SYSTEMATIC REVIEW

Measurement properties of the translations of instruments evaluating the subjective effects of tobacco- and nicotine-containing products: a systematic review of the literature

[version 1; peer review: 2 approved with reservations]

Catherine Acquadro ¹, Céline Desvignes-Gleizes², Nelly Mainy³, Matthew Hankins³, Rolf Weitkunat³, Christelle Chrea ³

¹Patient-Centered Sciences, ICON plc, Lyon, 69003, France

²Mapi Research Trust, Lyon, 69003, France

³PMI R&D, Philip Morris Products S.A, Neuchâtel, CH-2000, Switzerland

V1 First published: 04 Dec 2019, 8:2056
<https://doi.org/10.12688/f1000research.20595.1>
 Latest published: 04 Dec 2019, 8:2056
<https://doi.org/10.12688/f1000research.20595.1>

Abstract

Background: Several instruments are widely used for assessing dependence, craving, withdrawal symptoms, and reinforcing effects in users of tobacco- and nicotine-containing products (TNP), including the Fagerström Test for Nicotine Dependence (FTND), Questionnaire of Smoking Urges, original (QSU) and brief (QSU-b) versions; Minnesota Nicotine Withdrawal Scale, original (MNWS) and revised (MNWS-R) versions; and Cigarette Evaluation Questionnaire, original (CEQ) and modified (mCEQ) versions. Although these instruments have been translated extensively, their translations and corresponding measurement properties have not been systematically assessed. This study aimed to (1) identify the translations of these instruments for which psychometric properties have been published, (2) describe the methods used for translation, and (3) describe the measurement properties and the context in which these translations were evaluated (e.g., target population and TNP used).

Methods: Embase and MEDLINE databases were systematically searched.

Results: While no information could be found for the CEQ/mCEQ, several translations were available for the remaining instruments: FTND, 25; QSU and QSU-b, 4 each; QSU (12-item version), 1; MNWS, 4; and MNWS-R, 1. Cigarette smokers represented the main target population in which the validation studies were conducted. Information about the translation process was reported for 25 translations. In most cases, the properties of the translations mirrored those of the originals. Differential item functioning was explored in only one case.

Open Peer Review

Approval Status ? ?

	1	2
version 1		
04 Dec 2019	view	view

1. **Jennifer Rose**, Wesleyan University, Middletown, USA
2. **Maria Rosaria Galanti** , Karolinska Institutet, Stockholm, Sweden

Any reports and responses or comments on the article can be found at the end of the article.

Conclusions: There are few publications describing the measurement properties of the translations of the FTND, QSU/QU-b, and MNWS/MNWS-R. None of these translations have been validated for TNPs other than cigarettes, which suggests the need for greater development and validation of instruments in this area.

Keywords

Fagerström Test for Nicotine Dependence, Questionnaire of Smoking Urges, Minnesota Nicotine Withdrawal Scale, Cigarette Evaluation Questionnaire, translations, measurement properties; cross-cultural equivalence, tobacco research

Corresponding author: Catherine Acquadro (Catherine.Acquadro@iconplc.com)

Author roles: **Acquadro C:** Formal Analysis, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Desvignes-Gleizes C:** Conceptualization, Data Curation, Methodology, Writing – Review & Editing; **Mainy N:** Supervision, Writing – Review & Editing; **Hankins M:** Supervision, Writing – Review & Editing; **Weitekunat R:** Supervision, Validation, Writing – Review & Editing; **Chrea C:** Conceptualization, Funding Acquisition, Methodology, Supervision, Validation, Writing – Review & Editing

Competing interests: CC, MH, NM, and RW are employees of Philip Morris International. All other authors are contracted by Philip Morris International for consulting services on the work presented in this communication.

Grant information: The work presented in this communication was sponsored by Philip Morris International.

Copyright: © 2019 Acquadro C *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Acquadro C, Desvignes-Gleizes C, Mainy N *et al.* **Measurement properties of the translations of instruments evaluating the subjective effects of tobacco- and nicotine-containing products: a systematic review of the literature [version 1; peer review: 2 approved with reservations]** F1000Research 2019, 8:2056 <https://doi.org/10.12688/f1000research.20595.1>

First published: 04 Dec 2019, 8:2056 <https://doi.org/10.12688/f1000research.20595.1>

List of abbreviations

ABOUT: Assessment of Behavioral Outcomes related to Tobacco and nicotine products; **BAI:** Beck Anxiety Inventory; **BDI:** Beck Depression Inventory; **CDER:** Center for Drug Evaluation and Research; **CDS:** Cigarette Dependence Scale; **CEQ:** Cigarette Evaluation Questionnaire; **CESDS:** Center for Epidemiological Studies Depression Scale; **CO:** carbon monoxide; **COSMIN:** COnsensus-based Standards for the selection of health Measurement Instruments; **CTP:** Center for Tobacco Products; **CTT:** classical test theory; **DIF:** differential item functioning; **DSM:** Diagnostic and Statistical Manual; **F:** factor; **FDA:** US Food and Drug Administration; **FTND:** Fagerström Test for Nicotine Dependence; **IRT:** item-response theory; **mCEQ:** Modified Cigarette Evaluation Questionnaire; **MNWS:** Minnesota Nicotine Withdrawal Scale; **MNWS-R:** revised version of the Minnesota Nicotine Withdrawal Scale; **MRTP:** Modified risk tobacco product; **PREP:** potential reduced exposure products; **NDSS:** Nicotine Dependence Syndrome Scale; **PRO:** Patient-reported outcomes; **PROQOLID:** Patient-Reported Outcome and Quality of Life Instruments Database; **Q:** question; **QSU:** Questionnaire of Smoking Urges; **QSU-b:** brief version of the Questionnaire of Smoking Urges; **TDS:** Tobacco Dependence Scale (TDS); **TNP:** tobacco- and nicotine-containing products.

Introduction

On June 22, 2009, the US Congress enacted a legislation (US Congress, 2009) that granted the US Food and Drug Administration (FDA) the authority to regulate tobacco products and the advertising and promotion of such products. In March 2012, the FDA Center for Tobacco Products (CTP) issued a draft guidance regulating applications for modified risk tobacco products (MRTPs) (US Department of Health and Human Services, 2012). This draft guidance mandates that applications must include scientific evidence about the effects of the products on tobacco-use behavior among current tobacco users. In particular, the guidance clearly states that submissions should present “*nonclinical and/or human studies to assess the abuse liability and the potential for misuse of the product as compared to other tobacco products on the market.*” In this guidance, the FDA defines abuse liability as “*the likelihood that individuals will develop physical and/or psychological dependence on the tobacco product.*” Physical dependence encompasses a growing tolerance to product use and/or the inception of withdrawal symptoms when product use cessation occurs. Psychological dependence is mainly characterized by craving and persistent tobacco-seeking and tobacco-use behaviors.

Several authors (Carter *et al.*, 2009; Hanson *et al.*, 2009; Institute of Medicine, 2012) have extensively reviewed measures and methods for assessing dependence, craving, withdrawal symptoms, and reinforcing effects in tobacco- and nicotine-containing product (TNP) users. They have identified some measures either widely used or recommended in tobacco research for the evaluation of tobacco products in general and MRTPs in particular. The most commonly quoted are the

Fagerström Test for Nicotine Dependence (FTND) (Fagerström, 1978; Fagerström, 2012; Heatherton *et al.*, 1991), Questionnaire of Smoking Urges (QSU) (Kozlowski *et al.*, 1996; Tiffany & Drobes, 1991), Minnesota Nicotine Withdrawal Scale (MNWS) (Cox *et al.*, 2001; Hughes, 2017; Hughes & Hatsukami, 1986; Hughes, 1992; Hughes & Hatsukami, 1998), and Cigarette Evaluation Questionnaire (CEQ) (Cappelleri *et al.*, 2007; Rose *et al.*, 1998; Westman *et al.*, 1992). In terms of tobacco dependence assessment, the US Institute of Medicine report (2012) acknowledges that the FTND appears to contribute to a more precise estimation of dependence than the “Diagnostic and Statistical Manual of Mental Disorders” criteria. Regarding withdrawal symptoms, the same report mentions the MNWS as a well-characterized measure for assessing reduction of withdrawal symptoms. In their paper describing traditional tools and methods for abuse liability assessment, Carter *et al.* (2009) make references to the FTND for assessing the magnitude of nicotine dependence, the MNWS for assessing nicotine withdrawal signs and symptoms, and the QSU for measuring craving. In their review on questionnaires for measuring the subjective effects of potential reduced exposure products (PREP), Hanson *et al.* (2009) conclude that the most widely used scale has been the MNWS or its revised version (MNWS-R), followed by the QSU. They recommend that, at a minimum, these two scales should be included in a battery of assessment tests for PREPs. In addition, the authors also mention the CEQ and its modified version (mCEQ) as being widely used.

Table 1 describes these measures (FTND, QSU, MNWS, and CEQ) and their evolution over time (QSU-brief [QSU-b], MNWS-R, and mCEQ).

The content and measurement properties of the original versions of these four measures in different settings and populations have been documented in the literature (Buckley *et al.*, 2005; Burling & Burling, 2003; Cappelleri *et al.*, 2005; Cappelleri *et al.*, 2007; Cox *et al.*, 2001; Davies *et al.*, 2000; Etter & Hughes, 2006; Fagerström *et al.*, 2012; Haddock *et al.*, 1999; Heatherton *et al.*, 1991; Hudmon *et al.*, 2005; Hughes, 2017; Hughes & Hatsukami, 1986; Hughes, 1992; Hughes & Hatsukami, 1998; Hughes *et al.*, 2004; Kozlowski *et al.*, 1994; Kozlowski *et al.*, 1996; Okuyemi *et al.*, 2007; Payne *et al.*, 1994; Pomerleau *et al.*, 1994; Radzius *et al.*, 2003; Rose *et al.*, 1998; Sledjeski *et al.*, 2007; Steinberg *et al.*, 2005; Tiffany & Drobes, 1991; Toll *et al.*, 2006; Toll *et al.*, 2004; Weinberger *et al.*, 2007; West & Ussher, 2010; West *et al.*, 2006; Westman *et al.*, 1992). However, in the context of globalization of tobacco consumption (with an estimate of almost 972 million smokers in the world in 2012 (Ng *et al.*, 2014) and internationalization of tobacco control (Reubi & Berridge, 2016), it is fundamental to obtain information about the measurement properties of the translations of these self-report instruments since it is crucial to ensure cross-cultural equivalence between the original versions and their translations (Petersen *et al.*, 2003; Regnault & Herdman, 2015; US Department of Health and Human Services, 2009).

Table 1. Brief description of the Fagerström Test for Nicotine Dependence (FTND), Questionnaire of Smoking Urges (QSU), Minnesota Nicotine Withdrawal Scale (MNWS), and Cigarette Evaluation Questionnaire (CEQ) and the corresponding brief, revised, or modified versions.

Measure	History/content	Response scale
FTND	<p>Revised version of the Fagerström Tolerance Questionnaire (FTQ), which was developed in 1978 to provide a short, convenient self-reported measure of nicotine dependence (Fagerström, 1978).</p> <p>It includes 6 questions.</p> <p>In 2012, in an effort to integrate the total dependence panorama and the fact that the FTND has not been validated against all forms of tobacco use—from cigarettes to smokeless tobacco—Dr. Fagerström suggested that the FTND be renamed the Fagerström Test for Cigarette Dependence (FTCD) (Fagerström, 2012).</p>	<p>- Yes/No for Q2, Q5, and Q6</p> <p>- Four options for Q1 and Q4: ✓ Q1: Within 5 minutes, 6–30 minutes, 31–60 minutes, after 60 minutes ✓ Q4: ≤10, 11–20, 21–30, ≥31</p> <p>- Two options for Q3: The first one in the morning, All others.</p> <p>The scores enable the classification of nicotine dependence into five levels: very low (0–2 points); low (3–4 points); moderate (5 points); high (6–7 points); and very high (8–10 points).</p>
QSU (32 items)	<p>Developed in 1991 (Tiffany & Drobes, 1991) to assess the potential multidimensional nature of craving report. It originally consisted of 32 items. A two-factor item structure was shown, with factor 1 representing a desire and intention to smoke, with smoking anticipated as pleasurable (15 items of which 10 are negatively keyed), and factor 2 representing an anticipation of relief from negative affect and nicotine withdrawal, with an urgent desire to smoke (11 items positively keyed). The type of desire represented on the first factor was characterized by items such as “I have an urge for a cigarette,” and “I have no desire for a cigarette right now” (negatively keyed). In contrast, the second factor seemed to represent a more pressing and urgent state of desire as indicated by items such as “All I want right now is a cigarette,” and “My desire to smoke seems overpowering.”</p>	<p>Seven-point Likert-type scale (1 = <i>strongly disagree</i> and 7 = <i>strongly agree</i>)</p>
QSU (12 items)	<p>Kozlowski <i>et al.</i> (1996) proposed an alternative model using the 12 most robust items from the original analysis.</p>	<p>Seven-point Likert-type scale (1 = <i>strongly disagree</i> and 7 = <i>strongly agree</i>)</p>
QSU Brief Version (QSU-b) (10 items)	<p>Cox <i>et al.</i> (2001) developed a 10-item version, which they called the QSU-Brief (QSU-b), to facilitate use in laboratory and clinical settings. When used to derive a global measure of craving, QSU-b displayed high internal consistency across settings, providing a reliable assessment of desire to smoke. Factor analyses showed two distinct manifestations of verbal report of craving. Factor 1 represented a strong desire and intention to smoke, with smoking perceived as rewarding for active smokers, while factor 2 reflected an anticipation of relief from negative affect and an urgent desire to smoke.</p>	<p>100-point scale ranging (0 = <i>strongly disagree</i> and 100 = <i>strongly agree</i>)</p>
MNWS and MNWS revised version (MNWS-R)	<p>The MNWS was developed in 1986 when Hughes & Hatsukami (1986) provided a detailed description of tobacco withdrawal and listed several signs and symptoms to be assessed (seven to nine items) rated on a 4-point scale (not present, mild, moderate, or severe).</p> <p>This measure has evolved over the years (Hughes, 1992), and the scale is now composed of eight symptoms associated with nicotine withdrawal (i.e., craving, irritability, anxiety, difficulty concentrating, restlessness, increased appetite or weight gain, depression, and insomnia). In a short communication, Hughes & Hatsukami (1998) encouraged researchers to use a scale that includes only seven DSM items: depression, insomnia, irritability/frustration/anger, anxiety, difficulty concentrating, restlessness, and increased appetite/weight gain.</p> <p>Finally, a revised version was proposed—the MNWS-R (Hughes, 2017)—which includes 15 items. The first eight symptoms are well-validated items (and the ones to be used if calculating a total withdrawal discomfort score), with the first seven being the DSM original items and the eighth investigating craving. The remaining seven symptoms were considered promising candidate symptoms (impatience, constipation, dizziness, increased coughing, increased dreaming or nightmares, nausea, and sore throat).</p>	<p>Items can be rated on an ordinal scale (0 = not present, 1 = mild, 2 = moderate, and 3 = severe) or on a 0–4 scale with the additional descriptor of “slight” between not present and mild, or by using a 100-mm visual analogue scale.</p> <p>Five-point Likert scale with 0 = <i>none</i>, 1 = <i>slight</i>, 2 = <i>mild</i>, 3 = <i>moderate</i>, and 4 = <i>severe</i></p>
CEQ / Modified CEQ (mCEQ)	<p>The CEQ is a self-reported questionnaire containing 11 items covering both the reinforcing effects (i.e., smoking satisfaction, psychological reward, and enjoyment of respiratory tract sensations) and aversive effects (i.e., dizziness and nausea) of smoking (Rose <i>et al.</i>, 1998; Westman <i>et al.</i>, 1992).</p> <p>Cappelleri <i>et al.</i> (2007) developed a modified version (mCEQ) by adding one item on enjoying smoking.</p>	<p>Items are rated on a seven-point scale (1 = <i>not at all</i> and 7 = <i>extremely</i>).</p>

Abbreviations: DSM: Diagnostic and Statistical Manual.

The objectives of this paper were:

1. To identify translations of the FTND, QSU/QSU-b, MNWS/MNWS-R, and CEQ/mCEQ for which psychometric properties are available;
2. To describe the methods used for translation;
3. To describe the measurement properties and the context in which these translations were evaluated (i.e., study design, target population, and TNP used by the study population).

Methods

Search strategy

Embase and MEDLINE databases were searched (in March 2018) with no limitation in timeframe, by using the following keywords and Boolean operators: (1) translation OR language OR version or cross-cultural valid* OR internal consistency OR Cronbach's alpha OR reliability OR validation OR responsiveness OR validity, combined (AND) with (2) QSU OR Questionnaire on Smoking Urges OR Cigarette Evaluation Questionnaire OR Fagerström Test for Nicotine Dependence OR FTND OR Fagerström Test for Cigarette Dependence OR Minnesota Nicotine Withdrawing Scale OR MNWS. The combination of (1) and (2) was limited to Abstract, Human research, and English. We screened reference lists to identify supplemental pertinent studies.

Selection criteria

Abstracts retrieved through the search strategy were reviewed and excluded if they (1) did not refer to the instruments of interest; (2) referred to the original version of the instruments of interest; or (3) referred to a translation used (a) in an epidemiological or behavioral context (i.e., not reporting measurement properties) or (b) for validating another measure and not for assessing/reporting the internal consistency or structural validity of the instruments of interest. Conference abstracts were excluded.

The reference lists of the papers considered for inclusion were reviewed, and articles of interest were included if they (a) referred to a translation for which internal consistency or structural validity was assessed at minimum or (b) provided additional information on an existing translation identified through the first round of review.

Two independent reviewers performed the selection. Initial data were extracted by one reviewer and then reviewed (and complemented if needed) by another.

Measurement properties

We used the **C**onsensus-based **S**tandards for the selection of health **M**easurement **I**nstruments (COSMIN) categorization (Mokkink *et al.*, 2010a; Mokkink *et al.*, 2010b; Mokkink *et al.*, 2010c; Mokkink *et al.*, 2018), to classify the measurement properties as follows: reliability, validity, and responsiveness to change. A fourth category, sensitivity and specificity, was added where appropriate (e.g., when the instrument was used for screening).

Reliability is described as the overall consistency of a measure, i.e., the degree to which scores for subjects who have not changed are the same when the measurement is repeated over time [test–retest reliability], is done with different evaluators on the same occasion [inter-rater reliability] or with the same evaluator on different occasions [intra-rater reliability]. As for internal consistency reliability, this estimate assesses the consistency of scores across items within a measurement instrument.

Validity is the degree to which an instrument measures what it is supposed to measure and includes the following:

- Content validity: The extent to which the content of a questionnaire is an appropriate manifestation of the construct to be assessed. The key features are whether or not the items are relevant and that not important concept is missing, i.e., that the measure is comprehensive. As this review deals with translations, this part will include a description of the translation process and whether or not, on a qualitative level, the content of some items was changed to reflect cultural aspects.
- Construct validity: The extent to which the scores of a measure are in accordance with hypotheses based on the assumption that the questionnaire accurately measures the construct to be measured (Mokkink *et al.*, 2010a). We have included the following aspects in construct validity:
 - Structural validity: The degree to which the scores of an instrument are an adequate reflection of the dimensionality of the construct to be measured (Mokkink *et al.*, 2010a). Factor analysis should be performed to confirm the number of subscales present in a questionnaire.
 - Hypothesis testing: The extent to which an instrument relates to other instruments in a way that is expected if it is accurately measuring the supposed construct (i.e., in accordance with predefined hypotheses about the correlation or differences between the measures). We have included the following aspects in this category:
 - The degree to which the instrument scores correlate with changes in instruments assessing similar constructs, connected but dissimilar constructs, or unconnected constructs.
 - The degree to which the instrument scores correlate with biomarkers or measures of TNP consumption (consumption patterns).
 - Predictive validity: The degree to which the considered instrument score is predictive of a future outcome or event.
 - Cross-cultural validity: The extent to which the items performance in a translated or culturally

adapted instrument appropriately reflects the performance of the items in the original version of the instrument (Mokkink *et al.*, 2010a). This is evaluated using multi-group factor analysis or differential item functioning (DIF) by utilizing data from populations who completed the original version of the questionnaire and its translations.

Responsiveness to change (Hays & Hadorn, 1992) is the ability of an instrument to detect change over time in the construct to be measured. Responsiveness to change is considered an aspect of validity in a longitudinal context.

Sensitivity and specificity are used to evaluate the screening performance of a measure. Sensitivity is the proportion of true positives that are exactly identified, whereas specificity relates to the proportion of true negatives correctly identified. Generally, an optimal cutoff point for the score is selected to reduce the sum of false-positive and false-negative results.

Results

The search retrieved 193 articles (Table 2), of which 47 were selected for data extraction. While 46 of these articles described individual investigations on the measurement properties of translated versions of the FTND, QSU/QSU-b, and MNWS/MNWS-R, one was a review of the psychometric properties of the FTND (original and translations) (Meneses-Gaya *et al.*, 2009). No references were found on the CEQ or mCEQ. More details are presented in Figure 1 and Supplementary Material 1 (Table S1), which provides a list of references retrieved and reasons for inclusion/exclusion.

FTND results

The search retrieved 34 FTND studies and one review. We identified 25 different FTND translations (Table 3), including two different Arabic versions for Yemeni speakers [Yemeni immigrants in the UK (Kassim, *et al.*, 2012) and inhabitants of Yemen (Nakajima *et al.*, 2012)], two different Chinese versions [Taiwanese (Huang *et al.*, 2009; Huang *et al.*, 2006) and Chinese immigrants in the US (Yamada *et al.*, 2009)], two different Dutch versions (Breteler *et al.*, 2004; Vink *et al.*, 2005), two different Farsi versions (Iran) (Robabeh *et al.*, 2017; Sarbandi *et al.*, 2015), and two different Portuguese versions for Brazil (Carmo & Pueyo, 2002; de Meneses-Gaya *et al.*, 2009).

In case of different studies exploring the same language (e.g., Japanese versions reported in two independent studies or French versions reported in French and Swiss studies), we contacted the authors for clarification about the version used (personal communications by email). Authors confirmed using either an existing translation [Japanese (Kawada *et al.*, 2010)] or their own version [French (Chabrol *et al.*, 2003); Portuguese for Brazil (de Meneses-Gaya *et al.*, 2009); Spanish for Mexico (Moreno-Coutiño & Villalobos-Gallegos, 2017); and Dutch (Vink *et al.*, 2005)]. In cases where we did not receive any answer, we counted only one version. We did not include the study by de Leon *et al.* (2003), as the results were based on a combined sample of Spanish and American smokers [study reported in the review by Meneses-Gaya *et al.* (2009)].

The main objective of the majority of the studies was to investigate the psychometric properties of the FTND in a specific target language and population. In some instances, the studies aimed to develop other dependence scales by using the FTND as a reference instrument [Spanish (Becoña *et al.*, 2010); French (Etter, 2005); Italian (Grassi *et al.*, 2014); Japanese (Kawada *et al.*, 2010); and Arabic (Salameh *et al.*, 2013; Salameh & Khayat, 2014)] while also exploring the properties of the FTND (Table 3).

Supplementary Material 2 describes the sociodemographic/design characteristics and targeted country/language of each study retrieved (Table S2.1), translation process (steps and people involved if mentioned) as described in the paper (Table S2.2), and measurement properties of the translations (Table S2.3). Combustible cigarettes were the TNP evaluated in all studies. Studies in India had evaluated bidis (cigarettes made locally by wrapping coarse tobacco in dried temburni leaf) [Malayalam (Jayakrishnan *et al.*, 2012) and Hindi (Jhanjee & Sethi, 2010)].

A wide range of populations was investigated (Table S2.1). Participants were recruited from the general population in 24% of the studies (Becoña & Vázquez, 1998; Etter, 2005; John *et al.*, 2004; Klinsophon *et al.*, 2017; Nakajima *et al.*, 2012; Salameh & Khayat, 2014; Stavem *et al.*, 2008; Yamada *et al.*, 2009). While 12% of the studies investigated youth samples, such as students (de Meneses-Gaya *et al.*, 2009; Etter *et al.*, 1997; Nakajima *et al.*, 2012; Salameh & Jomaa, 2014), 20% studied patients with comorbidities or drug dependence (Becoña *et al.*, 2010; de Meneses-Gaya *et al.*, 2009; Jhanjee & Sethi, 2010;

Table 2. Results of the search strategy (Medline and Embase, March 8, 2018).

Search no.	Strategy	Results
1	Translation OR Language OR Version or Cross-Cultural Valid* OR Internal Consistency OR Cronbach alpha or Reliability OR Validation OR Responsiveness OR Validity	1862311
2	QSU OR Questionnaire on Smoking Urges OR Cigarette Evaluation Questionnaire OR Fagerström Test for Nicotine Dependence or FTND or Fagerström Test for Cigarette Dependence OR Minnesota Nicotine Withdrawing Scale or MNWS	2247
3	#1 AND #2 Limit to Abstract, Human and English	193 references without duplicates

Figure 1. PRISMA flow diagram: search and selection. Abbreviations: IC: internal consistency; Struct V: structural validity.

Mikami *et al.*, 1999; Osório Fde *et al.*, 2013; Park *et al.*, 2004; Robabeh *et al.*, 2017). The sex ratio was balanced, except in some languages or countries with predominantly male samples [e.g., Iran (Robabeh *et al.*, 2017; Sarbandi *et al.*, 2015), India (Jayakrishnan *et al.*, 2012; Jhanjee & Sethi, 2010), Japan (Kawada *et al.*, 2010; Mikami *et al.*, 1999), Korea (Park *et al.*, 2004), Malaysia (Yee *et al.*, 2011), Taiwan (Huang *et al.*, 2006, Huang *et al.*, 2009), Thailand (Klinsophon *et al.*, 2017), and the Yemeni population in the U.K. (Kassim *et al.*, 2012)]. With regard to cigarette consumption, participants in most countries were light to moderate smokers, except in India (Jhanjee & Sethi, 2010), Italy (Ferketich *et al.*, 2008; Grassi *et al.*, 2014; Svicher *et al.*, 2018), Japan (Kawada *et al.*, 2010; Mikami *et al.*, 1999),

Spain (Becoña & Vázquez, 1998), Taiwan (Huang *et al.*, 2006, Huang *et al.*, 2009), and Turkey (Uysal *et al.*, 2004; Uysal *et al.*, 2015), where the samples included a mix of moderate to heavy smokers. The original version of the FTND (Heatherton *et al.*, 1991) was developed with a sample of 254 adult visitors (male, 111; female, 143) at the Ontario Science Centre, who ranged in age from 17 to 77 years (mean age, 33.5 ± 12.7 years) and smoked an average of 20.7 cigarettes per day.

Translation process. Of the translations, 60% (15 out of 25) were documented with a description of the translation process used to develop each of them (Table S2.2). Out of these

Table 3. List of FTND translations and corresponding studies.

Language	Country	Study	Study objectives
Arabic	Lebanon	Salameh <i>et al.</i> , 2013	To validate the use of the FTND in a Lebanese university student population and to create the YACD scale
		Salameh & Khayat, 2014	To validate the use of the FTND in a Lebanese adult population and to create the LCD Score
Arabic	UK (Yemenite immigrants)	Kassim <i>et al.</i> , 2012	To explore the cross-cultural validity and reliability of the FTND among Arabic-speaking cigarette consumers who chew khat leaf
Arabic	Yemen	Nakajima <i>et al.</i> , 2012	To examine the reliability and validity of the FTND among concurrent users of tobacco and khat in Yemen
Chinese	Taiwan	Huang <i>et al.</i> , 2006	To examine the psychometric properties of the FTND Chinese version
		Huang <i>et al.</i> , 2009	To compare screening performances of the FTQ, FTND, and HSI
Chinese	USA	Yamada <i>et al.</i> , 2009	To investigate DIF of the English and Chinese versions of the FTND
Dutch	The Netherlands	Breteler <i>et al.</i> , 2004	To investigate the dimensionality of the FTND and explore the dimensional properties of the nicotine dependence construct by factor analysis and Rasch modeling
Dutch	The Netherlands	Vink <i>et al.</i> , 2005	To explore the performance of the FTND in a sample of daily smokers and ex-smokers
Farsi	Iran	Sarbandi <i>et al.</i> , 2015	To evaluate the psychometric properties of the Persian FTND and HSI in smokers
Farsi	Iran	Robabeh <i>et al.</i> , 2017	To verify the usefulness of the Persian version of the FTND in patients with opioid-use disorder/cigarette smokers undergoing methadone maintenance treatment
French	France	Chabrol <i>et al.</i> , 2003	To report the first study of the factorial structure of the FTND in a French population
French	Switzerland	Etter <i>et al.</i> , 1999	To assess the validity of the FTND and HSI in a population of relatively light smokers
		Etter, 2005	To compare the psychometrics of the CDS-12, FTND, CDS-5, and HSI
German	Germany	John <i>et al.</i> , 2004	To present the results of using a short version of the FTND in two population samples
Hindi	India	Jhanjee & Sethi, 2010	To verify the usefulness of the FTND in an Indian sample of daily smokers
Italian	Italy	Ferketich <i>et al.</i> , 2008	To examine the properties of the Italian version of the FTND
		Grassi <i>et al.</i> , 2014	To test the psychometric properties of an Italian version of the Severity of Dependence Scale with the FTND as a comparative measure
		Svicher <i>et al.</i> , 2018	To examine the psychometrics properties of the FTCD and HSI through IRT
Japanese	Japan	Mikami <i>et al.</i> , 1999	To examine the reliability and validity of the FTND in patients with smoking-related cancers
		Kawada <i>et al.</i> , 2010	To validate a 100-point scale for evaluating perceived tobacco dependence with the FTND as a comparative measure
Korean	Korea	Park <i>et al.</i> , 2004	To assess the validity of the Korean FTND
Malay	Malaysia	Yee <i>et al.</i> , 2011	To evaluate the validity and reliability of the Malay FTND version
Malayalam	India	Jayakrishnan <i>et al.</i> , 2012	To assess nicotine dependence among smokers in a selected rural population in Kerala, India
Norway	Norwegian	Stavem <i>et al.</i> , 2008	To compare the properties of four measures of dependence to nicotine/tobacco: the CDS-12, the FTND, and two shorter versions of the same measures
Portuguese	Brazil	Carmo & Pueyo, 2002	To present the results of an adaptation of the Brazilian version of the FTND
		Osório Fde <i>et al.</i> , 2013	To conduct a psychometric study of the FTND
Portuguese	Brazil	de Meneses Gaya <i>et al.</i> , 2009	To examine the psychometric properties of the Brazilian versions of the FTND and HSI

Language	Country	Study	Study objectives
Spanish	Spain	Becoña & Vázquez, 1998	To verify the usefulness of the FTND in a sample of Spanish smokers
		Becoña <i>et al.</i> , 2010	To validate the Spanish NDSS with the FTND as a comparative measure
Spanish	Mexico	Moreno-Coutiño & Villalobos-Gallegos, 2017	To assess the psychometric properties of the FTND in Spanish Mexican speakers
Thai	Thailand	Klinsophon <i>et al.</i> , 2017	To evaluate the test–retest reliability and internal consistency of the Thai version of the FTND
Turkish	Turkey	Uysal <i>et al.</i> , 2004	To study the reliability and present a factor analysis of the FTND in a Turkish sample
		Uysal <i>et al.</i> , 2015	To assess the psychometric properties of the Turkish version of the FTND

Abbreviations: CDS-5: Cigarette Dependence Scale (5-item version); CDS-12: Cigarette Dependence Scale (12-item version); DIF: differential item functioning; FTCD: Fagerström Test of Cigarette Dependence; FTND: Fagerström Test for Nicotine Dependence; FTQ: Fagerström Tolerance Scale; HSI: Heaviness of Smoking Index; IRT: item response theory; LCD: Lebanese Cigarette Dependence; NDSS: Nicotine Dependence Syndrome Scale; YACD: Young Adults' Cigarette Dependence.

15 translations, only three (Kassim *et al.*, 2012; Uysal *et al.*, 2004; Yee *et al.*, 2011) were presented with a brief report of the difficulties encountered and solutions found. References to guidelines or recommendations were given for six translations (Becoña & Vázquez, 1998; Kassim *et al.*, 2012; Klinsophon *et al.*, 2017; Sarbandi *et al.*, 2015; Yamada *et al.*, 2009; Yee *et al.*, 2011). Descriptions of the translation process were either minimal, with only a mention of the steps performed (Becoña & Vázquez, 1998; Jayakrishnan *et al.*, 2012; Jhanjee & Sethi, 2010; Mikami *et al.*, 1999; Nakajima *et al.*, 2012; Robabeh *et al.*, 2017), or more detailed, with information about the people involved in the process (Etter *et al.*, 1999; Kassim *et al.*, 2012; Klinsophon *et al.*, 2017; Park *et al.*, 2004; Salameh *et al.*, 2013; Salameh & Khayat, 2014; Sarbandi *et al.*, 2015; Uysal *et al.*, 2004; Yamada *et al.*, 2009; Yee *et al.*, 2011). Except for one translation, i.e., French for Switzerland (Etter *et al.*, 1999), all teams included in their process a backward translation step (i.e., the translation of the target language version back to the source language, English).

Measurement properties All translations were assessed using the classical test theory (CTT), except for the Chinese version developed for immigrants to the US (Yamada *et al.*, 2009), which was assessed only by an item response theory (IRT)-based approach. The Dutch (Breteler *et al.*, 2004) and Italian versions (Svicher *et al.*, 2018) were assessed by IRT supplemented with CTT. Table S2.3 provides detailed information on all properties.

Measurement equivalence was explored for only one translation—Chinese for immigrants to the US (Yamada *et al.*, 2009)—for which DIF was examined by using IRT. Question (Q) 2 (difficult to refrain) showed a significantly substantial DIF, indicating that users of the Chinese version were more likely to support this item and to report more difficulty in refraining from smoking at several public places even after controlling for the nicotine-dependence level. As this DIF item

in the Chinese version contributed minimally at the aggregate level, the authors concluded that its impact was negligible on scale scores. Neither unidimensional nor multidimensional results showed DIF for Q1 (time to first cigarette) or Q3 (cigarette hated most to give up), indicating that these two items are DIF-free. Authors concluded that these two items should be retained in the FTND to enable comparison between Chinese- and English-speaking smokers.

Structural validity was documented for 72% of the translations (18 of 25): Arabic for Lebanon (Salameh *et al.*, 2013; Salameh & Khayat, 2014); Arabic for Yemenites living in the UK (Kassim *et al.*, 2012) and Yemen (Nakajima *et al.*, 2012); Chinese for Taiwan (Huang *et al.*, 2006) and the US (Yamada *et al.*, 2009); Dutch (Breteler *et al.*, 2004); both Farsi versions (Robabeh *et al.*, 2017; Sarbandi *et al.*, 2015); French for France (Chabrol *et al.*, 2003) and Switzerland (Etter, 2005; Etter *et al.*, 1999); German (John *et al.*, 2004); Hindi (Jhanjee & Sethi, 2010); Italian (Grassi *et al.*, 2014; Svicher *et al.*, 2018); Korean (Park *et al.*, 2004); Malay (Yee *et al.*, 2011); Portuguese for Brazil (de Meneses-Gaya *et al.*, 2009); Spanish for Mexico (Moreno-Coutiño & Villalobos-Gallegos, 2017); and Turkish (Uysal *et al.*, 2004; Uysal *et al.*, 2015). Six of the translations had a monofactorial structure, similar to their original versions (Heatherton *et al.*, 1991): Arabic (Lebanon) (Salameh *et al.*, 2013); Farsi (Sarbandi *et al.*, 2015); French for France (Chabrol *et al.*, 2003) and Switzerland (Etter, 2005; Etter *et al.*, 1999); Italian (Svicher *et al.*, 2018); and Spanish for Mexico (Moreno-Coutiño & Villalobos-Gallegos, 2017). The bifactorial structure, described by Haddock *et al.* (1999) and Radzius *et al.* (2003) (i.e., one factor labeled “Smoking Pattern” with Q1, Q2, Q4, and Q6 and the other factor labeled “Morning Pattern” with Q3 and Q5, without and with cross-loading of Q1), was replicated in six studies: Chinese for Taiwan (Huang *et al.*, 2006) and the US (Yamada *et al.*, 2009); Farsi (Robabeh *et al.*, 2017); German (John *et al.*, 2004); Korean (Park *et al.*, 2004); and

Portuguese for Brazil (de Meneses-Gaya *et al.*, 2009). For the Arabic version for Lebanon (different from the original version), Salameh & Khayat (2014) reported a bifactorial structure in a general population sample but a monofactorial structure in a student sample (Salameh *et al.*, 2013).

Internal consistency was explored for all translations. In monofactorial structures, Cronbach's alpha ranged from 0.52 (Thai version, Klinphon *et al.*, 2017) to 0.82 (Portuguese for Brazil, Osório Fde *et al.*, 2013). For the French version (France, Chabrol *et al.*, 2003), Cronbach's alpha was 0.86 after deletion of Q3, which led the authors to propose a revised FTND without Q3. Removing Q3 improved the Cronbach's alpha of the French for Switzerland (Etter *et al.*, 1999), Portuguese for Brazil (Carmo & Pueyo, 2002) (slightly), and Turkish (Uysal *et al.*, 2004) versions as well. In the original version (Heatherton *et al.*, 1991), the alpha value was low (0.61), and Q3 loaded less than 0.30 (i.e., 0.23).

Test-retest reliability was assessed for 40% of the translations (10 of 25)—Dutch (Vink *et al.*, 2005); French (Switzerland) (Etter *et al.*, 1999); Farsi (Sarbandi *et al.*, 2015); Japanese (Mikami *et al.*, 1999); Malay (Yee *et al.*, 2011); Malayalam (Jayakrishnan *et al.*, 2012); Norwegian (Stavem *et al.*, 2008); Portuguese for Brazil (Carmo & Pueyo, 2002; de Meneses-Gaya *et al.*, 2009; Osório Fde *et al.*, 2013); Thai (Klinphon *et al.*, 2017); and Turkish (Uysal *et al.*, 2004)—with coefficients ranging from 0.50 (Yee *et al.*, 2011) to 0.92 (de Meneses-Gaya *et al.*, 2009) and intervals from 1 week (Moreno-Coutiño & Villalobos-Gallegos, 2017; Yee *et al.*, 2011) to 1.8 years (Vink *et al.*, 2005). In comparison, test-retest reliability analyses conducted on the original version showed coefficients ranging from 0.65 (smokers with schizophrenia, interval not specified) (Weinberger *et al.*, 2007) to 0.87 (a sample of young smokers during military training, 6-week interval) (Haddock *et al.*, 1999).

Inter-rater reliability was investigated for only one translation—Portuguese for Brazil (de Meneses-Gaya *et al.*, 2009)—with the authors reporting an intraclass coefficient of 0.99 (95% confidence interval [CI]: 0.98–1.0) for two raters (sample size, 40).

Correlations with biomarkers and consumption patterns were documented for 52% of the translations (13 of 25): Arabic (Lebanon) (Salameh *et al.*, 2013; Salameh & Khayat, 2014); Chinese (Taiwan) (Huang *et al.*, 2006; Huang *et al.*, 2009); Dutch (Vink *et al.*, 2005); Farsi (Sarbandi *et al.*, 2015); French (Switzerland) (Etter *et al.*, 1999); Hindi (Jhanjee & Sethi, 2010); Italian (Ferketich *et al.*, 2008; Grassi *et al.*, 2014); Japanese (Kawada *et al.*, 2010); Korean (Park *et al.*, 2004); Malay (Yee *et al.*, 2011); Malayalam (Jayakrishnan *et al.*, 2012); Norwegian (Stavem *et al.*, 2008); and Spanish [Spain (Becoña & Vázquez, 1998) and Mexico (Moreno-Coutiño & Villalobos-Gallegos, 2017)]. Correlations with biomarkers such as exhaled carbon monoxide and salivary and urinary cotinine levels were explored, as were correlations with duration of smoking (years), packs/year, packs/day, age of regular

smoking, age of first cigarette, and willingness to pay for a cigarette after a day of abstinence. Correlations with biomarkers ranged from weak to moderate [e.g., CO: $r = 0.288$ (Salameh & Khayat, 2014) to 0.535 (Salameh *et al.*, 2013)], in agreement with the correlations reported in studies that had used the original FTND [CO: $r = 0.210$ (Steinberg *et al.*, 2005), 0.40 (Buckley *et al.*, 2005), and 0.59 (Burling & Burling, 2003)].

Correlations with self-reported instruments investigating similar or dissimilar constructs was explored for 24% of the translations (6 of 25): Farsi (Robabeh *et al.*, 2017); French for Switzerland (Etter *et al.*, 1999); Japanese (Kawada *et al.*, 2010; Mikami *et al.*, 1999); Norwegian (Stavem *et al.*, 2008); and Spanish [Spain (Becoña *et al.*, 2010) and Mexico (Moreno-Coutiño & Villalobos-Gallegos, 2017)]. Correlations with the following scales were investigated: Beck Anxiety Inventory (BAI) and Beck Depression Inventory (BDI) ($r = 0.091$ and 0.116, respectively) (Moreno-Coutiño & Villalobos-Gallegos, 2017); Cigarette Dependence Scale (CDS) 12 and 5 ($r = 0.60$ and 0.72, respectively) (Stavem *et al.*, 2008); Diagnostic and Statistical Manual (DSM) diagnostic criteria for nicotine dependence (DSM II-R; $r = 0.70$) (Mikami *et al.*, 1999); DSM-V nicotine dependence (no correlation) (Robabeh *et al.*, 2017); Nicotine Dependence Syndrome Scale (NDSS) ($r = 0.58$) (Becoña *et al.*, 2010); Structured Clinical Interview for DSM-IV ($r = 0.38$) (Becoña *et al.*, 2010); Tobacco Dependence Scale (TDS) ($r = 0.352$) (Kawada *et al.*, 2010); and withdrawal symptoms (irritability, sensation of relief, and embarrassment: relative validity 93, 100, and 100, respectively) (Etter *et al.*, 1999). For the original version, Pomerleau *et al.* (1994) showed a correlation with the Classification of Smoking by Motives addictive factor ($r = 0.53$). No relationship was detected with depression measured by using the Center for Epidemiological Studies Depression Scale (CESDS) ($r = -0.24$).

Predictive validity was investigated for 12% of the translations (3 of 25): French for Switzerland (Etter, 2005); Italian (Ferketich *et al.*, 2008); and Spanish for Mexico (Moreno-Coutiño & Villalobos-Gallegos, 2017). The FTND predicted abstinence only at 7 weeks (Ferketich *et al.*, 2008) and not in the long term (Etter, 2005; Moreno-Coutiño & Villalobos-Gallegos, 2017). In contrast, Kozlowski *et al.* (1994) showed that the FTND could predict smoking cessation to a small degree (Study 2: 16-month follow-up, $r = -0.11$).

Sensitivity/specificity (Se/Sp) was explored for 28% of the translations (7 of 25): Chinese for Taiwan (Huang *et al.*, 2009); Italian (Svicher *et al.*, 2018); Japanese (Mikami *et al.*, 1999); Malay (Yee *et al.*, 2011); Portuguese for Brazil (de Meneses-Gaya *et al.*, 2009; Osório Fde *et al.*, 2013); and Spanish for Spain (Becoña *et al.*, 2010). Measures of reference varied among the studies: biomarkers (Huang *et al.*, 2009: salivary cotinine – cutoff score = +4, Se = 76.2%, Sp = 67.5%); status (Svicher *et al.*, 2018); measures of dependence (Becoña *et al.*, 2010: score under the curve, 0.69; Osório Fde *et al.*, 2013: cutoff score = +2, Se = 76%, Sp = 91%; de Meneses-Gaya *et al.*, 2009: cutoff score = +4, Se = 80%, Sp = 74%; Mikami *et al.*, 1999: cutoff score = +5/6, Se = 75%, Sp = 80%); and others (Yee *et al.*, 2011).

Responsiveness to change has never been assessed.

QSU/QSU-b results

We retrieved four translations for the QSU, four for the QSU-b, and one for the QSU-12 (Table 4).

Table 5 presents the general characteristics of each study. Combustible cigarettes were the TNP evaluated in all studies. Nearly all studies involved moderate smoker samples, on average. In comparison, the original QSU (Tiffany & Drobles, 1991) and QSU-b (Cox *et al.*, 2001) were developed with heavy smoker samples. Most studies had recruited samples from the general population, excepting studies in Spain [QSU (Cepeda-Benito *et al.*, 2004) and QSU-b (Cepeda-Benito & Reig-Ferrer, 2004)] and the Netherlands [QSU-b (Littel *et al.*, 2011)], which had recruited students. The sex ratio was balanced in all countries except China [QSU-b (Yu *et al.*, 2010)] and Malaysia [QSU-b (Blebil *et al.*, 2015)], where men were predominant (98% and 99%, respectively) and Spain [QSU (Cepeda-Benito *et al.*, 2004)], where the sample comprised a majority of women (81.5%).

In comparison, the original version of the QSU (Tiffany & Drobles, 1991) was developed in a sample of 230 daily cigarette smokers (141 men and 89 women) assigned to one of three levels of deprivation (0, 1, or 6 hours). The mean participant age in each subgroup was 20.91, 20.64, and 22.73 years (SD not shown), and the consumption rate was 23.3, 21.28, and 22.36 cig/day, respectively (SD not shown). The QSU-b (Cox *et al.*, 2001) was developed in two populations: Study 1 included 221 continuing smokers (111 men and 110 women; mean age, 30.23 ± 10.27 years; consumption rate,

26.95 ± 10.68 cig/day), while Study 2 included 112 smokers who were contemplating quitting (49 men and 63 women; mean age, 43.15 ± 11.60 years; consumption rate, 27.82 ± 12.57 cig/day).

Translation process. The translation process (Table 6) was not described for three translations: German QSU (Müller *et al.*, 2001); French QSU-12 (Dethier *et al.*, 2014); and Dutch QSU-b (Littel *et al.*, 2011). References to guidelines were provided for three translations: Brazilian (Araujo *et al.*, 2006); Chinese (Yu *et al.*, 2010); and Malay (Blebil *et al.*, 2015). The description for the Chinese QSU-b was minimal. Only the Brazilian team had provided some insight into the problems that arose during translation and the solutions they found.

Measurement properties. Table 7 reports the measurement properties explored for each translation. The results of the Spanish QSU-b (Cepeda-Benito & Reig-Ferrer, 2004) are not included, as the version developed by the authors is not a translation of the QSU-b but a new version with completely different content derived from the QSU, with items being positively keyed. It, therefore, represents a new version not comparable to the original US version of the QSU-b.

Measurement equivalence using DIF was never assessed. Structural validity was explored for all translations:

- QSU: The Brazilian (Araujo *et al.*, 2006) and French (Guillin *et al.*, 2000) versions of the QSU have a similar bifactorial structure as the original (Tiffany & Drobles, 1991) (i.e., with factor 1 [F1] representing a desire and intention to smoke, with smoking anticipated as pleasurable [15 items of which 10 are negatively keyed], and

Table 4. List of Questionnaire of Smoking Urges (QSU)/QSU-brief version (QSU-b) translations and corresponding studies.

Measure	Language	Country	Study	Study objectives
QSU-32	French	France	Guillin <i>et al.</i> , 2000	French translation and validation of the factorial structure of the QSU
	German	Germany	Müller <i>et al.</i> , 2001	To translate the QSU into German and examine its factorial structure, reliability, and validity
	Portuguese	Brazil	Araujo <i>et al.</i> , 2006	To validate the Brazilian version of the QSU
	Spanish	Spain	Cepeda-Benito <i>et al.</i> , 2004	To evaluate the factorial structure of the QSU across American and Spanish smokers
QSU-12	French	Belgium	Dethier <i>et al.</i> , 2014	To examine the psychometric properties of the 12-item French version of the QSU
QSU-10 (QSU-b)	Chinese	China	Yu <i>et al.</i> , 2010	To evaluate the reliability and validity of the Chinese versions of the MNWS and QSU-b in Chinese smokers
	Dutch	The Netherlands	Littel <i>et al.</i> , 2011	To investigate the reliability, validity, and factorial structure of the QSU-b in a Dutch smoker sample
	Malay	Malaysia	Blebil <i>et al.</i> , 2015	To evaluate the psychometric properties of the Malaysian version of the QSU-b
	Spanish	Spain	Cepeda-Benito & Reig-Ferrer, 2004	To develop a brief version of the QSU by using an all positively worded version of the QSU to avoid the confounding interpretations that arise from mixing positively and negatively worded items

Abbreviations: MNWS: Minnesota Nicotine Withdrawal Scale.

Table 5. Sociodemographic/design characteristics and targeted country/language of studies evaluating the measurement properties of the Questionnaire of Smoking Urges (QSU)/QSU-brief version (QSU-b) translations.

Measure	Authors	Study design	Language	Country	Sample details	Subject characteristics			
						Number of participants	Sex M/F (%)	Age in years (mean $\pm$ SD/range)	Number of cigarettes per day (mean $\pm$ SD)
QSU-32	Guillin <i>et al.</i> , 2000	Cross-sectional	French	France	General population	111 (all abstinent from 1.5 to 3 h)	37.8/62.2	38.7 (SD NS)/18–74	16.6 (SD NS)
	Müller <i>et al.</i> , 2001	Cross-sectional	German	Germany	General population	129 (three abstinent groups: 50 strongly deprived (12–15 h), 48 slightly deprived (2–3 h), and 31 subjects not deprived)	61/39	28.3 $\pm$ 0.7/range NS	NS
	Araujo <i>et al.</i> , 2006	Cross-sectional	Portuguese	Brazil	General population and staff from psychiatric hospital	201 (three abstinent groups: zero min (n = 69); 30 min (n = 60); and 60 min (n = 71) of tobacco abstinence)	33/67	38.15 $\pm$ 11.93/18–65	17.17 $\pm$ 11.0
QSU-12	Cepeda-Benito <i>et al.</i> , 2004	Cross-sectional	Spanish	Spain	Undergraduate psychology students	253	8.5/81.5	21.39 (SD NS)/17–30	12.5 (SD NS)
	Dethier <i>et al.</i> , 2014	Cross-sectional	French	Belgium	General population recruited through advertisements posted in specialized French forums and research networks	230 40 (abstinent for 1 h)	37.4/62.6 47.5/52.5	32.3 $\pm$ 11.4/range NS 38.9 $\pm$ 11.2/range NS	13.1 $\pm$ 10.0 14.0 $\pm$ 5.8
QSU-10	Yu <i>et al.</i> , 2010	Cross-sectional	Chinese	China	NS	355 (abstinent for ≥ 1 and ≤ 7 days)	98/2	39.6 $\pm$ 11.6/18–65	17.8 $\pm$ 7.7
	Littel <i>et al.</i> , 2011	Cross-sectional	Dutch	The Netherlands	Participants recruited by advertisements on internet forums and communities and by fliers distributed at the Erasmus University Rotterdam	208	41.3/58.7	24.4 $\pm$ 7.9/range NS	NS
	Blebil <i>et al.</i> , 2015	Cross-sectional	Malay	Malaysia	Smokers attending the Quit Smoking Clinic in Pulau Pinang Hospital, Penang State, Malaysia.	133	99.2/0.8	47.7 $\pm$ 14.0/18–76	14.92 $\pm$ 9.1
	Cepeda-Benito & Reig-Ferrer, 2004	Cross-sectional	Spanish	Spain	Student and staff of University of Alicante; Smokers from the province of Alicante	S1: 245 S2: 225	43.5/66.5 57.2/42.8	S1: 22.24 (SD NS)/16–50 S2: 32.6 (SD NS)/15–79	S1: 11.25 $\pm$ 6.91 S2: 16.21 $\pm$ 9.46

Abbreviations: F: female; M: male; NA: not available in this paper (reference to another publication); NS: not specified in this paper; SD: standard deviation; S1: Sample 1; S2: Sample 2.

Table 6. Description of translation processes used (steps and people involved if mentioned) for the Questionnaire of Smoking Urges (QSU)/QSU-brief version (QSU-b) translations.

Measure language/country of study	Conceptual evaluation	FT step (people involved in the process and number of FTs)	Consensus on the FT step to produce one FT (people involved in process)	BT step (people involved in the process and number of BTs)	Reconciliation / consensus (people involved in process)	Pilot/test (people involved in the process)	Report (i.e., description of issues, changes made for cultural, semantic, or syntactic reasons)	Reference (if any)
QSU-32 French/France <i>Guillin et al., 2000</i>	No	✓ (NS – verbatim: “we”; 1 FT)	No	✓ (1 independent translator, 1 BT)	✓ (review by author of original: Prof. Tiffany)	No	No	No
QSU-32 Portuguese/ Brazil <i>Araujo et al., 2006</i>	?	✓ (English teacher, graduated in Literature and knowledgeable about the purpose of the translation; 1FT: FT1)	✓ (FT1 tested on 10 subjects for understanding + brainstorming by 5 individuals to test clarity)	✓FT1 backtranslated (by a native English speaker fluent in Portuguese and unaware of the purpose of the translation; 1BT: BT1)	✓ 1. BT1 retranslated in Portuguese (by a Brazilian psychologist residing in the USA, fluent in English and knowledgeable about the purpose of the translation; 1FT: FT2) 2. Comparison FT1 and FT2 (by a Committee of Expert Judges, composed of five chemical-dependency specialists and 2 validators of psychological instruments, who compared the instrument versions, verifying that their items referred to the theme “craving”)	✓ (20 subjects)	Numbers 1 to 7 were added above the Likert scale points that would visually be related to these numbers in the original scale. Due to differences in the meaning of urge and craving, the initials in the English language (QSU) were used in the name of the scale, with the phrase “Brazilian version” being added. The term “craving” was not translated as “fissura” because of the latter being a popular term that suffers from regional influences and because its use is uncommon (according to the judges of this study) when reference is made to the desire to smoke. Therefore, “craving” was translated as “strong desire” (forte desejo).	<i>Ciconelli, 1997</i> <i>Pasquali, 1998</i>
QSU-32 Spanish/Spain <i>Cepeda-Benito et al., 2004</i>	No	✓ (a Spanish native fluent in both English and Spanish; 1 FT)	No	✓ (a US native fluent in Spanish as a second language; 1BT)	✓ (FT and BT translators)	✓ (small group of smokers)	No	No
QSU-10 Chinese/China <i>Yu et al., 2010</i>	✓ (called preparation)	✓ (NS)	✓ (NS)	✓ (NS)	✓ (NS)	✓ (23 subjects)	No	<i>Wild et al., 2005</i>

Measure language/ country of study	Conceptual evaluation	FT step (people involved in the process and number of FTs)	Consensus on the FT step to produce one FT (people involved in process)	BT step (people involved in the process and number of BTs)	Reconciliation / consensus (people involved in process)	Pilot/test (people involved in the process)	Report (i.e., description of issues, changes made for cultural, semantic, or syntactic reasons)	Reference (if any)
QSU-10 Malay/Malaysia** <i>Blebil et al., 2015</i>	✓ Not formal, but authors stress the need for conceptual equivalence between the original and translation	✓ (translators from the School of Language, Literacies and Translation, Universiti Sains Malaysia; native Malaysian; 2 FT)	✓ (two native Malaysian researchers)	✓ (third translator fluent in both languages)	✓ (NS)	✓ (20 smokers)	No	<i>Guillemin et al., 1993</i> <i>Herdman, et al., 1997</i> <i>Wild et al., 2005</i>
QSU-10 Spanish/Spain <i>Cepeda-Benito & Reig-Ferrer, 2004</i>	No	✓ (a Spanish native fluent in both English and Spanish; 1 FT)	No	✓ (a US native fluent in Spanish as a second language; 1BT)	✓ (FT and BT translators)	✓ (10 smokers)	No	No

Abbreviations: BT: backward translation; FT: forward translation.

* ?: Not clearly specified in the paper, but recommended in quoted guidelines; ** Translation Process referred to as Linguistic Validation Process.

Table 7. Measurement properties of the translations of the Questionnaire of Smoking Urges (QSU)/QSU-brief version (QSU-b).

Measure Language / country of study	Translation process described in paper	Reliability Internal consistency Cronbach's alpha	Validity Structural validity Factorial analyses (EFA, CFA), ITC, IRT	Correlations with biomarkers or consumption patterns	Correlations of total score/factors score with other measures assessing similar or dissimilar constructs
QSU-32 French / France Guillin <i>et al.</i> , 2000	Yes	F1: $\alpha = 0.89$; F2: $\alpha = 0.91$	Two factors F1: Urge to smoke/craving – 15 items (1, 2, 3, 7, 12, 13, 14, 15, 18, 19, 20, 23, 24, 29, 30) F2: Pleasure to smoke – 10 items (6, 10, 11, 16, 17, 21, 22, 27, 28, 32) Cross loading: 5 items (4, 5, 9, 25, 31) Poor loading: 2 items (8, 26)	No. of cig./day: $r = 0.54$	
QSU-32 German / Germany Müller <i>et al.</i> , 2001	No	F1: $\alpha = 0.91$; F2: $\alpha = 0.87$	Two factors, both same as the original F1: Desire and intention to smoke, with smoking anticipated as being pleasurable (15 items, with 10 negatively worded) F2: Anticipation of relief from negative affect and nicotine withdrawal, with an urgent desire to smoke (11 items, all positively worded)		Craving VAS: - Non-deprived smokers: F1 and craving VAS were significantly correlated with each other before ($r = 0.55$) and after ($r = 0.60$) smoking. F2 and craving VAS were also significantly correlated with each other before ($r = 0.45$) and after ($r = 0.44$) smoking. - Deprived smokers: F1 and craving VAS were significantly correlated ($r = 0.68$) after smoking but not before ($r = 0.31$). F2 and craving VAS were significantly correlated with each other only after smoking ($r = 0.50$); before smoking, their correlation was $r = 0.02$. FTQ: F1: $r = 0.04$; F2: $r = 0.16$
QSU-32 Portuguese / Brazil Araujo <i>et al.</i> , 2006	Yes	$\alpha = 0.97$ F1: $\alpha = 0.96$; F2: $\alpha = 0.92$	Two factors F1: Urgent and overwhelming desire to smoke – 17 items (2, 3, 5, 7, 12, 13, 14, 15, 18, 19, 20, 23, 24, 25, 29, 30, 31) F2: Desire to smoke and the anticipation of smoking pleasure – 13 items (4, 6, 8, 10, 11, 16, 17, 21, 22, 26, 27, 28, 32) Cross loading: 2 items (1, 2)	No. of cig./day: F1: $r = 0.197$; F2: $r = 0.182$ Age of 1 st cig.: F1: $r = 0.026$; F2: $r = 0.051$ No. of years smoked: F1: $r = -0.085$; F2: $r = -0.117$ No. of attempts to quit: F1: $r = -0.016$; F2: $r = -0.064$ Tobacco treatment: F1: $r = 0.103$; F2: $r = 0.034$	Craving VAS: F1: $r = 0.643$; F2: $r = 0.636$ FTND sum score: F1: $r = 0.244$; F2: $r = 0.163$ Q1 FTND: F1: $r = 0.188$; F2: $r = 0.104$ Q2 FTND: F1: $r = 0.242$; F2: $r = 0.268$ BAI: F1: $r = 0.249$; F2: $r = 0.090$ BDI: F1: $r = 0.249$; F2: $r = 0.076$

Measure Language / country of study	Translation process described in paper	Reliability Internal consistency Cronbach's alpha	Validity Structural validity Factorial analyses (EFA, CFA), ITC, IRT	Correlations with biomarkers or consumption patterns	Correlations of total score/factors score with other measures assessing similar or dissimilar constructs
QSU-32 Spanish / Spain <i>Cepeda-Benito et al., 2004</i>	Yes	F1: $\alpha = 0.95$; F2: $\alpha = 0.88$	Two factors. F1 and F2, same as the original (better fit). CFA exclusively by using either the positively worded or negatively worded items of FI showed that retention of only the negatively worded items in FI substantially improved model fit.		
QSU-12 French / Belgium <i>Dethier et al., 2014</i>	No	F1: $\alpha = 0.90$; F2: $\alpha = 0.80$	Two factors. F1: Relief of negative affect – 7 items (1, 2, 4, 7, 9, 10, 12) F2: Intention and desire to smoke – 5 items (3, 5, 6, 8, 11)	No. of cig./day: Total score: $r = 0.357$; F1: $r = 0.342$; F2: $r = 0.323$ Time since last cig.: Total score: $r = -0.025$; F1: $r = -0.058$; F2: $r = -0.009$ CO: Total score: $r = -0.136$; F1: $r = 0.201$; F2: $r = -0.050$	FTND: Total score: $r = 0.375$; F1: $r = 0.399$; F2: $r = 0.303$ Loss of control associated with smoking: Total score: $r = 0.169$; F1: $r = 0.152$; F2: $r = 0.167$ Frequency of intrusive thoughts related to smoking behaviors: Total score: $r = 0.291$; F1: $r = 0.335$; F2: $r = 0.211$
QSU-10 Chinese / China <i>Yu et al., 2010</i>	Yes	A = 0.92	Two factors. F1: Desire and intention to smoke with smoking anticipated as being pleasurable – 5 items (1, 3, 6, 7, 10) F2: Anticipation of relief from negative affect and nicotine withdrawal, with urgent desire to smoke – 3 items (4, 8, 9) ITC: $r = 0.57$ to 0.85		Patient-evaluated craving scores: $r = 0.75$

Measure Language / country of study	Translation process described in paper	Reliability	Validity	Correlations with biomarkers or consumption patterns	Correlations of total score/factors score with other measures assessing similar or dissimilar constructs
		Internal consistency Cronbach's alpha	Structural validity Factorial analyses (EFA, CFA), ITC, IRT		
QSU-10 Dutch / The Netherlands <i>Littel et al., 2011</i>	No	$\alpha = 0.83$ F1: $\alpha = 0.84$; F2: $\alpha = 0.84$	Two factors. F1: Anticipation of relief from negative affect with an urgent desire to smoke – 5 items (2, 4, 5, 8, 9) F2: Desire and intention to smoke – 5 items (1, 3, 6, 7, 10) (cross-loading items 1 and 6)	No. of cig./day: $r = 0.25$	0-100 craving rating scale: $r = 0.80$ Desire VAS: $r = 0.76$ Urge VAS: $r = 0.77$ FTND: $r = 0.14$ PANAS (subsample, $n = 84$): F1: $r = 0.25$ PANAS-NA; F2: $r = 0.16$ PANAS-NA F1: $r = -0.02$ PANAS-PA; F2: $r = -0.01$ PANAS-PA SHAPS (subsample, $n = 84$): F1: $r = 0.23$; F2: $r = 0.22$ FTND: $r = 0.24$
QSU-10 Malay / Malaysia <i>Biebil et al., 2015</i>	Yes	$\alpha = 0.81$	Two factors. F1: Desire and intention to smoke, with an anticipation of pleasure from smoking – 5 items (1, 3, 6, 7, 10) F2: Anticipation of relief from negative affect with an urgent desire to smoke – 5 items (2, 4, 5, 8, 9) ITC: $r = 0.29$ to 0.71	CO: $r = 0.024$ No. of cig./day: $r = 0.30$ Duration of smoking: $r = 0.06$ Chances for quitting: $r = -0.29$ Previous quit attempts: $r = 0.15$	

Abbreviations: BAI: Beck Anxiety Inventory; Beck Depression Inventory; Cig.: cigarettes; CFA: confirmatory factorial analysis; CO: carbon monoxide; EFA: exploratory factorial analysis; F1: factor 1; F2: factor 2; FTND: Fagerström Test for Nicotine Dependence; FTQ: Fagerström Tolerance Questionnaire; ITC: item-to-total correlation; ns: not significant; PANAS: Positive Affect Negative Affect Scales; PANAS-NA: PANAS Negative Affect; PANAS-PA: PANAS Positive Affect; SHAPS: Snaith–Hamilton Pleasure Scale; VAS: visual analog scale.

factor 2 [F2] representing an anticipation of relief from negative affect and nicotine withdrawal, with an urgent desire to smoke [11 items positively keyed]). However, the French and Brazilian translations show a difference in the order of factor extraction: Craving (urgent desire to smoke) was extracted first. According to the French authors, the duration of smoking abstinence of the French sample at the time of evaluation (1.5 to 3 hours) might explain the inversion of order of the two factors. Two-thirds of the subjects in the study of Tiffany & Drobes (1991) had been abstinent for 1 hour or less, which might have resulted in lesser craving in the US sample. The Spanish authors (Cepeda-Benito *et al.*, 2004) compared the factorial structures of the original US and translated Spanish versions and found that: (1) a better fit was found with the four-factor and two-factor models than with the one-factor model, and (2) the two-factor model provided a better fit than the four-factor model in both samples. In addition, their data suggested that the presence of mostly negatively worded items in F1 contributed largely to the two-factor structure of the QSU. Analysis with only negative items in F1 greatly improved the model fit in both data sets. According to the authors, these findings question the original interpretation of the nature of the dimensions measured by the two factors of the QSU.

- QSU-b: The authors of the Dutch and Malay versions reported differences from the original QSU-b (Cox *et al.*, 2001), which, when used to derive a global measure of craving, showed high internal consistency across settings and provided reliable assessment of the desire to smoke. In contrast, factor analyses generated two instances of verbal report of craving. F1 represented a strong desire and intention to smoke, with smoking perceived as satisfying for active smokers, when an anticipation of relief from negative affect and an urgent desire to smoke was reflected by F2.

The first factor (F1) of the Dutch version (Littel *et al.*, 2011) corresponded with the second factor (F2) of the English QSU-b (items 2, 4, 5, 8, and 9). F2 comprised items 1, 3, 6, 7, and 10. Items 2 and 5 loaded strongly on F1, whereas they had originally cross-loaded. The authors attributed this discrepancy to language differences. Items 2 and 5 (i.e., “nothing would be better than smoking a cigarette right now” and “all I want right now is a cigarette”) communicate quite extreme statements, especially when literally translated into Dutch. F2 corresponded with the first factor of the original QSU-b, although, in the Dutch study, items 1 and 6 loaded on two factors. Again, Dutch language might be an explanation for these items loading on both factors. Items 1 and 6 include the words “desire” and “urge.” Although phrases such as “I have a strong desire or urge for a cigarette,” might be used in Dutch, it is far more common to use less potent expressions (e.g., “I would like/fancy a cigarette”). Nevertheless, items 1 and 6 are less extreme than the items assigned to F1. The authors did not add “anticipation of pleasure from smoking” to the name of this factor, because the subscale was not significantly correlated with either positive or negative affect.

In the Malay version (Blebil *et al.*, 2015), factors 1 and 2 corresponded with those in the original version, with items 2 and 5 strongly loading on F2. The authors attributed this cross loading to the phrase “strong urge” conveying extreme utterances when literally translated into Malay.

Internal consistency was explored for all translations of the QSU. The alpha values for QSU F1 and F2 ranged from 0.89 (Guillin *et al.*, 2000) to 0.96 (Araujo *et al.*, 2006) and 0.87 (Cepeda-Benito *et al.*, 2004) to 0.92 (Araujo *et al.*, 2006), respectively, in line with the values of the original version, in which scores representing these two factors demonstrated strong internal consistency (Cronbach’s alpha = 0.95 and 0.93, respectively). The Cronbach’s alpha values of the QSU-b translations in Malay (Blebil *et al.*, 2015), Dutch (Littel *et al.*, 2011), and Chinese (Yu *et al.*, 2010) were 0.81, 0.83, and 0.92, respectively. When scored as a 10-item scale, the original QSU-b demonstrated high reliability as a measure of global craving in both initial and follow-up sessions (alpha = 0.89 and 0.87, respectively).

Test–retest reliability was never assessed.

Correlations with biomarkers and consumption patterns were explored for all translations (QSU/QSU-b) except the German (Müller *et al.*, 2001) and Spanish (Cepeda-Benito *et al.*, 2004) QSU versions, with the latter focusing only on factor analysis.

- QSU: As in the original development, correlations with biomarkers were not explored in the translations. The correlation with number of cigarettes per day was weak to moderate [Brazilian version: F1 and F2 QSU, $r = 0.197$ and 0.182 , respectively (Araujo *et al.*, 2006); French QSU, $r = 0.54$ (Guillin *et al.*, 2000) (not explored in the original)].
- QSU-b: Correlation with exhaled CO was explored for the Malay version ($r = 0.0024$; not explored in the original development) (Blebil *et al.*, 2015). The correlation with number of cigarettes per day was weak [Dutch QSU-b, $r = 0.25$ (Littel *et al.*, 2011); Malay QSU-b, $r = 0.30$ (Blebil *et al.*, 2015) (not explored in the original)].

Correlations with self-reported measures exploring similar or dissimilar constructs were not studied in the French (Guillin *et al.*, 2000) or Spanish versions of the QSU (Cepeda-Benito *et al.*, 2004) but were investigated in other translations (QSU/QSU-b):

- QSU: Correlations with the Craving Visual Analog Scale (VAS) German (Müller *et al.*, 2001) and Brazilian (Araujo *et al.*, 2006) versions, Fagerström Tolerance Questionnaire (FTQ) (Müller *et al.*, 2001), FTND (Araujo *et al.*, 2006), BAI, and BDI (Araujo *et al.*, 2006) were explored. Weak correlations were found with the FTQ (F2, $r = 0.16$), FTND (F1, $r = 0.244$; F2, $r = 0.163$), BAI (F1, $r = 0.249$), and BDI (F1 $r = 0.249$).

For the original QSU, correlations with the Withdrawal Symptoms Checklist (WSC) and Mood Form were assessed. F1 showed a strong correlation with the craving subscale of the WSC.

- QSU-b: Correlations with the FTND Malay and Dutch versions (Blebil *et al.*, 2015; Littel *et al.*, 2011), craving scales Dutch and Chinese versions (Littel *et al.*, 2011; Yu *et al.*, 2010), desire and urge VAS, the Positive Affect Negative Affect Scale (PANAS), and the Snaith–Hamilton Pleasure Scale (SHAPS) Dutch version (Littel *et al.*, 2011) were explored. Weak correlations were found with the FTND (Malay: $r = 0.24$; Dutch: $r = 0.14$), and strong correlations were found with the craving scales (Chinese: $r = 0.75$; Dutch: $r = 0.80$). Only a weak correlation with the Negative Affect scale of the PANAS was found for F1 of the Dutch version ($r = 0.25$). Weak correlations were found with the SHAPS (F1 and F2 of the Dutch version: $r = 0.23$ and 0.22 , respectively). Correlations with the Mood Form were assessed in the original version.

Responsiveness to change, predictive validity, sensitivity, and specificity were not assessed.

MNWS results

Four studies were retrieved for the MNWS (Table 8), corresponding to three translations of the MNWS (nine-item version) into Chinese (China) (Yu *et al.*, 2010), Korean (Kim *et al.*, 2007), and Malay (Blebil *et al.*, 2014) and one translation of the MNWS-R and MNWS (eight-item version) into Italian (same paper) (Svicher *et al.*, 2017).

In the MNWS, there were variations in the items included in the original versions that were used as a basis for translation. The Chinese version was based on the MNWS developed by Cappelleri *et al.* (2005), while the Italian version was based on that by Hughes (1992). The Malay and Korean versions included “impatience” and were based on the MNWS developed by Jorenby *et al.* (1996) (Table 9).

Table 10 presents the general characteristics of each study. Combustible cigarettes were the TNP evaluated in all studies.

Table 8. List of Minnesota Nicotine Withdrawal Scale (MNWS)/MNWS-revised (MNWS-R) translations and corresponding studies.

Measure	Language	Country	Study	Study objectives
MNWS 8-item version and MNWS-R	Italian	Italy	Svicher <i>et al.</i> , 2017	To perform factor analysis and explore the psychometric properties of the Italian version of the MNWS and MNWS-R
MNWS 9-item version	Chinese	China	Yu <i>et al.</i> , 2010	To evaluate the reliability and validity of the Chinese versions of the MNWS and QSU-b in Chinese smokers
	Korean	USA	Kim <i>et al.</i> , 2007	To develop and assess the psychometric properties of a Korean version of the MNWS for Korean Americans
	Malay	Malaysia	Blebil <i>et al.</i> , 2014	To evaluate the psychometric properties of the Malaysian version of the MNWS

Table 9. Items included in the Minnesota Nicotine Withdrawal Scale (MNWS) (original and translations).

Items	Original (Hughes, 1992)	Original (Hughes & Hatsukami, 1998)	Original (Cappelleri <i>et al.</i> , 2005)	Chinese (Yu <i>et al.</i> , 2010)	Italian (Svicher <i>et al.</i> , 2017)	Korean* (Kim <i>et al.</i> , 2007)	Malay* (Blebil <i>et al.</i> , 2014)
Craving	✓		✓	✓	✓	✓	
Depression	✓	✓	✓	✓	✓	✓	✓
Irritability/frustration/anger	✓	✓	✓	✓	✓	✓	✓
Anxiety	✓	✓	✓	✓	✓	✓	✓
Difficulty concentrating	✓	✓	✓	✓	✓	✓	✓
Restlessness	✓	✓	✓	✓	✓	✓	✓
Increased appetite/weight gain	✓	✓	✓	✓	✓	✓	✓
Insomnia	✓	✓			✓		
Difficulty going to sleep			✓	✓		✓	✓
Difficulty staying asleep			✓	✓			
Impatience						✓	✓

* Based on symptoms listed in Jorenby *et al.*, 1996.

Table 10. Sociodemographic/design characteristics and targeted country/language of studies evaluating the measurement properties of the Minnesota Nicotine Withdrawal Scale (MNWS)/MNWS-revised (MNWS-R) translations.

Measure	Authors	Study design	Language	Country	Subject characteristics Daily cigarette smokers				
					Sample details	Number of participants	Sex M/F (%)	Age in years (mean $\pm$ SD/range)	Number of cigarettes per day (mean $\pm$ SD)
MNWS and MNWS-R	MNWS 9-item version <i>Yu et al., 2010</i>	Cross-sectional	Chinese	China	NS	355	98/2	39.6 $\pm$ 11.6/18–65	17.8 $\pm$ 7.7 (based on consumption 1 month before quitting)
	MNWS-R and 8-item version <i>Svicher et al., 2017</i>	Longitudinal	Italian	Italy	General population	366 (133 for test-retest)	40.7/59.3	34.00 $\pm$ 11.29	13.0 $\pm$ 7.0 (in supplementary materials)
	MNWS 9-item version <i>Kim et al., 2007</i>	Cross-sectional	Korean	USA	Immigrants	118 (93 for test-retest)	100/0	42.11 $\pm$ 10.42/20–63	5.0 (SD NS) (current smokers, asked to rate themselves based on the time not smoking)
	MNWS 9-item version <i>Blebil et al., 2014</i>	Cross-sectional	Malay	Malaysia	Smokers who attended the Quit Smoking Clinic at Penang General Hospital	133 (75 for test-retest)	99.2/0.8	47.7 $\pm$ 14.0/18–76	14.92 $\pm$ 9.10 (more than 77% of sample had not attempted quitting previously)

Abbreviations: F: female; M: male; MNWS: Minnesota Nicotine Withdrawal Scale; MNWS-R: Minnesota Nicotine Withdrawal Scale revised version; NS: not specified in this paper; SD: standard deviation.

Most subjects reported previous attempts to quit, except in the Malay sample (Blebil *et al.*, 2014), where 77% of the subjects had not attempted to quit previously.

All studies were run with moderate smoker samples, on average, except for the study involving Koreans living in the US, which had recruited light smokers (Kim *et al.*, 2007). In comparison, the original MNWS was developed with heavy smokers (Hughes & Hatsukami, 1986). Mean participant age ranged from 34 to 47.7 years. Men were predominant in all studies except in that in Italy, where women were slightly preponderant (59%) (Svicher *et al.*, 2017).

Translation process. All four papers provided a description of the translation process used to develop each translation (Table 11). Only the Korean version (Kim *et al.*, 2007) presented a brief report of the difficulties encountered and solutions found. References to guidelines or recommendations were given for all translations except for the Italian version (Svicher *et al.*, 2017). Descriptions of the translation process were detailed for all translations except the Chinese version (Yu *et al.*, 2010).

Measurement properties. Table 12 reports the measurement properties explored for each translation. All translations were assessed for structural validity, with a one-factor structure reported for the Italian MNWS eight-item version (Svicher *et al.*, 2017) and the Malay nine-item version (Blebil *et al.*, 2014).

A two-factor structure was reported for the Chinese version of the MNWS nine-item version (Yu *et al.*, 2010) and the Korean nine-item version (Kim *et al.*, 2007). The structure of the Chinese version was identical to the two-factor structure of the original version reported by Cappelleri *et al.* (2005): negative effect (F1, four items: depressed mood; irritability, frustration, or anger; anxiety; and difficulty concentrating), insomnia (F2, two items: difficulty going to sleep and difficulty staying asleep), and three single items (craving, restlessness, and increased appetite). A review of the items showed a slight discrepancy in those used as originals, with impatience listed in the Korean version but not in the Chinese, where insomnia represents two items (difficulty going to sleep and difficulty staying asleep). For the Korean version, F1 represented early-occurring disorders in mental functioning, and F2 represented disorders in physiological functioning and late-occurring disorders in mental functioning (i.e., increased appetite, disturbed sleep, depression, and impatience), explaining 66% of the variance. A two-factor structure was also reported in the Italian MNWS-R.

Internal consistency was explored for all translations, with Cronbach's alpha values of 0.85 (Italian MNWS eight-item version) and 0.91 (Malay nine-item version) reported for the monofactorial structure. In comparison, the Cronbach's alpha values of the original eight-item MNWS explored by Toll *et al.* (2007) were 0.80 (abstinence study), 0.83 (framing study), and 0.82 (naltrexone + patch study) at the

initial time point after quitting. Internal consistency was not evaluated by Jorenby *et al.* (1996). With regard to the bifactorial structure, only a global alpha value was provided for the Chinese version (0.90). Cappelleri *et al.* (2005) reported alpha values ranging from 0.76 to 0.87 for the negative domain and 0.71 to 0.83 for the insomnia domain in the studies and time assessed.

Measurement equivalence using DIF was never assessed.

In total, 75% percent of the translations (3 of 4; Italian, Korean, and Malay) were assessed for test-retest reliability, with coefficients ranging from 0.51 (Kim *et al.*, 2007) to 0.88 (Blebil *et al.*, 2014) and intervals from 1 (Blebil *et al.*, 2014; Kim *et al.*, 2007) to 3 (Svicher *et al.*, 2017) months.

Correlations with consumption patterns and biomarkers were reported for three translations: Chinese, Italian, and Malay. Correlations with self-reported measures of dependence, craving, and anxiety were explored for all translations (Table 12).

Responsiveness to change, predictive validity, sensitivity, and specificity were not assessed.

Discussion

Given the globalization of tobacco research and control, we expected to retrieve more than 25, 9, 4, and 1 translations—documented with measurement properties—of the FTND, QSU/QSU-b, MNWS, and MNWS-R, respectively. A search on the Patient-Reported Outcome Quality Of Life Instrument Database (PROQOLID™; <https://eprovide.mapi-trust.org/>) reveals that there are 19 translations available for the QSU-b and 12 for the MNWS-R. No information was retrieved for the FTND, as it is not listed or documented on PROQOLID™.

- Among the 19 QSU-b translations listed on PROQOLID™ and indicated as translated by Mapi, we found two versions overlapping with our research (i.e., the Dutch and Spanish versions for Spain), indicating that there are at least two versions of the QSU-b in those languages. PROQOLID™ does not mention whether those 19 translations have undergone any evaluation of their measurement properties. Our review found three translations of the QSU-b (Chinese, Dutch, and Malay) and one Spanish version derived from the Spanish QSU, with content different from the original QSU-b.
- A review of the information on the Vermont University website reveals the existence of seven translations of the MNWS (nine-item versions: Chinese, Czech, Dutch, Japanese, and Korean; 11-item version: Arabic; 14-item version: Portuguese) and five of the MNWS-R (Bosnian, German, Italian, Russian, and Spanish), all listed under the acronym of MNWS. A comparison of the information on PROQOLID™ and the Vermont University website reveals no overlap for the MNWS-R translations, raising the number to 17 translations available. Our review found only one MNWS-R translation (Italian) with documented measurement properties.

Table 11. Description of translation processes used (steps and people involved if mentioned) for the Minnesota Nicotine Withdrawal Scale (MNWS)/MNWS-revised (MNWS-R) translations.

Language/country study	Conceptual evaluation	FT step (people involved in the process and number of FTs)	Consensus on the FT step to produce one FT (people involved in the process)	BT step (people involved in the process and number of BTs)	Reconciliation/consensus (people involved in the process)	Pilot/test (people involved in the process)	Report (i.e., description of issues, changes made for cultural, semantic, or syntactic reasons)	Reference (if any)
Chinese/China <i>Yu et al., 2010</i>	✓ (called preparation)	✓ (NS)	✓ (NS)	✓ (NS)	✓ (NS)	✓ (23 subjects)	No	<i>Wild et al., 2005</i>
Italian/Italy <i>Svicher et al., 2017</i>	No	✓ (two independent English lecturers; 2 FT)	✓ (NS)	✓ (another bilingual expert; 1 BT)	✓ (NS)	✓ (10 healthy volunteers)	No	No
Korean/USA <i>Kim et al., 2007</i>	?*	✓ (1st and 3rd authors of paper, native Korean speakers; 2 FT)	✓ (same as that for FT)	✓ (two research assistants, native English speakers; 2 BT) – 2 rounds of BT	✓ (expert review by 10 Korean American professionals in behavioral health)	No	Item 2: the Korean word <i>shyn-kyung-jil-juck-im</i> used for “irritability” was changed to <i>zta-seung-nam</i> after the panel review. Most of the members stated that the first Korean word was too strong to translate “irritability” because it is often used to describe a person with emotional instability. Item 5: “restlessness” was translated as <i>an-jul-boo-jul-mot-haam</i> , which was then translated back to agitation or feeling unstable.	<i>Flaherty et al., 1988</i>
Malay/Malaysia** <i>Blebil et al., 2014</i>	?*	✓ (two University lecturers, native Malaysian speakers; 2 FT)	✓ (two authors of the manuscript)	✓ (one translator, native Malay, proficient in English; 1 BT)	✓ (NS)	✓ (20 smokers)	No	<i>Guillemin et al., 1993</i> <i>Wild et al., 2005</i>

Abbreviations: BT: backward translation; FT: forward translation; NS: not specified.

*?: Not clearly specified in the paper, but recommended in quoted guidelines; **Translation Process referred to as Linguistic Validation Process.

Table 12. Measurement properties of the translations of the Minnesota Nicotine Withdrawal Scale (MNWS)/MNWS-revised (MNWS-R).

Measure language / country of study	Translation process described in paper	Reliability		Reliability – Test–retest Correlation coeff + interval - Inter-rater Correlation coeff + number of raters	Validity		Correlations with biomarkers or consumption patterns	Correlations of total score/factors score with other measures assessing similar or dissimilar constructs
		Internal consistency Cronbach's alpha			Structural validity Factor analyses (EFA, CFA) ITC IRT			
MNWS 9-item version Chinese / China <i>Yu et al., 2010</i>	Yes	$\alpha = 0.9$ (alpha for each factor not provided)			Two factors + 3 individual items (CTT): same as in <i>Cappellen et al., 2005</i> . F1: Negative effect (depressed mood; irritability, frustration, or anger; anxiety; difficulty concentrating) F2: insomnia (difficulty going to sleep; difficulty staying asleep) Single items: Craving, restlessness, and increased appetite ITC: $r = 0.54$ to 0.85		Patient-evaluated discomfort scores: $r = 0.68$	
MNWS 8-item version & MNWS-R Italian / Italy <i>Svicher et al., 2017</i>	Yes	MNWS $\alpha = 0.85$ MNWS-R $\alpha = 0.87$ (FI: 0.86, FI: 0.64)	MNWS Test-retest: $r = 0.59$ 3 months MNWS-R Test-retest: total score: $r = 0.60$ F1: $r = 0.59$; F2: $r = 0.45$ 3 months		MNWS: one factor ITC: $r > 0.30$ MNWS-R: two factors F1: Psychological symptoms (10 items) F2: Somatic features (5 items) ITC $r > 0.30$ (except for sore throat $r = 0.27$)	MNWS SCS: $r = 0.69$ SCS craving: $r = 0.34$ SCS abstinence: $r = 0.61$ FTCD: $r = 0.29$ ASI-3: $r = 0.53$ ASI-3 physical concerns: $r = 0.41$ ASI-3 mental concerns: $r = 0.44$ ASI-3 social concerns: $r = 0.42$ PANAS negative affect: $r = 0.54$ AUDIT: $r = 0.29$ MNWS-R SCS: $r = 0.56$ (F1: 0.62; F2: 0.22) SCS craving: $r = 0.31$ (F1: 0.32; F2: 0.18) SCS abstinence: $r = 0.55$ (F1: 0.62; F2: 0.20) FTCD: $r = 0.25$ (F1: 0.25; F2: 0.17) ASI-3: $r = 0.56$ (F1: 0.55; F2: 0.40) ASI-3 physical concerns: $r = 0.47$ (F1: 0.44; F2: 0.38) ASI-3 mental concerns: $r = 0.45$ ASI-3 social concerns: $r = 0.38$ (F1: 0.35; F2: 0.31) PANAS negative affect: $r = 0.57$ (F1: 0.57; F2: 0.27) AUDIT: $r = 0.29$ (F1: 0.28; F2: 0.22)		

Measure language / country of study	Reliability		Validity		Correlations with biomarkers or consumption patterns	Correlations of total score/factors score with other measures assessing similar or dissimilar constructs
	Translation process described in paper	Internal consistency Cronbach's alpha	Reliability - Test-retest interval - Inter-rater Correlation coeff + number of raters	Structural validity Factor analyses (EFA, CFA) ITC IRT		
MNWS 9-item version Korean / USA Kim <i>et al.</i> , 2007	Yes	$\alpha = 0.88$ (F1: 0.88, F2: 0.79)	Test-retest: ICC = 0.51 (95% CI: 0.70–0.73) 1 month	Two factors F1: Craving; irritability, frustration, or anger; anxiety; difficulty concentrating; restlessness F2: Increased appetite; disturbed sleep; depression; impatience ITC: $r = 0.39$ to 0.74	Attempts to quit in the past year: total scale $r = 0.25$, F1: $r = 0.21$, F2: $r = 0.26$	SERS: total scale $r = -0.23$, F1: $r = -0.20$, F2: $r = -0.22$
MNWS 9-item version Malay / Malaysia Blebil <i>et al.</i> , 2014	Yes	$\alpha = 0.91$	Test-retest: $r = 0.876$ 1 month	One factor ITC: 0.54 to 0.79	CO level: $r = 0.72$ No. of cig./day: $r = 0.68$ Self-rated chances to quit: $r = -0.38$ Duration of smoking: coeff. not shown Previous quit attempts: coeff. not shown	FTND-M total score: $r = 0.68$

Abbreviations: ASI-3: Anxiety Sensitivity Index-3; AUDIT: Alcohol Use Disorder Identification Test; CI: confidence interval; cig.: cigarettes; CO: carbon monoxide; Coeff.: coefficient; CTT: classical test theory; FTCD: Fagerström Test for Cigarette Dependence; F1: factor 1; F2: factor 2; FTND-M: Malay version of the Fagerström Test for Nicotine Dependence; ICC: intraclass coefficient; IRT: item-response theory; ITC: item-to-total correlation; ns: not significant; PANAS: Positive and Negative Affect Schedule; SCS: Smoker Complaint Scale; SERS: Self-Efficacy in Resisting Smoking Scale

Overall, our review showed that the process used to elaborate the translations of the FTND, QSU/QSU-b, and MNWS/MNWS-R is not standardized and is not always documented. This could prove to be a challenge if the US FDA CTP aligns any future guidance with the 2009 patient-reported outcome (PRO) guidance published by the US FDA Center for Drug Evaluation and Research (CDER) (US Department of Health and Human Services, 2009). Appendix VIII of this PRO guidance outlines that all translation documents should be provided for CDER review. This includes a report on the process(es) used and challenges encountered during the translation process, especially during testing of the translation on the target population.

There is a great heterogeneity in the populations recruited for each study, in terms of sample characteristics (e.g., sex [samples with mixed sexes or a majority of male subjects], age, or level of cigarette consumption [light to heavy smokers]). In addition, depending on the objectives of the research teams, not all properties are explored for each language. Our review showed that most of the translations have measurement properties similar to their original versions.

Results concerning the MNWS might be found to be problematic as the number of items used are different across the languages (e.g., eight for Italian vs. nine for Chinese, Korean, and Malay), and there is variation in which items are included, making it impossible to compare scores across studies. Translations of the FTND revealed the same concerns about the structural validity of the original version (mono vs. bifactorial), low internal consistency [except for some versions (de Meneses-Gaya *et al.*, 2009; Osório Fde *et al.*, 2013)], and validity of Q3 (hated the most to give up). Those translation measurement outcomes questioning the validity of the original instrument may raise questions about the need to modify the content of the original. In this context, there is a well-known precedent: The International Quality of Life Assessment Project (Aronson *et al.*, 1992) is a notorious example of the development of translations of a PRO measure (i.e., the Short Form-36 [SF-36] Health Survey, which led the developer to change the original US instrument). The development and validation of the translated versions contributed to improvements in item wording and response categories and to the creation of the SF-36v2 Health Survey (Ware, 2007).

Our review showed that cross-cultural validity is rarely explored. Measurement equivalence using an IRT-based approach for examining DIF is almost never applied. This is a concern, as it might make it difficult to know if the scores obtained with the translations of these measures are comparable across languages and cultures and whether or not it is relevant to aggregate data from studies conducted in different countries. Based on their extensive experience in cross-cultural evaluation, researchers from the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Group have suggested that DIF should be part of the validation of questionnaire translations (Petersen *et al.*, 2003; Scott *et al.*, 2006; Scott *et al.*, 2009). In their research, DIF analyses were

conducted to identify items answered differently by language administration, reflecting either linguistic issues (e.g., imperfect translation) or cultural differences. Overall, they showed that, although most of the EORTC QLQ-C30 items seemed to have good linguistic equivalence, several scales presented highly conflicting results for some translations. They implied that some of these effects might be substantial enough to affect the outcomes of clinical studies, as translation differences in an item could result in clinically important differences at the scale score level.

Finally, our review showed that none of the translations has been validated with candidate MRTPs, indicating that more research is needed to comply with regulatory recommendations on the development of self-reported measures for use in labeling claims (US Department of Health and Human Services, 2009).

The main limitation of our research lies in its descriptive design. We did not provide insights on the quality of the translated versions (i.e., ratings on the translation process and the quality of the measurement properties) (Schellingerhout *et al.*, 2011; Thoomes-de Graaf *et al.*, 2016). Further research is needed to critically appraise the quality of the translations and guide researchers in their search for the best translation for their studies.

These results showing (1) discrepancies between the number of translations available, with and without documented information about their measurement properties, (2) heterogeneity in the scope of measurement properties explored and in the characteristics of the samples recruited, and (3) lack of validation with TNPs other than conventional cigarettes raise the need for generating a new initiative with two main goals (i.e., information and development).

First, implementation of a centralized repository for measurement instruments (original version and translations) with a licensing structure (endorsed by the developers of the originals) would enable researchers to have access to the most up-to-date information about measures (i.e., development story and psychometric properties). By identifying existing translations and documenting them, this implementation might also help prevent the development of multiple translations for the same language and avoid concerns about which translation to use (Anfray *et al.*, 2009). Furthermore, engaging the developers of the original versions in this process might help protect the integrity of each measurement instrument included (Anfray *et al.*, 2018).

Second, if the original versions and translations of these measures are not appropriate for candidate MRTPs, fit-for-purpose measurement instruments (i.e., concept-driven instruments providing interpretable outcomes for the intended purpose) should be developed to enable comparison of combustible and noncombustible products on the same risk continuum. A similar initiative was launched several years ago, which led to the development of the ABOUT™ Toolbox (Assessment of

Behavioral Outcomes related to Tobacco and nicotine products) (Chrea *et al.*, 2018). The measurement instruments included in this Toolbox are at different degrees of development. With their dissemination on ePROVIDE™, researchers will be able to use instruments that are (1) developed and validated with state-of-the-art scientific methods to be psychometrically sound, straightforward to implement in clinical and population-based studies, and easy to interpret; (2) created to be relevant and applicable across the whole spectrum of TNPs and across various populations; and (3) designed to enhance standardization and comparison of data on perception and behaviors toward MRTPs across academic, industry, and public health research communities.

Data availability

Underlying data

All data underlying the results are available as part of the article and no additional source data are required.

Extended data

Open Science Framework. Measurement properties of the translations of instruments evaluating the subjective effects of tobacco- and nicotine-containing products: a systematic review of the literature. <https://doi.org/10.17605/OSF.IO/3Z2EV> (Acquadro, 2019).

This project contains the following extended data:

- Supplementary file 1: List of the 193 references retrieved during the literature search.
- Supplementary file 2: Tables presenting detailed information on the FTND translations.

- Table S2.1. Sociodemographic/design characteristics and targeted country/language of studies evaluating the measurement properties of the translations.
- Table S2.2. Description of translation processes used (steps and people involved if mentioned) for the FTND translations.
- Table S2.3. Measurement properties of the translations of the FTND.

Reporting guidelines

Open Science Framework: PRISMA checklist for: “Measurement properties of the translations of instruments evaluating the subjective effects of tobacco- and nicotine-containing products: a systematic review of the literature” <https://doi.org/10.17605/OSF.IO/3Z2EV> (Acquadro, 2019).

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

Author contributions

CA, CC and NM conceived the idea of the manuscript. CDG performed the literature search. CA and CDG reviewed the abstracts and selected the articles for review. CA developed and drafted the manuscript. CC, CDG, MH, NM, and RW reviewed critically the manuscript. All authors read and approved the final version of the manuscript.

References

- Aaronson NK, Acquadro C, Alonso J, *et al.*: **International Quality of Life Assessment (IQOLA) Project.** *Qual Life Res.* 1992; 1(5): 349–351.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Acquadro C: **Measurement properties of the translations of instruments evaluating the subjective effects of tobacco- and nicotine-containing products: a systematic review of the literature.** 2019.
<http://www.doi.org/10.17605/OSF.IO/3Z2EV>
- Anfray C, Arnold B, Martin M, *et al.*: **Reflection paper on copyright, patient-reported outcome instruments and their translations.** *Health Qual Life Outcomes.* 2018; 16(1): 224.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Anfray C, Emery MP, Conway K, *et al.*: **The advantages of a centralized dissemination strategy for health outcomes instruments and their translations: A case example with the Zarit Burden Interview (ZBI).** *Value Health.* 2009; 12: A399.
[Publisher Full Text](#)
- Araujo RB, Oliveira Mda S, Mansur MA: **Brazilian validation of the Questionnaire of Smoking Urges.** *Cad Saude Publica.* 2006; 22(10): 2157–2167.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Becoña E, López A, Del Rio EF, *et al.*: **Spanish adaptation of the NDSS (Nicotine Dependence Syndrome Scale) and assessment of nicotine-dependent individuals at primary care health centers in Spain.** *Span J Psychol.* 2010; 13(2): 951–960.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Becoña E, Vázquez F: **The Fagerström Test for Nicotine Dependence in a Spanish sample.** *Psychol Rep.* 1998; 83(3 Pt 2): 1455–1458.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bleibil AQ, S Sulaiman SA, A Hassali M, *et al.*: **Evaluation of the Psychometric Properties of the Malay Version of the Minnesota Nicotine Withdrawal Scale.** *Value Health Reg Issues.* 2014; 3: 19–23.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bleibil AQ, Sulaiman SA, Hassali MA, *et al.*: **The Malay version of the brief questionnaire on smoking urge: translation and psychometric properties of the questionnaire.** *J Epidemiol Glob Health.* 2015; 5(1): 15–22.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Breteler MH, Hilberink SR, Zeeman G, *et al.*: **Compulsive smoking: the development of a Rasch homogeneous scale of nicotine dependence.** *Addict Behav.* 2004; 29(1): 199–205.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Buckley TC, Mozley SL, Holohan DR, *et al.*: **A psychometric evaluation of the Fagerström Test for Nicotine Dependence in PTSD smokers.** *Addict Behav.* 2005; 30(5): 1029–1033.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Burling AS, Burling TA: **A comparison of self-report measures of nicotine dependence among male drug/alcohol-dependent cigarette smokers.** *Nicotine Tob Res.* 2003; 5(5): 625–633.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Cappelleri JC, Bushmakina AG, Baker CL, *et al.*: **Revealing the multidimensional framework of the Minnesota nicotine withdrawal scale.** *Curr Med Res Opin.* 2005; 21(5): 749–760.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Cappelleri JC, Bushmakina AG, Baker CL, *et al.*: **Confirmatory factor analyses and reliability of the modified cigarette evaluation questionnaire.** *Addict Behav.* 2007; 32(5): 912–923.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Carmo JT, Pueyo AA: **Adaptação do português do Fagerström test for nicotine**

dependence (FTND) para avaliar a dependência e tolerância à nicotina em fumantes brasileiros. *Rev Bras Med.* 2002; 59(1/2): 73–80.

Reference Source

Carter LP, Stitzer ML, Henningfield JE, *et al.*: **Abuse liability assessment of tobacco products including potential reduced exposure products.** *Cancer Epidemiol Biomarkers Prev.* 2009; 18(12): 3241–3262.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Cepeda-Benito A, Henry KC, Gleaves DH, *et al.*: **Cross-cultural investigation of the Questionnaire of Smoking Urges in American and Spanish smokers.** *Assessment.* 2004; 11(2): 152–159.

[PubMed Abstract](#) | [Publisher Full Text](#)

Cepeda-Benito A, Reig-Ferrer A: **Development of a brief questionnaire of smoking urges--Spanish.** *Psychol Assess.* 2004; 16(4): 402–407.

[PubMed Abstract](#) | [Publisher Full Text](#)

Chabrol H, Niezborala M, Chastan E, *et al.*: **A study of the psychometric properties of the Fagerstrom Test for Nicotine Dependence.** *Addict Behav.* 2003; 28(8): 1441–1445.

[PubMed Abstract](#) | [Publisher Full Text](#)

Chrea C, Acquadro C, Afolalu EF, *et al.*: **Developing fit-for-purpose self-report instruments for assessing consumer responses to tobacco and nicotine products: the ABOUT™ Toolbox initiative [version 1; peer review: 2 approved].** *F1000Res.* 2018; 7: 1878.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Ciconelli RM: **Tradução para o português e validação do questionário genérico de avaliação de qualidade de vida “medical outcomes study 36-item short-form health survey (SF-36)” [Tese de Doutorado].** São Paulo, Brazil: Escola Paulista de Medicina, Universidade Federal de São Paulo. 1997.

Reference Source

Cox LS, Tiffany ST, Christen AG: **Evaluation of the brief questionnaire of cigarette urges (QSU-brief) in laboratory and clinical settings.** *Nicotine Tob Res.* 2001; 3(1): 7–16.

[PubMed Abstract](#) | [Publisher Full Text](#)

Davies GM, Willner P, Morgan MJ: **Smoking-related cues elicit craving in tobacco “chippers”: a replication and validation of the two-factor structure of the Questionnaire of Smoking Urges.** *Psychopharmacology (Berl).* 2000; 152(3): 334–342.

[PubMed Abstract](#) | [Publisher Full Text](#)

de Leon J, Diaz FJ, Becoña E, *et al.*: **Exploring brief measures of nicotine dependence for epidemiological surveys.** *Addict Behav.* 2003; 28(8): 1481–1486.

[PubMed Abstract](#) | [Publisher Full Text](#)

de Menezes-Gaya C, Zuairi AW, de Azevedo Marques JM, *et al.*: **Psychometric qualities of the Brazilian versions of the Fagerström Test for Nicotine Dependence and the Heaviness of Smoking Index.** *Nicotine Tob Res.* 2009; 11(10): 1160–1165.

[PubMed Abstract](#) | [Publisher Full Text](#)

Dethier V, Heeren A, Galanti L, *et al.*: **Probing smoking craving with a multidimensional approach: validation of the 12-item French-language version of the Questionnaire on Smoking Urges.** *Neuropsychiatr Dis Treat.* 2014; 10: 1459–1468.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Etter JF: **A comparison of the content-, construct- and predictive validity of the cigarette dependence scale and the Fagerström test for nicotine dependence.** *Drug Alcohol Depend.* 2005; 77(3): 259–268.

[PubMed Abstract](#) | [Publisher Full Text](#)

Etter JF, Duc TV, Perneger TV: **Validity of the Fagerström test for nicotine dependence and of the Heaviness of Smoking Index among relatively light smokers.** *Addiction.* 1999; 94(2): 269–281.

[PubMed Abstract](#) | [Publisher Full Text](#)

Etter JF, Hughes JR: **A comparison of the psychometric properties of three cigarette withdrawal scales.** *Addiction.* 2006; 101(3): 362–372.

[PubMed Abstract](#) | [Publisher Full Text](#)

Etter JF, Perneger TV, Ronchi A: **Distributions of smokers by stage: international comparison and association with smoking prevalence.** *Prev Med.* 1997; 26(4): 580–585.

[PubMed Abstract](#) | [Publisher Full Text](#)

Fagerström KO: **Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment.** *Addict Behav.* 1978; 3(3-4): 235–241.

[PubMed Abstract](#) | [Publisher Full Text](#)

Fagerström K: **Determinants of tobacco use and renaming the FTND to the Fagerstrom Test for Cigarette Dependence.** *Nicotine Tob Res.* 2012; 14(1): 75–78.

[PubMed Abstract](#) | [Publisher Full Text](#)

Fagerström KO, Russ C, Yu CR, *et al.*: **The Fagerström Test for Nicotine Dependence as a predictor of smoking abstinence: a pooled analysis of varenicline clinical trial data.** *Nicotine Tob Res.* 2012; 14(12): 1467–1473.

[PubMed Abstract](#) | [Publisher Full Text](#)

Ferketich AK, Fossati R, Apolone G: **An evaluation of the Italian version of the Fagerström Test for Nicotine Dependence.** *Psychol Rep.* 2008; 102(3): 687–694.

[PubMed Abstract](#) | [Publisher Full Text](#)

Flaherty JA, Gaviria FM, Pathak D, *et al.*: **Developing instruments for cross-cultural psychiatric research.** *J Nerv Ment Dis.* 1988; 176(5): 257–263.

[PubMed Abstract](#) | [Publisher Full Text](#)

Grassi MC, Enea D, Ferketich AK, *et al.*: **Validity of the Italian version of the Severity of Dependence Scale (SDS) for nicotine dependence in smokers intending to quit.** *Psychol Rep.* 2014; 114(1): 1–13.

[PubMed Abstract](#) | [Publisher Full Text](#)

Guillemin F, Bombardier C, Beaton D: **Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines.** *J Clin Epidemiol.* 1993; 46(12): 1417–1432.

[PubMed Abstract](#) | [Publisher Full Text](#)

Guillin O, Krebs MO, Bourdel MC, *et al.*: **[Validation of the French translation and factorial structure of the Tiffany and Drobes Smoking Urge Questionnaire].** *Encephale.* 2000; 26(6): 27–31.

[PubMed Abstract](#)

Haddock CK, Lando H, Klesges RC, *et al.*: **A study of the psychometric and predictive properties of the Fagerström Test for Nicotine Dependence in a population of young smokers.** *Nicotine Tob Res.* 1999; 1(1): 59–66.

[PubMed Abstract](#) | [Publisher Full Text](#)

Hanson K, O'Connor R, Hatsukami D: **Measures for assessing subjective effects of potential reduced-exposure products.** *Cancer Epidemiol Biomarkers Prev.* 2009; 18(12): 3209–3224.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Hays RD, Hadorn D: **Responsiveness to change: an aspect of validity, not a separate dimension.** *Qual Life Res.* 1992; 1(1): 73–75.

[PubMed Abstract](#) | [Publisher Full Text](#)

Herdman M, Fox-Rushby J, Badia X: **‘Equivalence’ and the translation and adaptation of health-related quality of life questionnaires.** *Qual Life Res.* 1997; 6(3): 237–247.

[PubMed Abstract](#) | [Publisher Full Text](#)

Heatherton TF, Kozlowski LT, Frecker RC, *et al.*: **The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire.** *Br J Addict.* 1991; 86(9): 1119–1127.

[PubMed Abstract](#) | [Publisher Full Text](#)

Huang CL, Cheng CP, Lin HH, *et al.*: **Psychometric testing of the Chinese version of the Hooked on Nicotine Checklist in adolescents.** *J Adolesc Health.* 2009; 45(3): 281–285.

[PubMed Abstract](#) | [Publisher Full Text](#)

Huang CL, Lin HH, Wang HH: **The psychometric properties of the Chinese version of the Fagerstrom Test for Nicotine Dependence.** *Addict Behav.* 2006; 31(12): 2324–7.

[PubMed Abstract](#) | [Publisher Full Text](#)

Hudmon KS, Pomerleau CS, Brigham J, *et al.*: **Validity of retrospective assessments of nicotine dependence: a preliminary report.** *Addict Behav.* 2005; 30(3): 613–7.

[PubMed Abstract](#) | [Publisher Full Text](#)

Hughes JR: **Tobacco withdrawal in self-quitters.** *J Consult Clin Psychol.* 1992; 60(5): 689–97.

[PubMed Abstract](#) | [Publisher Full Text](#)

Hughes JR: **Background on the Minnesota Tobacco Withdrawal Scale-Revised (MTWS-R).** 2017;

Reference Source

Hughes JR, Hatsukami D: **Signs and symptoms of tobacco withdrawal.** *Arch Gen Psychiatry.* 1986; 43(3): 289–94.

[PubMed Abstract](#) | [Publisher Full Text](#)

Hughes JR, Hatsukami DK: **Errors in using tobacco withdrawal scale.** *Tob Control.* 1998; 7(1): 92–93.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Hughes JR, Oliveto AH, Riggs R, *et al.*: **Concordance of different measures of nicotine dependence: two pilot studies.** *Addict Behav.* 2004; 29(8): 1527–1539.

[PubMed Abstract](#) | [Publisher Full Text](#)

Institute of Medicine: **Scientific Standards for Studies on Modified Risk Tobacco Products.** Washington: The National Academies Press. 2012;

[Publisher Full Text](#)

Jayakrishnan R, Mathew A, Lekshmi K, *et al.*: **Assessment of nicotine dependence among smokers in a selected rural population in Kerala, India.** *Asian Pac J Cancer Prev.* 2012; 13(6): 2663–2667.

[PubMed Abstract](#) | [Publisher Full Text](#)

Jhanjee S, Sethi H: **The Fagerström test for nicotine dependence in an Indian sample of daily smokers with poly drug use.** *Nicotine Tob Res.* 2010; 12(11): 1162–1166.

[PubMed Abstract](#) | [Publisher Full Text](#)

John U, Meyer C, Schumann A, *et al.*: **A short form of the Fagerström Test for Nicotine Dependence and the Heaviness of Smoking Index in two adult population samples.** *Addict Behav.* 2004; 29(6): 1207–1212.

[PubMed Abstract](#) | [Publisher Full Text](#)

Jorenby DE, Hatsukami DK, Smith SS, *et al.*: **Characterization of tobacco withdrawal symptoms: transdermal nicotine reduces hunger and weight gain.** *Psychopharmacology (Berl).* 1996; 128(2): 130–138.

[PubMed Abstract](#) | [Publisher Full Text](#)

Kassim S, Salam M, Croucher R: **Validity and reliability of the Fagerstrom Test for Cigarette Dependence in a sample of Arabic speaking UK-resident Yemeni khat chewers.** *Asian Pac J Cancer Prev.* 2012; 13(4): 1285–1288.

[PubMed Abstract](#) | [Publisher Full Text](#)

- Kawada T, Hirata K, Inagaki H, *et al.*: **Significance of the 100-point scale to evaluate perceived tobacco dependence.** *Work*. 2010; **35**(2): 183–189.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kim SS, Gulick EE, Kim SH, *et al.*: **Psychometric properties of the Minnesota Nicotine Withdrawal Scale: a Korean version.** *J Nurs Meas*. 2007; **15**(2): 121–32.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Klinsophon T, Janwantanakul P, Thaveeratitham P: **Reliability of the Thai version of the Fagerstrom Test for Nicotine Dependence (FTND).** *J Med Assoc Thai*. 2017; **100**(10): 1130.
[Reference Source](#)
- Kozlowski LT, Pillitteri JL, Sweeney CT, *et al.*: **Asking questions about urges or cravings for cigarettes.** *Psychol Addict Behav*. 1996; **10**(4): 248–260.
[Publisher Full Text](#)
- Kozlowski LT, Porter CQ, Orleans CT, *et al.*: **Predicting smoking cessation with self-reported measures of nicotine dependence: FTQ, FTND, and HSI.** *Drug Alcohol Depend*. 1994; **34**(3): 211–6.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Littel M, Franken IHA, Muris P: **Psychometric properties of the brief Questionnaire on Smoking Urges (QSU-Brief) in a Dutch smoker population.** *Neth J Psychol*. 2011; **66**.
[Reference Source](#)
- Meneses-Gaya IC, Zuairi AW, Loureiro SR, *et al.*: **Psychometric properties of the Fagerström Test for Nicotine Dependence.** *J Bras Pneumol*. 2009; **35**(1): 73–82.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Mikami I, Akechi T, Kugaya A, *et al.*: **Screening for nicotine dependence among smoking-related cancer patients.** *Jpn J Cancer Res*. 1999; **90**(10): 1071–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mokkink LB, Prinsen CAC, Patrick DL, *et al.*: **COSMIN methodology for systematic reviews of Patient-Reported Outcomes (PROMs).** – User manual Version 1.0 dated February 2018. 2018.
[Reference Source](#)
- Mokkink LB, Terwee CB, Patrick DL, *et al.*: **The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes.** *J Clin Epidemiol*. 2010a; **63**(7): 737–45.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Mokkink LB, Terwee CB, Patrick DL, *et al.*: **The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study.** *Qual Life Res*. 2010b; **19**(4): 539–49.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mokkink LB, Terwee CB, Knol DL, *et al.*: **The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content.** *BMC Med Res Methodol*. 2010c; **10**: 22.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Moreno-Coutiño A, Villalobos-Gallegos L: **Psychometric properties of the Fagerström Test for Nicotine Dependence in a Sample of Mexican Smokers.** *J Addict Nurs*. 2017; **28**(1): 27–33.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Müller V, Mucha RF, Ackermann K, *et al.*: **The assessment of craving in smokers with a German version of the “Questionnaire on smoking urges” (QSU-G) (In German).** *Zeitschrift für klinische Psychologie, Psychiatrie und Psychotherapie*. 2001; **30**(3): 164–171.
[Publisher Full Text](#)
- Nakajima M, al’Absi M, Dokam A, *et al.*: **An examination of the Fagerström Test for Nicotine Dependence among concurrent tobacco and khat users.** *J Psychoactive Drugs*. 2012; **44**(5): 437–441.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ng M, Freeman MK, Fleming TD, *et al.*: **Smoking prevalence and cigarette consumption in 187 countries, 1980–2012.** *JAMA*. 2014; **311**(2): 183–192.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Okuyemi KS, Pulvers KM, Cox LS, *et al.*: **Nicotine dependence among African American light smokers: a comparison of three scales.** *Addict Behav*. 2007; **32**(10): 1989–2002.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Osório Fde L, de Carvalho AC, Crippa JA, *et al.*: **Screening for smoking in a general hospital: scale validation, indicators of prevalence, and comorbidity.** *Perspect Psychiatr Care*. 2013; **49**(1): 5–12.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Park SM, Son KY, Lee YJ, *et al.*: **A preliminary investigation of early smoking initiation and nicotine dependence in Korean adults.** *Drug Alcohol Depend*. 2004; **74**(2): 197–203.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pasquall L: **Princípios de elaboração de escalas psicológicas.** *Rev Psiq Clin*. 1998; **25**: 206–213.
[Reference Source](#)
- Payne TJ, Smith PO, McCracken LM, *et al.*: **Assessing nicotine dependence: a comparison of the Fagerström Tolerance Questionnaire (FTQ) with the Fagerström Test for Nicotine Dependence (FTND) in a clinical sample.** *Addict Behav*. 1994; **19**(3): 307–317.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Petersen MA, Groenvold M, Bjorner JB, *et al.*: **Use of differential item functioning analysis to assess the equivalence of translations of a questionnaire.** *Qual Life Res*. 2003; **12**(4): 373–385.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pomerleau CS, Carton SM, Lutzke ML, *et al.*: **Reliability of the Fagerstrom Tolerance Questionnaire and the Fagerstrom Test for Nicotine Dependence.** *Addict Behav*. 1994; **19**(1): 33–39.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Radzins A, Gallo JJ, Epstein DH, *et al.*: **A factor analysis of the Fagerström Test for Nicotine Dependence (FTND).** *Nicotine Tob Res*. 2003; **5**(2): 255–240.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Regnault A, Herdman M: **Using quantitative methods within the Universalist model framework to explore the cross-cultural equivalence of patient-reported outcome instruments.** *Qual Life Res*. 2015; **24**(1): 115–124.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Reubi D, Berridge V: **The Internationalisation of Tobacco Control, 1950–2010.** *Med Hist*. 2016; **60**(4): 453–472.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Robabeh S, Mohammad JM, Mahnaz F, *et al.*: **Psychometric properties of the Persian version of the Fagerström test for nicotine dependence in patients with opioid use disorder / cigarette smokers under methadone maintenance treatment.** *Neuroquantology*. 2017; **15**: 253–260.
[Publisher Full Text](#)
- Rose JE, Behm FM, Westman EC: **Nicotine-mecamylamine treatment for smoking cessation: The role of pre-cessation therapy.** *Exp Clin Psychopharmacol*. 1998; **6**(3): 331–343.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Salameh P, Jomaa L, Farhat G, *et al.*: **The Young Adults’ Cigarette Dependence (YACD) score: an improved tool for cigarette dependence assessment in university students.** *Addict Behav*. 2013; **38**(5): 2174–2179.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Salameh P, Jomaa L, Issa C, *et al.*: **Assessment of health risk behaviours among university students: a cross-sectional study in Lebanon.** *International Journal of Adolescent Youth*. 2014; **19**: 203–216.
[Publisher Full Text](#)
- Salameh P, Khayat G, Waked M: **The Lebanese Cigarette Dependence (LCD) Score: a comprehensive tool for cigarette dependence assessment.** *Int J Behav Med*. 2014; **21**(2): 385–393.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Sarbandi F, Niknami S, Hidarnia A, *et al.*: **Psychometric properties of the Iranian version of the Fagerström Test for Nicotine Dependence and of Heaviness of Smoking Index.** *J Res Health*. 2015; **5**: 96–103.
[Reference Source](#)
- Schellingerhout JM, Heymans MW, Verhagen AP, *et al.*: **Measurement properties of translated versions of neck-specific questionnaires: a systematic review.** *BMC Med Res Methodol*. 2011; **11**: 87.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Scott NW, Fayes PM, Aaronson NK, *et al.*: **The practical impact of differential item functioning analyses in a health-related quality of life instrument.** *Qual Life Res*. 2009; **18**(18): 1125–1130.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Scott NW, Fayes PM, Bottomley A, *et al.*: **Comparing translations of the EORTC QLQ-C30 using differential item functioning analyses.** *Qual Life Res*. 2006; **15**(6): 1103–1115.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Sledjeski EM, Dierker LC, Costello D, *et al.*: **Predictive validity of four nicotine dependence measures in a college sample.** *Drug Alcohol Depend*. 2007; **87**(1): 10–19.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Stavem K, Røgeberg OJ, Olsen JA, *et al.*: **Properties of the Cigarette Dependence Scale and the Fagerström Test of Nicotine Dependence in a representative sample of smokers in Norway.** *Addiction*. 2008; **103**(9): 1441–1449.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Steinberg ML, Williams JM, Steinberg HR, *et al.*: **Applicability of the Fagerström Test for Nicotine Dependence in smokers with schizophrenia.** *Addict Behav*. 2005; **30**(1): 49–59.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Svicher A, Beghè A, Mangiaracina G, *et al.*: **Factor Analysis and Psychometric Properties of the Minnesota Nicotine Withdrawal Scale and the Minnesota Nicotine Withdrawal Scale-Revised: Italian Version.** *Eur Addict Res*. 2017; **23**(3): 157–162.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Svicher A, Cosci F, Giannini M, *et al.*: **Item Response Theory analysis of Fagerström Test for Cigarette Dependence.** *Addict Behav*. 2018; **77**: 38–46.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Thoomes-de Graaf M, Scholten-Peeters GG, Schellingerhout JM, *et al.*: **Evaluation of measurement properties of self-administered PROMs aimed at patients with non-specific shoulder pain and “activity limitations”: a systematic review.** *Qual Life Res*. 2016; **25**(9): 2141–2160.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Tiffany ST, Drobes DJ: **The development and initial validation of a questionnaire on smoking urges.** *Br J Addict*. 1991; **86**(11): 1467–1476.
[PubMed Abstract](#) | [Publisher Full Text](#)

Toll BA, Katulak NA, McKee SA: **Investigating the factor structure of the Questionnaire on Smoking Urges-Brief (QSU-Brief)**. *Addict Behav.* 2006; **31**(7): 1231–1239.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Toll BA, McKee SA, Krishnan-Sarin S, *et al.*: **Revisiting the factor structure of the questionnaire on smoking urges**. *Psychol Assess.* 2004; **16**(4): 391–395.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Toll BA, O'Malley SS, McKee SA, *et al.*: **Confirmatory factor analysis of the Minnesota Nicotine Withdrawal Scale**. *Psychol Addict Behav.* 2007; **21**(2): 216–225.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

U.S. Congress: **Family Smoking Prevention and Tobacco Control Act of 2009, Public Law 111-31, 123 Stat. 1776 (June 22, 2009)**. 2009.

[Reference Source](#)

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Tobacco Products: **Guidance for industry - Modified risk tobacco product applications - Draft Guidance**. 2012.

[Reference Source](#)

U.S. Department of Health and Human Services, Food and Drug Administration: **Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims**. 2009. Federal Register, 74, 65132–65133.

[Reference Source](#)

Uysal MA, Kadakal F, Karşıdağ C, *et al.*: **Fagerstrom test for nicotine dependence: reliability in a Turkish sample and factor analysis**. *Tüberk Toraks.* 2004; **52**(2): 115–121.

[PubMed Abstract](#)

Uysal MA, Öztuna D, Bahadır A, *et al.*: **Psychometric properties of the Turkish version of the Fagerström test for nicotine dependence**. *Tüberk Toraks.* 2015; **63**(4): 250–256.

[PubMed Abstract](#) | [Publisher Full Text](#)

Vink JM, Willemsen G, Beem AL, *et al.*: **The Fagerström Test for Nicotine Dependence in a Dutch sample of daily smokers and ex-smokers**. *Addict Behav.* 2005; **30**(3): 575–579.

[PubMed Abstract](#) | [Publisher Full Text](#)

Ware J: **User's Manual for the SF-36v2 Health Survey**. Second ed. Lincoln, RI QualityMetric Incorporated. 2007.

Weinberger AH, Reutenauer EL, Allen TM, *et al.*: **Reliability of the Fagerström Test for Nicotine Dependence, Minnesota Nicotine Withdrawal Scale, and Tiffany Questionnaire for Smoking Urges in smokers with and without schizophrenia**. *Drug Alcohol Depend.* 2007; **86**(2–3): 278–282.

[PubMed Abstract](#) | [Publisher Full Text](#)

West R, Ussher M: **Is the ten-item Questionnaire of Smoking Urges (QSU-brief) more sensitive to abstinence than shorter craving measures?**

Psychopharmacology (Berl). 2010; **208**(3): 427–432.

[PubMed Abstract](#) | [Publisher Full Text](#)

West R, Ussher M, Evans M, *et al.*: **Assessing DSM-IV nicotine withdrawal symptoms: A comparison and evaluation of five different scales**.

Psychopharmacology (Berl). 2006; **184**(3–4): 619–627.

[PubMed Abstract](#) | [Publisher Full Text](#)

Westman EC, Levin ED, Rose JE: **Smoking while wearing the nicotine patch: Is smoking satisfying or harmful?** *Clinical Research.* 1992; **40**: 871A.

[Reference Source](#)

Wild D, Grove A, Martin M, *et al.*: **Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation**. *Value Health.* 2005; **8**(2): 94–104.

[PubMed Abstract](#) | [Publisher Full Text](#)

Yamada H, Acton GS, Tsoh JY: **Differential item functioning of the English and Chinese versions of the Fagerstrom Test for Nicotine Dependence**. *Addict Behav.* 2009; **34**(2): 125–133.

[PubMed Abstract](#) | [Publisher Full Text](#)

Yee A, Ng CG, Rusdi AR: **Validation of the Malay version of Fagerström Test for Nicotine Dependence (FTND-M) among a group of male staffs in a university hospital**. *Malaysian Journal of Psychiatry.* 2011; **20**: 1.

[Reference Source](#)

Yu X, Xiao D, Li B, *et al.*: **Evaluation of the Chinese versions of the Minnesota nicotine withdrawal scale and the questionnaire on smoking urges-brief**. *Nicotine Tob Res.* 2010; **12**(6): 630–634.

[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 30 May 2022

<https://doi.org/10.5256/f1000research.22644.r136674>

© 2022 Galanti M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Maria Rosaria Galanti

Department of Global Public Health, Centre for Epidemiology and Community Medicine (CES), Karolinska Institutet, Stockholm, Sweden

The authors did a considerable effort in tracing and summarizing scientific articles related to the use of instruments for the assessment of nicotine dependence in languages different from the one in which they were originally created.

Despite this effort, I find the review quite difficult to follow because it is lacking a definite focus, ultimately resulting in point-by-point description of what was or was not done in different studies. Indeed, the discussion of the results leads to rather general statements, some of which could be anticipated: not standardized processes translation/adaptation; heterogeneity of populations assessed; deficit of analysis of differential item functioning (DIF); lack of cross-cultural validation.

I suggest that the authors re-organize and summarize their valuable material along one (or both) of the following themes:

1. Compare the procedures and the analyses done for the several language-specific versions of the scales with pre-defined gold-standards (as described in the methods section), for instance how should a sound translation be conducted and reported; how should content, construct validity, predictive validity, etc. be minimally investigated. As a result of this summary the reader will be informed on how many of the x-language versions of the y-instrument meet these criteria.
2. Compare the psychometric properties of the different language-specific translations of a given instrument with the original one(s) analyzing the possible reasons for discrepancies, above all the characteristics of the involved samples. This work would be of crucial importance, since smoking populations are increasingly diversifying across different countries, with the emergence of both low-frequency smokers and heavy smokers with psychiatric co-morbidity. Suggested comparisons: population vs. clinical samples; heavy smokers vs light/moderate smokers.

Apart from these overarching considerations I have two additional points:

1. Title: would be better to rephrase it so that it would be clear that “subjective effects” refer to symptoms of dependence (even dizziness after smoking is a subjective effect).
2. I wonder why the Cigarette Dependence Scale (Etter *et al.*, 2003¹) was not included in the search/results. I am aware of at least one study (Rydell *et al.*, 2016²) dealing with both translation of the original instrument and its adaptation to smokeless tobacco use.

References

1. Etter JF, Le Houezec J, Perneger TV: A self-administered questionnaire to measure dependence on cigarettes: the cigarette dependence scale. *Neuropsychopharmacology*. 2003; **28** (2): 359-70
[PubMed Abstract](#) | [Publisher Full Text](#)
2. Rydell M, Sundin E, Ramstedt M, Galanti MR: Measuring tobacco dependence in the general population: Evaluation of the Cigarette Dependence Scale (CDS-12) and its adaptation to smokeless tobacco use (STDS-12) in two Swedish surveys. *Scand J Public Health*. 2016; **44** (2): 134-42
[PubMed Abstract](#) | [Publisher Full Text](#)

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Partly

Is the statistical analysis and its interpretation appropriate?

Not applicable

Are the conclusions drawn adequately supported by the results presented in the review?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Population and clinical studies of tobacco use and cessation

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 23 March 2021

<https://doi.org/10.5256/f1000research.22644.r77538>

© 2021 Rose J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Jennifer Rose

Quantitative Analysis Center, Wesleyan University, Middletown, CT, USA

This manuscript reports on a comprehensive review of the psychometric properties of the translated versions of some frequently used nicotine dependence (ND) assessments. The range of psychometric properties used to test equivalence of the translated ND measures was as rigorous as possible, given that there was considerable variability in the translation process and reporting of psychometric properties across the studies.

The review found a surprising lack of rigorous assessment of measurement invariance (differential item functioning) for translated versions compared to the original English language versions. This is mentioned in the Discussion, but it seems important to more strongly emphasize that the psychometric properties examined in this study can evaluate measurement equivalence, but not measurement invariance. Although factor structure, reliability, and validity appeared similar for most of the translations, it is possible that the meaning of the ND measure was different for the translated versions compared to the English version. This potential lack of measurement invariance might be a reason for the psychometric differences that were found. Without testing for measurement invariance, the possibility of unaccounted noninvariance remains a potential confounder in subsequent statistical analyses using the translated ND measures, and may be an important factor in frequent failure to replicate findings across studies using the same measures.

In the Introduction, the authors indicate that it is crucial to ensure “cross-cultural equivalence between the original versions and their translations”. This suggests that observed translation differences in the ND measures are related to cultural differences. However, the reasons for the differences are indeterminate. The differences may be attributed to cultural differences, but also they could be due to changes in the wording in the translated measures, study design, or a myriad number of other factors. This should be clarified.

The drawback of such a comprehensive review is that it is very difficult to keep track of all the psychometric differences, and there is no clear way to determine when the identified differences may be considered significant or meaningful and when they may be insignificant or inconsequential. In addition, the supplemental tables provide more information, but they were not easy to locate. Perhaps there could be more detailed instructions for accessing those tables.

Are the rationale for, and objectives of, the Systematic Review clearly stated?

Yes

Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

Is the statistical analysis and its interpretation appropriate?

Not applicable

Are the conclusions drawn adequately supported by the results presented in the review?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Nicotine dependence, statistics, data analysis, psychometric assessment.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research