3.3 Measurement of nicotine dependence

Introduction

In this section, evidence of the validity of self-report measures of nicotine/tobacco dependence in adults is examined. Measures are concentrated on that are potentially appropriate for population-based/ epidemiologic research, as nicotine dependence is often assessed as a potential moderator of programme and policy effects. The Working Group (WG) has focused mainly on scales measuring cigarette dependence, as cigarette smoking accounts for most of the health damage caused by tobacco, and because the most widely used and best studied scales measure cigarette dependence. This section has not attempted to review evidence evaluating measures to assess nicotine dependence of other types of smoked tobacco products (e.g. cigars, pipe tobacco, bidis, hookah), although adaptations of measures used to assess cigarette smoking dependence would be reasonable to consider. The WG did include a review of measures of dependence on smokeless tobacco products, since the pattern of compulsive use of these products is similar to that observed for cigarette smoking (IARC, 2007b). Persistent use of nicotine medications has been described, but it is very rare (Shiffman *et al.*, 2003). Also, long-term use of nicotine medications has no documented untoward health effects, so therefore measurement of dependence to nicotine medications will not be included in this review. Finally, while dependence on tobacco products is clearly evident among some youth, research on measures of nicotine dependence in adolescents is limited, and will not be considered in this section. For those interested in a measure of nicotine dependence among youth, please refer to the paper which describes the measurement properties of the Hooked on Nicotine Checklist (DiFranza *et al.*, 2002b).

Nicotine dependence is a hypothetical construct that is designed to explain and predict societally-important outcomes, such as an inability to quit smoking, heavy use, and other problems occasioned by smoking or tobacco use (Piper *et al.*, 2006). Assessing tobacco dependence is difficult and is made even more so in population-based epidemiologic research by the need for efficient assessment (valid and brief). Ideally, a measure should reflect the nature or domain of the construct of interest (i.e. tobacco dependence), predict important outcomes (e.g. likelihood of quitting, problems encountered through use), and be relatively brief to assess.

Measures of cigarette-induced nicotine dependence

The following section provides a brief review of data on the measurement properties of seven self-report measures developed to assess the construct of cigarette-induced nicotine dependence: 1) Fagerström Test for Nicotine Dependence (FTND); 2) Heaviness of Smoking Index (HSI); 3) Diagnostic and Statistical Manual-IV (DSM-IV) criterion of dependence; 4) International Statistical Classification and Related Health Problems-10 (ICD-10) criteria; 5) Cigarette Dependence Scale (CDS); 6) Nicotine Dependence Syndrome Scale (NDSS); and 7) Wisconsin Inventory of Smoking Dependence Motives (WISDM).

Each measure will be evaluated based on a review of the items that constitute the scales in terms of their reading level, face validity, coverage of the dependence domain, and cross-cultural applicability. The WG will review the psychometrics of each scale, including its reliability (e.g. internal consistency) and factor structure, and will examine the predictive validity of each measure, focusing on two specific tobacco dependence criteria: a pattern of pervasive and heavy smoking and the ability to quit smoking.

Pervasive and heavy smoking could be assessed using self-report measures (e.g. cigarettes smoked per day or lifetime cigarettes smoked), or using biomarkers of exposure (e.g. carbon monoxide (CO), cotinine, puff topography) (see Section 3.1), and the ability to quit smoking could be assessed using a number of strategies as well (see Section 3.1). These criteria reflect the sheer volume of tobacco products consumed and the intransigence of drug use, both of which have significant effects on the health and economics of both the individual and society. Although it is not a validation criterion, the evidence of genetic linkages to the various measures of tobacco dependence will be examined. This information may be helpful for researchers who are interested in using epidemiological measures to make inferences regarding etiology.

It is important to note that other criteria could be used to evaluate the performance of dependence measures. For instance, such measures could be evaluated with respect to prediction of withdrawal severity or other outcomes theoretically linked to dependence (Piper *et al.*, 2006). However, such outcomes seem less relevant than the ones selected for measures to be used in epidemiologic research. For the purposes of epidemiologic research, a measure should reflect or predict outcomes of societal import, such as degree of tobacco exposure and use, the intransigence of use, and the likelihood of important negative outcomes of use. Obviously, a pattern of heavy, pervasive smoking will capture the degree of exposure to nicotine and the harmful constituents of tobacco/cigarettes. Moreover, a relative inability to quit smoking will forecast the likely continued exposure to such elements. Evidence shows that past, current, and future use of tobacco directly predict outcomes of societal import, such as money expended in buying tobacco products and disease outcomes (and associated costs) caused by tobacco use (US Department of Health and Human Services, 2004; Centers for Disease Control and Prevention, 2005).

Overarching issues:

It is important to note that dependence is a construct (i.e. a hypothetical entity). It is not, in theory, equivalent to any single measure or criterion (Piper *et al.,* 2006); although single items can be used to estimate a person's standing on the construct. Thus, dependence is an inferred influence or force that produces the outcomes associated with it (e.g. high rates of smoking, relapse), although it is not the only predictor of such outcomes. Generally it takes multiple variables or items to adequately assess a complex, hypothetical entity such as nicotine dependence (Clark & Watson, 1995). In this section, however, considerable attention is devoted to very brief measures of dependence, as evidence shows that such measures (i.e. number of cigarettes smoked per day) can predict outcomes, such as relapse, as well as longer measures (e.g. DSM-III-R, FTQ, and FTND) (Razavi *et al.,* 1999; Breslau & Johnson, 2000; Dale, *et al.*, 2001).

When considering the information comprised here, it is important to remember that reliability and validity are not inherent in measures. It can not be assumed that one can generalize psychometric properties across different use contexts, or that validity for one use of a measure is generalizeable for a different use (e.g. predicting relapse likelihood versus withdrawal severity). Rather, these features are estimated based on patterns of statistical covariation and are influenced by the nature of the population being assessed (Nunnally & Bernstein, 1994; McDonald, 1999). For instance, there may be less variance in item scores, or item scores might have a less skewed distribution, when a dependence measure is used in a clinical population rather than a nationally representative population. This could easily affect both reliability and validity estimates. Different populations might yield different psychometric data because of true differences in the severity or range of dependence. However, differences might also arise because of other factors, such as secular or environmental events that might affect scores on dependence measures, while not actually changing the dependence per se. One study showed that US smokers had higher frequencies of severe nicotine dependence

(FTND ≥ 6) than did Spanish smokers (de Leon *et al*., 2002). It is possible that such population differences reflect different degrees or sources of error across the two populations (restrictions in smoking in the home, the amount of discretionary income, gender differences in smoking across the populations, the ways the smokers answer the questions and, indeed, understand them and so on) rather than differences in the biological/psychological internal processes that make up dependence. There are numerous environmental or social sources of error variance that could differentially affect the validity of a measure across populations: smoking policies in the workplace, taxes, religious or social norms, to list few.

In recognition of the dependence of psychometric properties on the population being assessed, reliability and validity data from both clinical trials and epidemiologic studies conducted around the world, and present data relating to the heritability of dependence as it is assessed using the different measures, will be presented. The tobacco dependence measures will be divided into two groups: *unidimensional* and *multidimensional*. Unidimensional measures are intended to assess dependence as a single dimension (although some, it turns out, may actually be multifactorial). Such measures are useful, because the best of them are fairly efficient in that they possess significant validity given their length/response burden. In fact, as efficient as some of the unidimensional measures are, some data suggest that particular items from these measures possess predictive validities that meet or exceed those of the whole measure (Storr *et al.,* 2005). Such items might be especially valuable for epidemiologic research.

A review of multidimensional measures of nicotine dependence are included despite their length and reduced efficiency, because they have the potential to provide information about the mechanism underlying nicotine dependence not supplied by unidimensional measures. For instance, multidimensional measures are intended to assess particular facets of dependence or dependence processes (e.g. particular motives for drug use). Thus, these measures may provide greater insights into the nature of tobacco dependence than do unidimensional measures. They also may provide greater discrimination amongst smokers/tobacco users to the extent that smokers may be distinguished on the basis of something other than a single intensity dimension (which might be well captured by a single severity dimension). For instance, some scales appear to reflect motives associated with initial versus extensive use of tobacco (Piper *et al*., 2004), and other scales differ in sensitivity to use patterns of highly dependent users versus “chippers” (those who engage in periodic or light tobacco use) (Shiffman & Sayette, 2005). Since the subscales of multidimensional measures tend to ask about relatively discrete processes (e.g. a taste motive for smoking) rather than global consequences of smoking (e.g. smoking causing problems in life), these multidimensional measures may be more suitable for genetics research, as they may tap processes that reflect a stronger genetic signal (Baker *et al*., in press). Finally, because multidimensional measures tend to ask about internal and subjective phenomena (e.g. role of affect regulation) rather than externally referenced events (e.g. latency to smoke in the morning, number of cigarettes consumed each day), these measures may be less susceptible to biasing by error due to regional secular or policy influences. Workplace smoking restrictions, for example, might exert a more direct and larger effect on number of cigarettes smoked per day than on the smokers liking of the taste of cigarettes. On the other hand, multidimensional scales tend to ask about relatively subtle, psychological variables (e.g. asking individuals to attribute smoking urges or affect), and it is possible, indeed probable, that cultures may differ in how they make attributions or label internal phenomena. Of course, while entire multidimensional scales can be quite lengthy, individual items or subscales can be selected for use (Lerman *et al*., 2006); thus, this section will review relevant subscale data.

The foregoing discussion should make clear that blanket

recommendations cannot be given regarding dependence. Rather, the investigator must both weigh practical issues (e.g. response burden) and clearly identify the goals of assessment (e.g. predict probability of relapse) in order to select an appropriate dependence instrument or assessment strategy.

Unidimensional measures of tobacco dependence

Fagerström Test for Nicotine Dependence and the Heaviness of Smoking Index

The first unidimensional measure of tobacco dependence is actually a group of measures arising from the Fagerström Tolerance Questionnaire (FTQ) (Fagerström, 1978): these comprise the FTQ itself, as well as the 6-item Fagerström Test for Nicotine Dependence (FTND) (Heatherton *et al*., 1991) and the 2-item Heaviness of Smoking Index (HSI) (Kozlowski *et al*., 1994). See Appendix 1 for the items and scoring. These measures are based on the construct of physical dependence, which includes facets such as the need to smoke early in the morning to alleviate overnight withdrawal, the need to smoke numerous cigarettes per day, and the invariance of smoking behaviour (i.e. smoking even when you are ill) (Fagerström, 1978). The Flesch-Kincaid Reading Grade Level is 4.4 for the FTND and 4.2 for the HSI.

The FTND has been translated and used with population samples in Germany (John *et al*., 2003a; John *et al*., 2004a), Switzerland (Etter *et al.,* 1999), Australia (Pergadia *et al*., 2006a), Canada (Howard *et al*., 2003), Austria (Lesch *et al.,* 2004), and Brazil, Mexico, Poland, and China (Blackford *et al*., 2006; Huang *et al*., 2006). The HSI has also been used in research in Spain (Diaz *et al*., 2005), Australia, Canada, UK, and the USA (Heatherton *et al*., 1991; Hymowitz *et al*., 1997; Hyland *et al*., 2006). One of the questions on the FTND concerns smoking in forbidden places. The validity of this question may be affected by regional differences in environmental restrictions in smoking (Huang *et al*., 2006). In addition, two questions in this scale assume a pattern of daily smoking (e.g. questions 1 & 4, the two questions in the HSI). It is very likely that scores on these items will have reduced validity if used with non-daily smokers. An important goal of future research is to identify dependence measures that are appropriate for non-daily smokers.

Reliability and structure: Compared with the FTQ, the FTND has demonstrated better psychometric properties, such as internal consistency (Payne *et al.,* 1994; Pomerleau *et al*., 1994; Haddock *et al*., 1999); however, these improved reliability coefficients are still low (Etter, 2005) and below traditionally accepted standards for clinical use (α =0.80) (Nunnally & Bernstein, 1994). Using a French translation of the FTND with light smokers found internal consistencies of approximately 0.70 (Etter *et al*, 1999), while a study with a German population found low internal consistency for the FTND (α =.57) in two separate samples (John *et al*., 2004b), and a study in China found that FTQ had low internal consistency as well (α =.58) (Huang *et al*., 2006).

Some studies have shown that the FTND has a two-factor structure, suggesting that it does not measure a unitary construct of physical dependence (Payne *et al*., 1994; Etter *et al*., 1999; Haddock *et al*., 1999; Radzius *et al.,* 2003; Breteler *et al*., 2004; John *et al*., 2004b). A population-based study in France found that while a two-factor model fit the data well, the two factors were highly correlated (Chabrol *et al*., 2003). Inter-item correlations also reveal that not all items are highly related (r = 0.06-0.39) (Transdisciplinary Tobacco Use Research Center (TTURC) Tobacco Dependence Phenotype Workgroup, 2007). These studies suggest that the two factors reflect *morning smoking* (i.e. whether one smokes more in the morning and whether one would rather give up the first cigarette of the day or all others), and *smoking pattern* (i.e. the number of cigarettes smoked per day, time to first cigarette, difficulty refraining from smoking, and smoking when ill), although some data indicate that time to first cigarette loaded on both factors (Radzius *et al*., 2003). Latent class analyses suggest that the FTND divides smokers into groups based on severity of dependence (Storr *et al*., 2005); that is the two factors do not

appear to "pick-out" smokers who differ in terms of types of dependence.

The HSI is comprised of only two items, which limits the relevance of internal consistency estimates. However, zero-order correlations between the two items in the measure indicate moderate levels of association (e.g. r's ≈ 0.30) (TTURC Tobacco Dependence Phenotype Workgroup, 2007).

Validation: The FTND and HSI predict both behavioural and biochemical indices of smoking in Chinese-, English-, French-, and German-speaking populations (e.g. CO, cotinine, lifetime amount smoked) (Heatherton *et al*., 1989, 1991; Kozlowski *et al*., 1994; Etter *et al*., 1999; John *et al*., 2003a; Huang *et al*., 2006). This should not be surprising, given that the FTND and HSI directly assess smoking heaviness. However, it is encouraging to note that smokers are indeed able to estimate their amount of smoking as indexed by biochemical tests in response to single items (e.g. Question #4 on the FTND, "How many cigarettes/day do you smoke?"). The FTND has demonstrated an ability to predict cessation outcomes in smoking cessation studies (Campbell *et al.,* 1996; Westman *et al.,* 1997; Alterman *et al.,* 1999; Patten *et al*., 2001; TTURC Tobacco Dependence Phenotype Workgroup, 2007), and with college students in a population-based study (Sledjeski *et al*., 2007). In addition, the FTND has been shown to index a heightened risk for psychiatric comorbidities in a large population sample in Germany (John *et al*., 2005).

Some data indicate that the standard scoring method used with the FTND (adding up item responses) may not produce an optimal scaling of dependence level. Latent class analysis suggested that some items are particularly important to the assessment of dependence level (those that capture variance due to morning smoking) and that they are relatively underweighted in the typical scoring method (Storr *et al*., 2005). Therefore, investigators using the FTND may wish to explore alternative, empirically-based scoring or cut-score determination methods (e.g. latent class analysis, Receiver Operating Characteristic curves (Swets *et al*., 2000)).

While the FTND certainly can predict future smoking or likelihood of cessation, the HSI appears to account for much of the predictive validity of that measure (Breslau & Johnson, 2000; Heatherton *et al*., 1989; TTURC Tobacco Dependence Phenotype Workgroup, 2007). Population-based studies conducted in Australia, Canada, the UK, and the USA found that the two HSI items (number of cigarettes smoked and time to first cigarette in the morning) were the strongest predictors of quitting (Hymowitz *et al*., 1997; Hyland *et al*., 2006). Furthermore, recent research has shown that a single item on the FTND and HSI (Item #1 – latency to first cigarette in the morning) predicts relapse vulnerability, as well as, or better than, much longer multidimensional instruments (TTURC Tobacco Dependence Phenotype Workgroup, 2007). Recent population-based research shows that a single item on the HSI (item #1) is highly effective in predicting the likelihood of future cessation (TTURC Tobacco Dependence Phenotype Workgroup, 2007).

Heritability: In a study of young adult Australian Twins, HSI-assessed dependence was found to be highly heritable (71%) (Lessov *et al*., 2004). In addition, the FTND and HSI were both related to the dopa decarboxylase gene, which is involved in the synthesis of dopamine, norepinephrine, and serotonin (Ma *et al*., 2005). One haplotype was significantly related to dependence in both African-American and Euro-American smokers, while another was related to dependence only in Euro-American smokers (Ma *et al*., 2005). Additional studies link FTND-defined dependence to particular genetic variants (Bierut *et al*., 2007; Gelernter *et al*., 2007; Saccone *et al*., 2007).

Summary: The FTND has been widely used in a number of different countries and a number of different languages. It is short and has an accessible reading level. In addition, while there are concerns regarding its structure and reliability, it has been found to predict smoking heaviness and cessation outcome. However, it

appears that the HSI is a more efficient predictor of outcome than is the FTND (using only two items). FTND and HSI scores have also been found to be heritable and related to specific dependence-linked genetic variants.

The Diagnostic and Statistical Manual, International Statistical Classification of Diseases and Related Health Problems, 10th Revision and the Tobacco Dependence Screener

Two different diagnostic systems are commonly used to diagnose tobacco dependence: both are typically considered to be unidimensional measures of tobacco dependence. One is the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) (American Psychiatric Association, 1995)[1] which is based on an empirically driven, syndromal medical model, rather than on a theoretical model of dependence (see Appendix 2 for the criteria). The second is the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10), an international diagnostic classification system that was endorsed by the 43rd World Health Assembly in May 1990 and came into use by WHO Member States as of 1994 (see Appendix 3 for the criteria (WHO, 1993)). The Tobacco Dependence Screener (TDS) (Kawakami *et al.*, 1999) is a 10-item, self-report questionnaire designed to assess ICD-10, DSM-III-R (the 1987 revision of DSM-II), and DSM-IV symptoms of dependence with a Flesch-Kincaid Reading Grade Level of 8.1 (see Appendix 4 for items and scoring). To the best of our knowledge, this is the only published, self-report DSM/ICD questionnaire of tobacco/nicotine dependence. Most of the existing research has utilised the DSM criteria, and that will be the focus of this Handbook's review of diagnostic classifications of tobacco dependence.

DSM and ICD structured clinical interviews, such as the World Mental Health Survey Initiative version of the Composite International Diagnostic Interview (CIDI), or the National Institute of Mental Health Diagnostic Interview Schedule (DIS), have been translated into various languages and used in at least 11 population-based studies (Hughes *et al.*, 2006) in countries including: Germany (John *et al.,* 2003b (DSM); John *et al*., 2004a (DSM); Hoch *et al.*, 2004 (DSM)), Australia (Pergadia *et al.*, 2006b (DSM)), Canada and Taipei (Howard *et al.*, 2003 (DSM)), Spain (de Leon *et al.*, 2002 (DSM)), Austria (Lesch *et al.,* 2004 (DSM & ICD)), Switzerland (Angst *et al.*, 2005 (DSM)), Japan (Yoshimura, 2000 (ICD)), Korea (Lee *et al.*, 1990 (DSM)), and the USA (Breslau *et al.*, 2004 (DSM); Hughes *et al.*, 2004a (DSM & ICD)). The ICD-10 criteria are available in 42 languages, including Arabic, Chinese, English, French, Russian, and Spanish. The DIS, CIDI, and other diagnostic interviews comprise a series of branching questions that are aimed at eliciting information about features relevant to nicotine dependence.

Some aspects of the DSM-derived interviews and similar instruments may cause problems in any sample, or when using the instrument with culturally diverse populations. Another important caveat to observe, in regards to the DSM measure of dependence, is that the scoring algorithm used in establishing formal DSM diagnoses does not appear to yield decision rules that agree with empirical methods, such as latent class analysis (Muthen & Asparouhov, 2006). Thus, the investigator may wish to explore different methods for item-weighting and cut-score estimation if a categorical outcome is desired. In addition, it should be noted that the tobacco sections of DIS and CIDI are quite long (over 30 items), and were designed to be administered either in a face-to-face interview or by a trained professional. New technology has made it possible to have individuals respond to text-based presentations of the questions, but it is unknown how valid this presentation method would be and it would remain quite time consuming.

[1]There has been a text revision of the DSM-IV (American Psychiatric Association, 2000), however this revision did not alter any diagnostic criteria for any diagnostic categories, including the substance dependence diagnosis

Reliability and structure: Data on the reliability and structure of diagnostic interview measures of nicotine dependence arise from studies using face-to-face administration strategies. Therefore, the following conclusions cannot be generalized to a different administration format. There is evidence that the various structured diagnostic measures yield reliable diagnoses as assessed by test-retest reliability (κ = 0.63, Grant *et al.*, 2004; κ = 0.88, Hughes *et al.*, 2004a; κ = 0.73, Koenen *et al.*, 2005). One-factor analysis indicated that responses to the CIDI had a strong single factor structure (Strong *et al.*, 2003); although other factor analyses of the structured diagnostic items found that a two-factor structure was a better fit (Johnson *et al.*, 1996; Radzius *et al.*, 2004; Muthen & Asparouhov, 2006). Patterns of covariation that were found amongst the symptoms could be best accounted for by two factors (Muthen *et al.*, 2006). The first accounted for covariance in the "tolerance," "larger amounts," and "time spent using" items (see Appendix 2). Thus, this factor seems to be highly related to sheer amount smoked. The second factor was related to "persistent desired/unsuccessful efforts to cut down or quit," and "continued use despite emotional/-physical problems." Confidence in this solution is bolstered by the fact that it was obtained in three separate groups of individuals. It is also consistent with other recent factor analyses (Lessov *et al.*, 2004). Investigators might wish to analyze these item parcels separately since they may be addressing somewhat distinct constructs.

The TDS, a written questionnaire assessing the presence of diagnostic criteria, has demonstrated acceptable internal consistency in Japanese smokers (α = 0.74-0.81) (Kawakami *et al.*, 1999), but was less internally consistent among smokers in the USA (α = 0.64) (Piper *et al.*, 2008). To date, there have been no studies comparing the reliability of the interview measures with the paper-pencil measure. Therefore, one cannot assume that the psychometric data generated by the interview-format delivery of DSM or ICD items would generalize to a self-administered format.

Validation: Evidence suggests that the small set of dichotomous DSM items can distinguish between light versus heavy smoking (Strong *et al.*, 2003). An epidemiological study found that the DSM-III-R (as assessed by the DIS), was a significant, though weak, predictor of cigarette abstinence over one year, but that the FTND was a better predictor and that number of cigarettes smoked per day was the best predictor (Breslau & Johnson, 2000). Another study showed that DSM-IV diagnoses of nicotine dependence predicted heaviness of use and cessation outcome in a population-based study of college students (Sledjeski *et al.*, 2007). Several studies have shown that DSM-IV nicotine dependence diagnosis is associated with greater risk of psychiatric comorbidities in adults and youth (Grant *et al.*, 2004; John *et al.,* 2004a; Dierker *et al.*, 2006). In addition, DSM diagnoses of nicotine dependence were significantly associated with self-rated general health in a population sample in Germany (John *et al.*, 2005). In sum, there is substantial evidence that DSM/ICD diagnoses are meaningfully related to smoking heaviness and a variety of health outcomes.

Studies have shown that the TDS is associated with the smoking heaviness measures (e.g. number of cigarettes smoked per day, CO levels) and years of smoking (Kawakami *et al.*, 1999; Piper *et al.*, 2004). With respect to relapse, one study found that Japanese smokers with lower TDS scores were more likely to quit smoking after a health risk appraisal (Kawakami *et al.*, 1999). However, data from smokers who participated in smoking cessation studies in the USA, revealed that the TDS did not predict abstinence at 1-week or 6-months post-quit (TTURC Tobacco Dependence Phenotype Workgroup, 2007).

Heritability: There has been considerable research supporting the heritability of DSM/ICD-diagnosed nicotine dependence. In the Australian Twin sample study, analyses revealed that all of the DSM-IV symptoms and diagnosed DSM-IV dependence were meaningfully heritable (45-73%), and that the DSM-IV criteria of

tolerance, withdrawal, and difficulty quitting were the most highly heritable symptoms of nicotine dependence for both men and women (Lessov *et al*., 2004). A study of twin fathers, using the Vietnam Era Twin Registry, found that paternal DSM diagnosis of nicotine dependence was significantly associated with offspring DSM diagnosis of nicotine dependence (Volk *et al.,* 2007). However, one study found that DSM nicotine dependence was not related to familial liability to smoking persistence, because familial density of persistence was not associated with smoking persistence among nicotine-dependent daily smokers (Johnson *et al*., 2002). Other genetics research has linked DSM-diagnosed nicotine dependence with the CYP2E1 genotype, which codes for a protein that metabolizes alcohol and tobacco smoke nitrosamines, and is implicated in creating metabolic cross-tolerance between alcohol and tobacco (Howard *et al.,* 2003).

Summary: There is evidence that diagnostic measures effectively index smoking heaviness, smoking-related health and mental health risks, and likelihood of future cessation. There is also strong evidence of heritability of DSM-diagnosed nicotine dependence. It is unclear whether paper-pencil versions of such measures (the TDS) are comparable to the interview versions of such measures. Moreover, there is evidence that the FTND may predict cessation and health outcomes as well as the diagnostic measures (e.g. John *et al*., 2004a). In terms of the prediction of likelihood of future cessation, it is unclear that diagnostic measures possess any incremental validity relative to briefer measures, such as the HSI. The diagnostic scales have relatively high reading levels, which may hinder their use with certain populations (even if administered orally).

Cigarette Dependence Scale

The Cigarette Dependence Scale (CDS) is another unidimensional tobacco dependence measure (Etter *et al*., 2003b). This assay was developed using smokers' reports of signs that they believed indicated addiction to cigarettes. Both a 5- and 12-item version of the CDS were developed (see Appendix 5). The items overlap somewhat with the Fagerström tests (e.g. they both assess number of cigarettes smoked per day and time to first cigarette in the morning). The Flesch-Kincaid Reading Grade Levels were 4.9 for the CDS-12 and 6.8 for the CDS-5.

Reliability and structure: To date, only two published studies have reported data on the two versions of the CDS, using data collected via the mail or Internet (Etter *et al.,* 2003b; Etter, 2005). The CDS-12 had strong internal consistency, the CDS-5 was within the acceptable range, and both scales were slightly skewed toward higher values. Test-retest correlations were ≥ 0.60 for all items and ≥ 0.83 for the full scales. Factor analysis suggested a unidimensional structure for the CDS-12.

Validation: The CDS scales were significantly correlated with number of cigarettes smoked per day (whether a smoker was a daily or occasional smoker), strength of urges during the last quit attempt, and cotinine level (Etter *et al*., 2003b). Curiously, the CDS-5 was more strongly correlated with cotinine levels than was the CDS-12. This was probably due to the fact that the question about smoking heaviness (Question #2) determined a greater portion of total scale variance in the 5-item version. In one study, none of the three dependence measures (i.e. the FTND, CDS-5, or CDS-12) was a significant predictor of relapse likelihood (Etter *et al*., 2003b); however, only a third of potential respondents participated in the follow-up study, which might have produced considerable response bias. In a second study, the CDS-12 weakly predicted smoking abstinence at 1-month post-quit, but in a counterintuitive direction (e.g. higher CDS-12 scores predicted abstinence) (Etter, 2005).

Heritability: To date, no data regarding heritability or genetics have been published using the CDS.

Summary: While the CDS scales do index smoking heaviness well, there is little evidence that they

predict likelihood of cessation effectively, or that they index other health outcomes of public health importance. Further, there is little evidence that they possess incremental validity relative to other measures, such as the diagnostic measures or the FTND. Overall, this measure is promising in that it can be used with paper-pencil administration and it has good reliability, but a meaningful evaluation must await additional validity research.

Multidimensional Measures of Tobacco Dependence

Nicotine Dependence Syndrome Scale

The Nicotine Dependence Syndrome Scale (NDSS) (Shiffman *et al.*, 2004) is a 19-item multidimensional scale based on Edwards and Gross' 1976 theory of the alcohol dependence syndrome. The NDSS was intended to complement, not replace, traditional dependence measures, such as the DSM-based assessments, and therefore there is little content overlap between the NDSS and the unidimensional measures. The NDSS assesses five dimensions of nicotine dependence: "Drive" reflects craving, withdrawal, and smoking compulsions; "Priority" reflects preference for smoking over other reinforcers; "Tolerance" reflects reduced sensitivity to the effects of smoking; "Continuity" reflects the regularity of smoking rate; and "Stereotypy" reflects the invariance of smoking (Appendix 6). The Flesch-Kincaid Reading Grade Level is 7.7. This reading level is somewhat elevated relative to other self-administered scales, which may reflect the fact that some items contain unusual words and require integration of more than one sentence or statement. For instance, the item, "My smoking pattern is very irregular throughout the day. It is not unusual for me to smoke many cigarettes in an hour, then not have another one until hours later," involves three negatives over its two sentences. In addition, some questions are double-barrelled, such as "It's hard to estimate how many cigarettes I smoke per day because the number often changes." If a person answers no, it is unclear whether the answer refers to difficulty of estimation per se, or because the number of cigarettes smoked per day does not change. Some items may be significantly influenced by cultural factors, such as eating in restaurants that are smoke-free or experiences during air travel. These features may make the NDSS somewhat less appropriate than some other measures for individuals of modest reading abilities or educational status. The NDSS has been translated into Finnish (Broms *et al.,* 2007).

Reliability and structure: To date, four studies of adult smokers have generated data on the NDSS; one study has reported on the NDSS in adolescents aged 12-18 (Clark *et al.*, 2005).

Psychometric data discussed here are based on the revised 30-item scale. The internal consistency for the NDSS total scale, the NDSS-T, is good (Shiffman *et al.*, 2004); however, data show that the internal consistencies of individual subscales are problematic (Piper *et al.*, 2006). Principal components analysis revealed a 5-factor structure for the NDSS (Shiffman *et al.*, 2004) as predicted by the underlying theory. Significant differences in the scores on the subscales between White and African-American smokers suggest the scale may operate differently in subpopulations, although there were no ethnic differences in the total NDSS score (Shiffman *et al.*, 2004). A more recent study, using the 19-item questionnaire with the Finnish Twin Cohort Study population, found that a 3-factor structure (priority/drive, continuity/stereotypy, and tolerance) best fit the data, with the internal consistencies of the three factors ranging from 0.83 to 0.92 (Broms *et al.*, 2007).

Validation: Much of the initial validation work was done with the 30- and 23-item NDSS, prior to its being refined to the 19-item version. These results indicated that the NDSS-T predicted time to lapse and time to relapse, but no individual subscale predicted lapse or relapse (Shiffman *et al.*, 2004). However, new data suggest that the NDSS subscales are significantly, though modestly, related to cigarettes smoked per day ($r = 0.12$-0.26) and that the Tolerance and Continuity subscales are modestly related to CO

level (r = 0.12 and 0.13, respectively) (Piper, *et al*., 2008). In samples of treatment-seeking smokers, the NDSS Priority and the Stereotypy subscales were found to predict cessation outcomes for up to 6-months post-quit (TTURC Tobacco Dependence Phenotype Workgroup, 2007; Piper, *et al*., 2008). The NDSS Drive, Tolerance, and the total score were found to predict heaviness of smoking and cessation outcome in a population-based sample of college students (Sledjeski *et al*., 2007). In Finnish smokers, the NDSS was significantly correlated with both FTND and DSM-IV, as assessed by the CIDI measures of dependence (Broms *et al*., 2007). The NDSS subscales accounted for 51% of the variance in self-reported difficulty abstaining among "chippers" (light/non-daily smokers) (Shiffman & Sayette, 2005), with the Drive subscale having the strongest relation (β = 0.61), relative to the other scales (β = 0.13-0.28).

Heritability: In the Finnish cohort, NDSS was found to have a significant heritability estimate of 0.30, relative to a heritability estimate of 0.40 for the FTND (Broms *et al*., 2007).

Summary: Like the CDS, the NDSS is a relatively new scale and it is not yet possible to draw firm conclusions about its validity relative to other dependence instruments. In its favour is the fact that some of its subscales have been shown to predict smoking heaviness measures, other dependence measures, and smoking cessation likelihood (Broms *et al.,* 2007; Piper *et al.,* 2008). The majority of this research has been done on clinical populations and it is not known how well these results would generalize to population-based samples. There is evidence that the various subscales of the measure are differentially related to various dependence criteria (Shiffman & Sayette, 2005; Broms *et al*., 2007; TTURC Tobacco Dependence Phenotype Workgroup, 2007). This suggests that some of the subscales possess discriminative validity with respect to different dimensions or aspects of dependence. However, there is evidence that the NDSS is not able to predict the major dependence criteria of smoking heaviness or cessation likelihood better than shorter measures (TTURC Tobacco Dependence Phenotype Workgroup, 2007). In addition, the marginal reliabilities of some of the subscales, and the reading level and complexity of some of the items, may discourage use in large population-based samples.

Wisconsin Inventory of Smoking Dependence Motives

The Wisconsin Inventory of Smoking Dependence Motives (WISDM) (Piper *et al*., 2004) is a 68-item measure developed to assess the discrete motivational basis of dependence. This measure has 13 theoretically-based subscales designed to tap different smoking dependence motives: Affiliative Attachment, Automaticity, Behavioral Choice/ Melioration, Cognitive Enhancement, Craving, Cue Exposure/Associative Processes, Loss of Control, Negative Reinforcement, Positive Reinforcement, Social and Environmental Goads, Taste and Sensory Properties, Tolerance, and Weight Control (see Appendix 7 for the items and scoring). The Flesch-Kincaid Reading Grade Level is 4.6; however, balanced against this easy reading level is the fact that the total scale is quite long. Therefore, investigators might wish to use individual, theoretically targeted subscales in epidemiologic research (subscales range from 4-7 items) (Lerman *et al*., 2006). Finally, relatively subtle psychological concepts are addressed in this measure, such as thinking of cigarettes as a friend or experiencing a loss of control, and this may affect the validity of such items in some cultures. There are English and Spanish versions of the WISDM (D.W. Wetter, personal communication, December 12, 2006).

While all subscales assess dependence, it should be noted that some of the subscales (i.e. Cue Exposure/Associative Processes, Social/Environmental Goads, and Taste/Sensory Properties) represent early-onset motives, which are present for all smokers even at modest levels of smoking experience, while other subscales represent late-onset motives (i.e. Affiliative Attachment, Automaticity, Behavioral Choice/ Melio-

ration, Cognitive Enhancement, Craving, and Tolerance), which are present only in individuals who smoke at a moderate daily rate or have at least moderate smoking experience (Piper *et al*., 2004).

Reliability and structure: To date, only one study has published data on the WISDM (Piper *et al*., 2004). Across two different samples all 13 subscales had strong internal consistencies that were evident across gender and across Whites and African-Americans. A new study found that the internal consistency of the subscales ranged from 0.74-0.94 with the total scale having a Chronbach's alpha of 0.96 (Piper *et al*., 2008). Factor analytic strategies indicated that the WISDM-68 is multidimensional, although some scales hit on related or overlapping dimensions of dependence. Thus, it is safe to say that some of the subscales are tapping the same underlying dimensions.

Validation: The total WISDM was correlated with smoking heaviness (cigarettes per day r = 0.63; CO r = 0.55) (Piper *et al*., 2004). Data also indicated that WISDM Total predicted outcome at both 1-week and 6-months post-quit (TTURC Tobacco Dependence Phenotype Workgroup, 2007). Thus, there is evidence that the whole scale is meaningfully related to the major dependence criteria. However, as with the NDSS, it appears that some shorter measures, such as the HSI, predict smoking heaviness and cessation likelihood as well or better than the longer WISDM (TTURC Tobacco Dependence Phenotype Work-group, 2007).

The various WISDM subscales show different patterns of relations with the dependence criteria. For instance, the Tolerance subscale was the best predictor of CO level, but the Craving, Cue Exposure/ Associative Processes, and Tolerance subscales were the best predictors of DSM-IV dependence when entered together into a multiple regression equation (Piper *et al*., 2004). One study found that although the total score was not a significant predictor of relapse after controlling for treatment, the combination of Automaticity, Behavioral Choice/-Melioration, Cognitive Enhancement, and Negative Reinforcement subscales all predicted relapse by the end of treatment in a multivariate model (Piper *et al*., 2004). Data from two different smoking cessation trials found that WISDM Automaticity and Tolerance were predictive of outcome at 6-months post-quit (TTURC Tobacco Dependence Phenotype Workgroup, 2007).

Heritability: There is evidence that the Taste/Sensory Properties subscale was significantly related to a genetic variant that determines sensitivity to bitter tastes (the phenylthiocarbamide (PTC) haplotype) (Cannon *et al*., 2005). Data have also revealed a significant relation between the WISDM Tolerance subscale with the ratio of 3-hydroxycotinine to cotinine (Piper *et al*., 2008). These data suggest that some WISDM subscales may code for biological diversity so as to permit genetic mapping.

Summary: Like the CDS and the NDSS, the WISDM is a relatively new scale and it is too soon to draw firm conclusions about its validity relative to other dependence instruments. However, data reveal that some of its subscales predict smoking heaviness measures and smoking cessation likelihood (Piper *et al*., 2008). There is also evidence that the various subscales of the measure are differentially related to various dependence criteria (TTURC Tobacco Dependence Phenotype Workgroup, 2007; Piper *et al*., 2008), suggesting that this measure is able to capture different dimensions or aspects of dependence. However, there is evidence that the WISDM is not able to predict the major dependence criteria of smoking heaviness or cessation likelihood better than shorter measures (TTURC Tobacco Dependence Phenotype Workgroup, 2007). Some WISDM subscales have been related to various dependence-linked genetic components. It is important to note that the WISDM research has been done on clinical populations and it is not known how well these results would generalize to population-based samples.

Summary:

Assessment of cigarette-induced nicotine dependence is an important goal for three reasons. First, the human and economic

costs of cigarette-induced, nicotine dependence is significant. Second, only a portion of cigarette smokers are "dependent" (as defined by traditional instruments), and those who are dependent are indeed distinguishable from other smokers on the basis of factors, such as likelihood of future cessation and amount smoked daily. Finally, cigarette-induced nicotine dependence may serve to moderate individuals' responses to different tobacco control programmes and policies, as well as the proximal and distal effects of these interventions.

It is important to note that there is considerable evidence that the various measures of nicotine dependence are not highly related to one another, and can have very different relations with validity measures (Hughes *et al.,* 2004a; Piper *et al.*, 2006). Thus it is critical that investigators select measure(s) that are psychometrically sound, appropriate for the intended population, and target the constructs in which the researchers are interested. If the goal is to assess a central core of nicotine dependence as a predictor of cigarette use cessation likelihood, or as an index of associated health risks, then the FTND or HSI appear best suited for this purpose (Tables 3.22 and 3.23). These instruments are brief and have relatively impressive predictive validities, and their reading level should make them appropriate for a broad range of populations. However, it is important to note that none of the dependence measures accounts for a large proportion of variance in outcomes in cessation likelihood. This is no doubt due to the fact that cessation likelihood is affected by countless situational/environmental factors, and other person factors. In addition, if one uses a brief measure, such as the HSI, it is important to recognize that it does not tap all dependence factors. It also does not appear to predict certain core features of dependence well, such as withdrawal, and it may be inappropriate in populations that do not smoke daily or have significant restrictions on smoking (e.g. restrictions that constrain smoking in certain contexts or times of day).

There may be situations when there is a need to assess particular, relatively discrete, facets of nicotine dependence. For example, identifying specific tobacco dependence mechanisms may facilitate: identification of a more proximal phenotype (Cannon *et al.*, 2005), identification of specific dependence dimensions with which one could create treatment algorithms, monitoring of the development of tobacco dependence, or identification of a specific group of dependent tobacco users for whom a policy is particularly effective or ineffective. If this is the goal of the research, then a multifactorial measure (i.e. the NDSS and the WISDM-68, and their subscales) would be optimal, despite the fact that there is little evidence for incremental validity in predicting important public health outcomes. However, the relative lack of validity information on these scales may mean that researchers should use these instruments only in the context of exploratory research. They might be most appropriate for research addressing etiology and cultural or population-based differences in smoking determinants.

Measures of smokeless tobacco-induced nicotine dependence

Like cigarettes, smokeless tobacco (ST) products contain nicotine, although the levels vary considerably across products (Hatsukami *et al.*, 1992; IARC, 2007b). Data on patterns of use of ST support the conclusion that many users are nicotine dependent (Henningfield *et al.* 1997; IARC, 2007b). Many ST users experience withdrawal symptoms upon abstinence (Hatsukami *et al.*, 1992; 1999). Studies have used a biomarker of nicotine uptake, cotinine, to show that daily users of ST exhibit levels of nicotine absorption that are equivalent to daily cigarette smokers (Gritz *et al.*, 1981).

Dependence on smokeless tobacco has often been assessed with questionnaires derived from FTND, with the addition of specific items, in particular, swallowing the tobacco juice (Boyle *et al.*, 1995; Ebbert *et al.,* 2006). In three different samples, use of ST within 30 minutes of waking and swallowing the tobacco juice were

Construct	Tobacco Dependence
Measure 1	Fagerström Test of Nicotine Dependence (FTND) – 6 items
Source	Heatherton *et al.*, 1991
Variation	It is possible to change the wording of the items to be culturally appropriate or to reflect non-cigarette tobacco use. However, these changes may affect the reliability and validity of the data obtained.
Validity	• Predicts both behavioural (e.g. lifetime amount smoked) and biochemical (e.g. CO, cotinine) indices of smoking in multiple countries • Predicts cessation • Evidence of linkage to specific dependence-linked genetic variants
Comments	This measure is recommended as an assessment of dependence's ability to predict cessation and heavy use • Brief and well-known • Strong predictive validity of heavy use and cessation • Internal consistency is modest, which may reflect a 2-factor structure • Some items may be influenced by smoking restrictions in the environment • Has been translated into a number of different languages
Measure 2	Heaviness of Smoking Index (HSI) – 2 items from the FTND: number of cigarettes smoked per day and time to first cigarette in the morning
Source	Kozlowski *et al.*, 1994
Variation	It is possible to change the wording of the items to be culturally appropriate or to reflect non-cigarette tobacco use. However, these changes may affect the reliability and validity of the data obtained.
Validity	• Predicts both behavioural (e.g. lifetime amount smoked) and biochemical (e.g. CO, cotinine) indices of smoking in multiple countries • Predicts cessation – the HSI appears to be the strongest predictor of cessation, accounts for much of the predictive validity of the FNTD • Highly heritable (71%) and linked to specific dependence-linked genetic variants
Comments	This measure is recommended as the most efficient measure to assess dependence's ability to predict cessation. • Brief • Using this measure may only involve the addition of item (time to first cigarette) if number of cigarettes per day is already being collected • Strong predictive validity of heavy use and cessation • Items may be influenced by smoking restrictions in the environment • Has been translated into a number of different languages

Table 3.22 Measures of Cigarette-Induced Nicotine Dependence

Construct	Tobacco Dependence
Measure	Fagerström Test of Nicotine Dependence (FTND) – 6 items
Sources	Boyle *et al.*, 1995; Ebbert *et al.*, 2006
Variation	It is possible to change the wording of the items to be culturally appropriate or to reflect non-cigarette tobacco use. However, these changes may affect the reliability and validity of the data obtained.
Validity	• Predicts both behavioural (e.g. lifetime amount smoked) and biochemical (e.g. CO, cotinine) indices of smoking in multiple countries • Predicts cessation • Evidence of linkage to specific dependence-linked genetic variants
Comments	This measure is recommended as an assessment of dependence's ability to predict cessation and heavy use • Brief and well-known • Strong predictive validity of heavy use and cessation • Internal consistency is modest, which may reflect a 2-factor structure • Some items may be influenced by smoking restrictions in the environment • Has been translated into a number of different languages

Table 3.23 A Measure of Smokeless Tobacco-Induced Nicotine Dependence

the variables most consistently associated with cotinine level (Boyle *et al.*, 1995) (see Appendix 8 for the items and scoring).

Summary:

Like cigarettes, smokeless tobacco can result in nicotine dependence. While less research has been done to validate self-report measures of ST-induced nicotine dependence, questionnaires derived from FTND appear to provide a means for identifying ST users who are nicotine dependent.

Summary and recommendations

Nicotine dependence is an important construct to assess as a moderator for the effects of tobacco control programmes and policies. In this section the evidence was reviewed on the validity of various proposed measures of cigarette and smokeless tobacco nicotine dependence. For cigarette smoking, the 2-item Heaviness of Smoking Index is recommended for use in population level studies. If only a single item measure is possible, the use of "time to first cigarette in the morning" is recommended. For smokeless tobacco, the FTND-ST appears to be a useful measure of nicotine dependence.