drug use disorders, the AUC value was .92 or above in psychotic and nonpsychotic men and women. For psychotic men, at a DUDIT score of 5, the analysis yielded a sensitivity of 1.00 and specificity of .88. For nonpsychotic men, a DUDIT score of 1 implied a sensitivity of .92 and a specificity of .85. For psychotic women, at a DUDIT score of 1, the analysis yielded a sensitivity of 1.00 and specificity of .92. For nonpsychotic women, a DUDIT score of 1 implied a sensitivity of 1.00 and a specificity of .85. Gundersen et al. (2013) found that across all groups it was even possible to elevate the cutoff scores considerably above the conventional scores without violating the limit of a sensitivity and specificity at or above .80.

Durbeej et al. (2010) explored the concurrent validity of the DUDIT in relation to substance abuse and dependency diagnoses in a sample of 181 suspected offenders with signs of mental health problems. Drug abuse and dependence diagnoses were determined based on the

SCAN ($n = 16$) or the Mental disorder, Substance Abuse and Crime (MSAC) diagnostic assessment tool ($n = 165$). The ROC analysis revealed that the DUDIT could screen for dependency diagnoses of drugs in this population ($AUC = .93$). At a DUDIT cutoff score of 12 points, the analysis yielded a sensitivity and specificity of .85. With regard to screening for a drug abuse disorder, an unsatisfactory AUC of .61 was found suggesting that it is more difficult to identify a drug abuse diagnosis with the DUDIT than correctly identifying a dependency diagnosis, a result that was also reported by Berman et al. (2005). Evren, Ogel et al. (2014) found evidence for the predictive validity of the DUDIT in a sample of 202 prisoners. With regard to screening for any drug use disorder, an AUC value of .95 was found. At the optimal DUDIT cutoff score of 10 points, sensitivity was .95, specificity .79, positive predictive power .88, and negative predictive power .91 (against the SCID-I criterion of any drug use disorder).

Berg and Asskildt (2011) compared scores on the DUDIT with substance abuse diagnoses stated in medical records (primary and secondary diagnoses were recorded by the first author from the hospital medical records after patients were discharged) in a sample of 74 acute patients in a psychiatric intensive care unit. The sensitivity and specificity of the DUDIT in this sample were .78 and .13, respectively. Because of the discrepancies found between diagnostic confirmation of substance abuse and the scores on the DUDIT, the authors concluded that “such tests may improve diagnostic accuracy” (p. 15). However, as stated before and summarized above, the other five studies examining the sensitivity and specificity of the DUDIT against DSM-IV or ICD-10 criteria found quite favorable results, with sensitivity $\geq .85$ and specificity $\geq .75$, thus showing evidence for accurate diagnostic accuracy.

Finally, we found one study that explored the relative utility of the DUDIT compared to selected biomarkers for drug use disorders. In a preliminary investigation of the association between the DUDIT and biomarkers of drugs in urine and hair among 30 HIV positive patients who attended an HIV community health clinic, Kader et al. (2012) found that, at the suggested cut-points of 6 for men and 2 for women, sensitivity and specificity were .00 ($n = 0$ for both true positives and true negatives) and .70, respectively (hair), and 1.00 and .66, respectively (urine). However, due to the very small sample size, these results should be interpreted with caution.

5. Discussion

In this article, a concise review of the literature that addressed issues on the reliability, factor structure, sensitivity and specificity of the DUDIT, a screening instrument for detecting problematic drug use/abuse that has gained some popularity in the last decade, especially in Northern Europe, is provided. In terms of internal consistency, all studies that have reported Cronbach’s α for the DUDIT found that value to be above .70, which is one of the standard definitions of adequate reliability (Nunnally, 1978). In fact, most of the studies revealed internal consistencies close to .90 or above, which suggests that the DUDIT is a homogeneous screening tool, in diverse samples, in a broad range of settings, and in different countries. Test–retest reliability has only been investigated in three studies, and two of these studies provides preliminary evidence of short-term stability of the DUDIT (Evren, Ovali et al., 2014; Matuszka et al., 2014). Although the DUDIT was also designed to be applied by raters and not only as a self-report instrument, thus far, only one study examined the interrater reliability of DUDIT scores. Hildebrand and Noteborn (2015) reported initial evidence for the interrater reliability when the DUDIT was administered as an oral interview using a joint interview approach.

Results are equivocal when it comes to the factor structure of the DUDIT. Berman et al. (2005) reported a three-factor solution in a relatively small sample of drug users ($N = 154$) in their initial investigation of the psychometric properties of the DUDIT. They also explored the factor structure in a large ($N = 1109$) general population sample, and for this sample a two-factor solution was found. Matuszka et al. (2014)

also found a two-factor solution that was similar to an extent to that produced by Berman et al. (2005) in their general population sample. In contrast, the other four studies that investigated the factor structure of the DUDIT (Evren, Ogel et al., 2014; Evren, Ovali et al., 2014; Hildebrand & Noteborn, 2015; Voluse et al., 2012), found support for a unidimensional construct in samples of substance abusers in outpatient and residential treatment (Voluse et al., 2012), probationers (Hildebrand & Noteborn, 2015), prisoners (Evren, Ogel et al., 2014) and inpatients with drug use disorder (Evren, Ovali et al., 2014). Of the factor analytic studies included in this review, five of six samples had sizes between 153 and 358 subjects (Table 2). As stated by Tabachnick and Fidell (2007), factor analysis is highly sensitive to the size of correlations, and correlation coefficients tend to be less reliable when estimated from small samples. In other words, the reliability of factor analysis will depend on sample size and therefore it is important that the overall sample size be large enough that correlations are reliably estimated (Tabachnick & Fidell, 2007). They argue that “it is comforting to have at least 300 cases” (p. 613) to provide a stable factor solution. Comrey and Lee (1992) also give as a guide sample size of 300 as good – and 100 for poor and 1000 for excellent reliabilities. Only one study in our review had more than 1000 subjects but, as described earlier, the results of this factor analysis is questionable because of highly skewed positive distribution and probable selective dropout of individuals (Berman et al., 2005).⁴

Using different optimal cutoff scores and different populations, as well as different methods to establish diagnostic criteria, most research groups that investigated the sensitivity and specificity of the DUDIT revealed favorable results. To obtain maximum sensitivity on the DUDIT, a lower cutoff score is recommended when screening for drug abuse, and a higher cutoff score is recommended when screening for nondrug abuse. The specificity of the DUDIT is increased when the cutoff score is high(er) and, as a result, sensitivity decreases. According to Tiet, Finney, and Moos (2008), a cutoff score for a screening instrument like the DUDIT should, ideally, have both sensitivity and specificity >.80 to be considered as clinically useful, although others define suitable cutoff scores as DUDIT scores yielding sensitivity above .80 and specificity above .70 (e.g., Nesvåg et al., 2010). As stated before, optimal cutoff scores may very well differ depending on the population studied (Berman et al., 2005; Gundersen et al., 2013). The results of this review revealed that optimal cutoff scores varied from 2 (Matuszka et al., 2014) to 25 (Berman et al., 2005). Low cutoff points may be indicative of a less extensive drug use among these participants. In the end it is the responsibility of the professional (clinicians, researchers) to select what cutoff score to use depending on the screening purpose (Carey, Carey, & Chandra, 2003; Cocco & Carey, 1998).

Some limitations of the current review should be mentioned. Although we systematically searched the data bases there is the possibility that relevant, especially unpublished data or non-English studies has not been identified, although we believe the potential impact of this selection criterion to be limited. In addition, it needs to be noted that all searching was conducted by the author. While the search terms for including studies in this review appear objective and straight forward, the possibility that others would obtain different results must be acknowledged. We also need to consider the possibility that publication bias could explain the relative paucity of identified research. However, because it is unclear in this context whether studies demonstrating non-reliability or non-validity would be any less likely (or more likely) to be published compared to studies demonstrating reliability or

validity, we believe publication bias is unlikely to have caused us to draw erroneous conclusions about the overall reliability or validity of the DUDIT. Another potential criticism may be that we did not compare the psychometric properties of the DUDIT in depth to other screening measures such as the DAST-10 or UNCOPE but the goal was not to determine equal validity or relative superiority of one measure over the other but rather a concise review of the data available for the DUDIT. Last but not least one could raise that we did not systematically evaluate the quality of the studies. For example, we only looked at aspects whether reliability, factor structure or sensitivity/specificity was reported. We felt that at this stage it was more important to deliver a more narrative review approach incorporating all the evidence available.

Despite the growing level of research activity on the DUDIT in the last decade, several issues clearly need further research or are as yet unresolved. In reviewing the present state of knowledge, a couple of additional issues were identified that merit further attention in future DUDIT studies. Test–retest reliability questions have been addressed only in a preliminary manner by the three studies we reviewed. Additional research that confirms the stability of the DUDIT is needed, especially to determine its long-term (i.e., >6 months) test–retest reliability. Future research should also examine test–retest reliability for scoring above versus below the cutoff score in classifying patients. Also, further factor analytic studies using large and heterogeneous samples, and attempts to extract varying numbers of factors are needed to better understand the factor structure of the DUDIT and to determine if the factor structure of the instrument varies as a function of the nature of the sample along dimensions such as severity of drug abuse, mental illness, age, gender, and culture. The studies conducted so far were primarily conducted in (Northern) Europe. Also, seven of the of the 18 included studies (39%) were (co)-authored by a researcher who had also been an author of the original DUDIT article (Berman et al., 2005). It must be borne in mind that there is potential to introduce bias whereby developers examining the scale’s psychometric properties may have an interest in it performing well. Therefore, more investigations, by other researchers, are desirable. Finally, although the DUDIT has been shown to be effective in identifying subjects in various settings with hazardous or harmful drug use, it would also be worthwhile to know more about how well the DUDIT helps to give both client and clinician a sense of the drug-related issues that could be focused upon in later treatment (Berman et al., 2003). Berman et al. (2005) already suggested that future clinical research also could explore the utility of the DUDIT as a re-test instrument following appropriate treatment.

6. Conclusion

As a screening test for current problematic drug use, the DUDIT seems to be a sensitive and specific instrument, performing at a level at least comparable to, if not higher than, other drug screening instruments. Some of the strengths of the DUDIT are that it is brief and easy to administer, it focuses on current (i.e., past year) behavior, and its content is relatively free of cultural bias. Obviously, many research questions remain, but it seems fair to conclude that the DUDIT has proven itself to be reliable, valid, and practical.

References⁵

- Addington, J., & Addington, D. (2007). Patterns, predictors and impact of substance use in early psychosis: A longitudinal study. *Acta Psychiatr Scand*, 115, 304–309. <http://dx.doi.org/10.1111/j.1600-0447.2006.00900.x>.
- *Berg, J. E., & Asskildt, T. J. (2011). Diagnosis after an acute psychiatric inpatient stay: How do psychotic and non-psychotic diagnoses relate to the results of psychometric tests of substance abuse? *J Psychiatr Intensive Care*, 7, 11–16. <http://dx.doi.org/10.1017/S1742646410000087>.

⁴ It should be noted that research has demonstrated that the general rule of thumb of the minimum sample size is in fact nothing more than a general rule of thumb (e.g., Guadagnoli, & Velicer, 1988; MacCallum, Widaman, Zhang, & Hong, 1999). The minimum sample size (also) depends on other aspects of design, such as size of the factor loadings and communalities (e.g., Field, 2013).

⁵ References marked with an asterisk indicate studies included in the review.

- Berman, A. H., Bergman, H., Palmstierna, T., & Schlyter, F. (2003). *DUDIT — The Drug Use Disorders Identification Test, MANUAL Version 1.0*. Stockholm: Karolinska Institutet, Department of Clinical Neuroscience, Section for Alcohol and Drug Dependence Research.
- *Berman, A. H., Bergman, H., Palmstierna, T., & Schlyter, F. (2005). Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. *Eur Addict Res*, 11, 22–31, <http://dx.doi.org/10.1159/000081413>.
- Berman, A. H., Palmstierna, T., Källmén, H., & Bergman, H. (2007). The self-report Drug Use Disorders Identification Test-Extended (DUDIT-E): Reliability, validity, and motivational index. *J Subst Abuse Treat*, 32, 357–369, <http://dx.doi.org/10.1016/j.jsat.2006.10.001>.
- Bland, J. M., & Altman, D. G. (1997). Statistic notes: Cronbach's alpha. *Br Med J*, 314, 572, <http://dx.doi.org/10.1136/bmj.314.7080.572>.
- Carey, K. B., Carey, M. P., & Chandra, P. S. (2003). Psychometric evaluation of the alcohol use disorders identification test and short drug abuse screening test with psychiatric patients in India. *J Clin Psychiatry*, 64, 767–774, <http://dx.doi.org/10.4088/JCP.v64n0705>.
- Cocco, K. M., & Carey, K. B. (1998). Psychometric properties of the Drug Abuse Screening Test in psychiatric outpatients. *Psychol Assess*, 10, 408–414, <http://dx.doi.org/10.1037/1040-3590.10.4.408>.
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cruce, G., Nordström, L. G., & Öjehagen, A. (2007). Risky use and misuse of alcohol, drugs and cigarettes detected by screening questionnaires in a clinical psychosis unit. *Nord J Psychiatry*, 61, 92–99, <http://dx.doi.org/10.1080/08039480701226062>.
- Cruce, G., & Öjehagen, A. (2007). Risky use of alcohol, drugs and cigarettes in a psychosis unit: A 1 1/2 year follow-up of stability and changes after initial screening. *BMC Psychiatry*, 7, 31, <http://dx.doi.org/10.1186/1471-244X-7-31>.
- *Durbecq, N., Berman, A. H., Gumpert, C. H., Palmstierna, T., Kristiansson, M., & Alm, C. (2010). Validation of the Alcohol Use Disorders Identification Test and the Drug Use Disorders Identification Test in a Swedish sample of suspected offenders with signs of mental health problems: Results from the mental disorder, substance abuse and crime study. *J Subst Abuse Treat*, 39, 364–377, <http://dx.doi.org/10.1016/j.jsat.2010.07.007>.
- *Evren, C., Ogel, K., Evren, B., & Bozkurt, M. (2014). Psychometric properties of the Turkish versions of the Drug Use Disorders Identification Test (DUDIT) and the Drug Abuse Screening Test (DAST-10) in the prison setting. *J Psychoactive Drugs*, 46, 140–146, <http://dx.doi.org/10.1080/02791072.2014.887162>.
- *Evren, C., Ovali, E., Karabulut, V., & Cetingok, S. (2014). Psychometric properties of the Drug Use Disorders Identification Test (DUDIT) in heroin dependent adults and adolescents with drug use disorder. *Bull Clin Psychopharmacol*, 24, 39–48, <http://dx.doi.org/10.5455/bcp.20130310124522>.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Thousand Oaks, CA: SAGE.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995). *Structured Clinical Interview for DSM-IV Axis I Disorders—Patient Edition (SCID-I/P, Version 2.0)*. New York: Biometrics Research Department, New York State Psychiatric Institute.
- Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychol Bull*, 103, 265–275, <http://dx.doi.org/10.1037/0033-2909.103.2.265>.
- *Gundersen, Ø. H., Mordal, J., Berman, A. H., & Bramness, J. G. (2013). Evaluation of the Alcohol Use Disorders Identification Test and the Drug Use Disorders Identification Test among patients at a Norwegian psychiatric emergency ward. *Eur Addict Res*, 19, 252–260, <http://dx.doi.org/10.1159/000343485>.
- *Hildebrand, M., & Noteborn, M. G. C. (2015). Exploration of the (interrater) reliability and latent factor structure of the Alcohol Use Disorders Identification Test (AUDIT) and Drug Use Disorders Identification Test (DUDIT) in a sample of Dutch probationers. *Substance Use & Misuse*, <http://dx.doi.org/10.3109/10826084.2014.998238> (Advance online publication).
- *Hillege, S., Das, J., & de Ruiter, C. (2010). The Youth Psychopathic Traits Inventory: Psychometric properties and its relation to substance use and interpersonal style in a Dutch sample of non-referred adolescents. *J Adolesc*, 33, 83–91, <http://dx.doi.org/10.1016/j.adolescence.2009.05.006>.
- Hoffmann, N. G., Hunt, D. E., Rhodes, W. M., & Riley, K. J. (2003). UNCOPE: A brief substance dependence screen for use with arrestees. *J Drug Issues*, 33, 29–44, <http://dx.doi.org/10.1177/002204260303300102>.
- *Kader, R., Seedat, S., Koch, J. R., & Parry, C. D. (2012). A preliminary investigation of the AUDIT and DUDIT in comparison to biomarkers for alcohol and drug use among HIV-infected clinic attendees in Cape Town, South Africa. *Afr J Psychiatry*, 15, 346–351, <http://dx.doi.org/10.4314/ajpsy.v15i5.43>.
- *Landheim, A., Bakken, K., & Vaglum, P. (2006). Impact of comorbid psychiatric disorders on the outcome of substance abusers: a six year prospective follow-up in two Norwegian counties. *BMC Psychiatry*, 6, 44, <http://dx.doi.org/10.1186/1471-244X-6-44>.
- Lange, E. H., Nesvåg, R., Ringen, P. A., Bhandari Hartberg, C., Haukvik, U. K., Andreassen, O. A., et al. (2014). One year follow-up of alcohol and illicit substance use in first-episode psychosis: Does gender matter? *Compr Psychiatry*, 55, 274–282, <http://dx.doi.org/10.1016/j.comppsy.2013.08.018>.
- *Lobmaier, P. P. K., Berman, A. H., Gossop, M., & Ravndal, E. (2013). Substance use and problem awareness among drug-involved prisoners in Norway. *J Subst Use*, 18, 211–220, <http://dx.doi.org/10.3109/14659891.2012.661022>.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychol Methods*, 4, 84–99, <http://dx.doi.org/10.1037/1082-989X.4.1.84>.
- Margolese, H. C., Malchy, L., Negrete, J. C., Tempier, R., & Gill, K. (2004). Drug and alcohol use among patients with schizophrenia and schizoaffective disorder: Levels and consequences. *Schizophr Res*, 67, 157–166, [http://dx.doi.org/10.1016/S0920-9964\(02\)00523-6](http://dx.doi.org/10.1016/S0920-9964(02)00523-6).
- *Martin, L., Viljoen, M., Kidd, M., & Seedat, S. (2014). Are childhood trauma exposures predictive of anxiety sensitivity in school attending youth? *J Affect Disord*, 168, 5–12, <http://dx.doi.org/10.1016/j.jad.2014>.
- *Matuszka, B., Bácskai, E., Berman, A. H., Czobor, P., Sinadinovic, K., & Gerevich, J. (2014). Psychometric characteristics of the Drug Use Disorders Identification Test (DUDIT) and the Drug Use Disorders Identification Test-Extended (DUDIT-E) among young drug users in Hungary. *Int J Behav Med*, 21, 547–555, <http://dx.doi.org/10.1007/s12529-013-9336-8>.
- Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *J Consult Clin Psychol*, 62, 783–792, <http://dx.doi.org/10.1037/0022-006X.62.4.783>.
- *Nesvåg, R., Lange, E. H., Faerden, A., Barret, E. A., Emilsson, B., Ringen, P. A., et al. (2010). The use of screening instruments for detecting alcohol and other drug use disorders in first-episode psychosis. *Psychiatry Res*, 177, 228–234, <http://dx.doi.org/10.1016/j.psychres.2010.01.007>.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- *Reddy, S., Dick, A. M., Gerber, M. R., & Mitchell, K. (2014). The effect of a yoga intervention on alcohol and drug abuse risk in veteran and civilian women with posttraumatic stress disorder. *J Altern Complement Med*, 10, 750–756, <http://dx.doi.org/10.1089/acm.2014.0014>.
- Regier, D. A., Farmer, M. E., Rae, D. S., Locke, B. Z., Keith, S. J., Judd, L. L., et al. (1990). Comorbidity of mental disorders with alcohol and other drug abuse. Results from the epidemiological catchment area (ECA) study. *JAMA*, 264, 2511–2518, <http://dx.doi.org/10.1001/jama.1990.03450190043026>.
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*, 59(Suppl. 20), 22–33.
- *Sinadinovic, K., Berman, A. H., Hasson, D., & Wennberg, P. (2010). Internet-based assessment and self-monitoring of problematic alcohol and drug use. *Addict Behav*, 35, 464–470, <http://dx.doi.org/10.1016/j.addbeh.2009.12.021>.
- *Sinadinovic, K., Berman, A. H., & Wennberg, P. (2011). Population screening of risky alcohol and drug use via Internet and Interactive Voice Response (IVR): A feasibility and psychometric study in a random sample. *Drug Alcohol Depend*, 114, 55–60, <http://dx.doi.org/10.1016/j.drugalcdep.2010.09.004>.
- Skinner, H. A. (1982). The drug abuse screening test. *Addict Behav*, 7, 363–371, [http://dx.doi.org/10.1016/0306-4603\(82\)90005-3](http://dx.doi.org/10.1016/0306-4603(82)90005-3).
- Stuart, G. L., Moore, T. M., Kahler, C. W., & Ramsey, S. E. (2003). Substance abuse and relationship violence among men court-referred to batterers' intervention programs. *Subst Abuse*, 24, 107–122, <http://dx.doi.org/10.1023/A:1023736732766>.
- Stuart, G. L., Moore, T. M., Kahler, C. W., Ramsey, S. E., & Strong, D. (2004). Cigarette smoking and substance use among men court-referred to domestic violence treatment programs. *Am J Addict*, 13, 319–320, <http://dx.doi.org/10.1080/10550490490460283>.
- Stuart, G. L., Moore, T. M., Ramsey, S. E., & Kahler, C. W. (2003). Relationship aggression and substance use among women court-referred to domestic violence intervention programs. *Addict Behav*, 28, 1603–1610, <http://dx.doi.org/10.1016/j.addbeh.2003.08.038>.
- Stuart, G., Moore, T. M., Ramsey, S. E., & Kahler, C. W. (2004). Hazardous drinking and relationship violence perpetration and victimization in women arrested for domestic violence. *J Stud Alcohol*, 65, 46–53.
- Stuart, G. L., Temple, J. R., Follansbee, K., Bucossi, M. M., Hellmuth, J. C., & Moore, T. M. (2008). The role of drug use in a conceptual model of intimate partner violence in men and women arrested for domestic violence. *Psychol Addict Behav*, 22, 12–24, <http://dx.doi.org/10.1037/0893-164X.22.1.12>.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.
- Tiet, Q. Q., Finney, J. W., & Moos, R. H. (2008). Screening psychiatric patients for illicit drug use disorders and problems. *Clin Psychol Rev*, 28, 578–591, <http://dx.doi.org/10.1016/j.cpr.2007.08.002>.
- *Voluse, A. C., Gioia, C. J., Carter Sobell, L., Dum, M., Sobell, M. B., & Simco, E. R. (2012). Psychometric properties of the Drug Use Disorders Identification Test (DUDIT) with substance abusers in outpatient and residential treatment. *Addict Behav*, 37, 36–41, <http://dx.doi.org/10.1016/j.addbeh.2011.07.030>.
- WHO ASSIST Working Group (2002). The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Development, reliability and feasibility. *Addiction*, 97, 1183–1194, <http://dx.doi.org/10.1046/j.1360-0443.2002.00185.x>.
- World Health Organization (1999). *Schedules for Clinical Assessment in Neuropsychiatry, Version 2.1: Interview, Glossary and Code Book*. Geneva: World Health Organization, Division of Mental Health, Assessment, Classification and Epidemiology.
- Wüsthoff, L. E., Waal, H., & Gråve, R. W. (2014). The effectiveness of integrated treatment in patients with substance use disorders co-occurring with anxiety and/or depression – a group randomized trial. *BMC Psychiatry*, 14, 67, <http://dx.doi.org/10.1186/1471-244X-14-67>.