Title

Development of the PROMIS nicotine dependence item banks.

Journal

Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco, 16 Suppl 3(Suppl 3)

ISSN

1462-2203

Authors

Shadel, William G
Edelen, Maria Orlando
Tucker, Joan S
et al.

Publication Date

2014-09-01

DOI

10.1093/ntr/ntu032

Peer reviewed

ORIGINAL INVESTIGATION

Development of the PROMIS® Nicotine Dependence Item Banks

William G. Shadel PhD[1], **Maria Orlando Edelen** PhD[2], **Joan S. Tucker** PhD[3], **Brian D. Stucky** PhD[3], **Mark Hansen** MPH[4], **Li Cai** PhD[4]

[1]*RAND Health, RAND Corporation, Pittsburgh, PA;* [2]*RAND Health, RAND Corporation, Boston, MA;* [3]*RAND Health, RAND Corporation, Santa Monica, CA;* [4]*CSE/CRESST, Graduate School of Education and Information Studies, University of California, Los Angeles, CA*

Corresponding Author: William G. Shadel, PhD, RAND Health, RAND Corporation, 4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213, USA. Telephone: 412-683-2300, ext. 4489; Fax: 412-683-2800; E-mail: shadel@rand.org

Received July 8, 2013; accepted February 12, 2014

ABSTRACT

Introduction: Nicotine dependence is a core construct important for understanding cigarette smoking and smoking cessation behavior. This article describes analyses conducted to develop and evaluate item banks for assessing nicotine dependence among daily and nondaily smokers.

Methods: Using data from a sample of daily ($N = 4,201$) and nondaily ($N = 1,183$) smokers, we conducted a series of item factor analyses, item response theory analyses, and differential item functioning analyses (according to gender, age, and race/ethnicity) to arrive at a unidimensional set of nicotine dependence items for daily and nondaily smokers. We also evaluated performance of short forms (SFs) and computer adaptive tests (CATs) to efficiently assess dependence.

Results: A total of 32 items were included in the Nicotine Dependence item banks; 22 items are common across daily and nondaily smokers, 5 are unique to daily smokers, and 5 are unique to nondaily smokers. For both daily and nondaily smokers, the Nicotine Dependence item banks are strongly unidimensional, highly reliable (reliability = 0.97 and 0.97, respectively), and perform similarly across gender, age, and race/ethnicity groups. SFs common to daily and nondaily smokers consist of 8 and 4 items (reliability = 0.91 and 0.81, respectively). Results from simulated CATs showed that dependence can be assessed with very good precision for most respondents using fewer than 6 items adaptively selected from the item banks.

Conclusions: Nicotine dependence on cigarettes can be assessed on the basis of these item banks via one of the SFs, by using CATs, or through a tailored set of items selected for a specific research purpose.

INTRODUCTION

Nicotine dependence is considered one of the most important constructs for understanding cigarette smoking behavior and smoking cessation. There is no debate that the nicotine in cigarettes is dependence producing (Stolerman & Jarvis, 1995). Dependence is a construct with neurobiological underpinnings with a core set of cognitive and behavioral expressions or features (Rosenthal, Weitzman, & Benowitz, 2011). Features that characterize nicotine dependence (on cigarettes) include craving, withdrawal that occurs upon cessation of smoking, compulsive use, and tolerance, which are features that correspond somewhat with the more formal diagnostic criteria (American Psychiatric Association, 2000). Nicotine dependence is viewed as a primary motivator of smoking behavior in adolescents (Colby, Tiffany, Shiffman, & Niaura, 2000) and adults (Shadel, Shiffman, Niaura, Nichter, & Abrams, 2000) and is cited as a prime reason that smokers of all ages have difficulty quitting (Hughes, 2011). For cigarette smokers, the central features of nicotine dependence (e.g., withdrawal, craving, tolerance) are the targets of many behavioral and pharmacological treatment approaches (Fiore et al., 2008). For example, treatment with the nicotine patch is designed to reduce the withdrawal symptoms most smokers experience upon quitting, thereby making cessation more attainable (Shiffman, Ferguson, Gwaltney, Balabanas, & Shadel, 2006).

Considering its central clinical importance, reliably and validly measuring nicotine dependence has been a core goal for cigarette smoking research for nearly 40 years. Indeed, there are at least eight assessments of nicotine dependence that have appeared in this time. However, significant issues make choosing from among these assessments a less than straightforward task. The earliest and most widely used self-report instrument developed to measure nicotine dependence in cigarette smokers, the Fagerström Tolerance Questionnaire (FTQ; Fagerström, 1978), and even its modifications (see Heatherton, Kozlowski, Frecker, & Fagerström, 1991), have been criticized for lacking adequate reliability (Pomerleau, Carlton, Lutzke,

doi:10.1093/ntr/ntu032

Flessland, & Pomerleau, 1994), for lacking construct validity because they do not represent the diversity of features that characterize the nicotine dependence syndrome (e.g., the FTQ lacks items tapping into compulsive use and craving) (Shiffman, Waters, & Hickcox, 2004), and for not reflecting the diagnostic criteria for nicotine dependence (Etter, Le Houezec, & Perneger, 2003). Other assessments, like the Nicotine Dependence Syndrome Scale (NDSS; Shiffman et al., 2004) and the Wisconsin Inventory of Smoking Dependence Motives (WISDM; Piper et al., 2004), were systematically developed as psychometrically stronger alternatives to the FTQ and were also designed to capture the diversity of features that characterize the construct of nicotine dependence for cigarette smokers. However, despite theoretical rigor and psychometric strengths, a common limitation of both the NDSS and WISDM is that they are relatively long assessments. Thus, widespread adoption of these scales may not be practical, particularly in clinical contexts or research settings where the time that is needed to complete an assessment battery is limited. Briefer assessments exist (e.g., Heaviness of Smoking Index; Heatherton, Kozlowski, Frecker, Rickert, & Robinson, 1989), but these assessments lack the content coverage and breadth of the longer assessments. Although there is a dearth of studies comparing different measures of nicotine dependence for cigarette smokers, findings from extant comparative studies have yielded mixed results with some measures better at predicting some outcomes compared with other measures (Piper, McCarthy, & Baker, 2006). As such, deciding how to assess nicotine dependence is complicated by the sheer number of measures available, and the fact that there is almost no empirically grounded guidance available as to which measure(s) one should choose. Thus, the decision on which measure of nicotine dependence to use is likely determined by the time allocated for assessment as much as for conceptual reasons (Niaura & Shadel, 2003).

The primary goal of the PROMIS® Smoking Initiative is to develop psychometrically sound item banks to measure nicotine dependence for cigarette smokers as well as other biopsychosocial constructs associated with smoking (see other articles in this supplement) that can be effectively and efficiently administered in a variety of contexts. In the initial phase of the PROMIS Smoking Initiative, items from existing measures of cigarette nicotine dependence, as well as items from measures of other aspects of smoking (e.g., craving, nicotine effects, positive and negative expectancies of smoking), were subjected to rigorous qualitative review and exploratory and confirmatory factor analyses (see Edelen, Tucker, Shadel, Stucky, & Cai, 2012, for details of this approach). Results from those analyses identified a subset of items with content representative of the nicotine dependence construct (e.g., craving, withdrawal, and behavioral priority; see Shadel et al., 2000). The current article describes a series of analyses that were conducted on this set of nicotine dependence items to arrive at a unidimensional "bank" of items that can serve as the basis for reliable assessment of cigarette nicotine dependence, functioning in the same way for smokers of either gender and across various racial/ethnic and age groups. Because nicotine dependence may express itself differently in daily and in intermittent or nondaily cigarette smokers (Shiffman, Ferguson, Dunbar, & Scholl, 2012; Shiffman & Paty, 2006), these analyses were conducted with data from daily smokers and nondaily smokers separately and included a final cocalibration of all items across daily and nondaily smokers to link the scales of the two item sets. We also evaluated the performance of short forms (SFs) and computer adaptive tests (CATs) in order to efficiently, yet reliably, assess cigarette nicotine dependence. Our analysis plan follows closely the many procedures described by Reeve et al. (2007) in their psychometric evaluation and calibration of health-related quality of life item banks for PROMIS. A more complete description of the analytic process used to develop the daily and nondaily smoker item banks for the PROMIS Smoking Initiative is presented by Hansen et al. in this supplement.

METHODS

Sample and Procedure

A national sample of smokers ($N_{(total)}$ = 5,384; $N_{(daily)}$ = 4,201; $N_{(nondaily)}$ = 1,183) was recruited by Harris Interactive through their online panel membership, and all assessments were completed via the Internet. All procedures were IRB approved. Individuals were eligible if they were 18 years or older, had been smoking for at least a year, had smoked in the past 30 days, and did not have plans to quit in the next 30 days. Based on their response to number of days smoked in past 30 days, those participants indicating smoking 28–30 of the past 30 days were classified as daily smokers; respondents smoking less than 28 of the past 30 days were classified as nondaily smokers. Sample recruitment was targeted to reflect the demographic composition of U.S. adult smokers in terms of gender, race/ethnicity, and age. The survey was fielded between July and September 2011 via a randomized block design (Reeve et al., 2007). The block design was constructed to minimize respondent burden while maximizing the interitem covariance coverage. To cross-validate the dimensionality of the Nicotine Dependence item bank, the daily smoker sample was randomly split into exploratory ($N_{(exploratory)}$ = 3,021) and confirmatory ($N_{(confirmatory)}$ = 1,180) subsamples.

Demographic information and the characteristics of individual smoking behaviors for the daily and nondaily samples are contained in Table 1. As anticipated, differences between the exploratory and confirmatory daily smoker subsamples were minor and varied by chance; thus characteristics of the two daily smoker subsamples are combined in the table. The nondaily and daily samples differed significantly on all demographic characteristics. Compared with the daily smokers, nondaily smokers were more likely to be male, less likely to be Caucasian, were slightly younger, tended to be more educated, were more likely to be employed and never married, and tended to have higher incomes. As expected, daily and nondaily smokers also differed with respect to their smoking behaviors. Most notably, the nondaily smokers had not smoked for as long as the daily smokers and reported more lifetime quit attempts.

Measures

Smoking Items

A total of 277 unique smoking items were administered. These items were developed according to PROMIS procedures from extant items in the literature as well as direct feedback from smokers. This process, described in more detail in Edelen et al. (2012), employed a rigorous qualitative approach that included systematic literature review, binning and winnowing of items,

Table 1. Characteristics of Daily and Nondaily Smokers

Characteristic	Daily smokers (N = 4,201)	Nondaily smokers (N = 1,183)	Difference test
Female, %	54.8	47.0	$p < .0001$
Race/ethnicity, %			$p < .0001$
Non-Hispanic White	72.2	55.2	
African American	12.1	15.5	
Hispanic	11.3	24.4	
Asian	1.8	1.0	
Other	2.6	3.9	
Age, mean (*SD*)	46.4 (11.6)	44.1 (11.9)	$p < .0001$
Education, %			$p < .0001$
<High school graduate	3.1	2.0	
High school graduate	16.5	14.1	
Some college	38.2	32.1	
AA degree	12.5	9.8	
BA/BS degree	17.5	21.9	
Graduate degree	12.3	20.2	
Employment, %			$p < .0001$
Full time	52.9	60.6	
Part time	12.2	14.4	
Unemployed/retired/student/homemaker	34.2	24.3	
Missing	0.8	0.8	
Marital status, %			$p < .001$
Never married	20.5	26.1	
Married or civil union	45.6	43.9	
Divorced/separated/widowed	21.8	18.7	
Living with partner	12.1	11.2	
Income, %			$p < .0001$
<$25,000	22.5	19.4	
$20,000–$49,999	32.3	27.1	
$50,000–$99,999	33.8	34.9	
$100,000+	11.5	18.7	
Years smoked, %			$p < .0001$
1–10 years	11.7	29.2	
More than 10 years	88.3	70.8	
Number of days smoked in past 30, %			$p < .0001$
1 or 2 days	0.0	15.8	
3–5 days	0.0	9.6	
6–9 days	0.0	9.6	
10–19 days	0.0	23.2	
20–27 days	0.0	41.9	
28–30 days	100.0	0.0	
Average number of cigarettes per day in past 30 days, %			$p < .0001$
<1 per day	0.2	13.0	
1–5	8.0	48.3	
6–10	22.0	22.3	
11–20	47.3	13.5	
20+	22.6	3.9	
Number of times quit for at least 24 hr, %			$p < .0001$
Never	18.0	14.7	
1 time	12.3	6.2	
2–3 times	30.7	19.1	
4–5 times	19.7	12.7	
6–9 times	7.4	7.8	
10 or more times	12.0	40.1	
Quitting contemplation, %			$p < .0001$
Not thinking about quitting	40.1	42.3	
Thinking about quitting, but no plans to quit	37.1	29.0	
Plans to quit in next 6 months	22.7	28.7	

item standardization, solicitation of feedback from smokers via focus groups and cognitive interviews, and final item revisions. All respondents completed 13 of the 277 smoking items which assessed their smoking behavior and quitting history. The remaining 264 items were candidate items that were being considered for inclusion in one of the smoking item banks. These items were distributed across 26 overlapping forms containing an average of 147 items (range = 134–158); each respondent was randomly assigned one of the 26 forms.

Other Measures

All respondents supplied basic demographic information and completed one of eight PROMIS health-related quality of life SF measures (alcohol consumption, anger, anxiety, depression, fatigue, physical functioning, sleep disturbance, and global health; Cella et al., 2007).

Item Factor Analyses

Previous analyses of the daily smoker exploratory subsample identified a set of 55 items to be considered for inclusion in the Nicotine Dependence item bank for daily smokers (Edelen et al., 2012). These items were drawn from established scales that included the Fagerstrom Tolerance Scale (Fagerstrom, 1978; and by extension, the Fagerstrom Test for Nicotine Dependence [Heatherton et al., 1991] and Heaviness of Smoking Index [Heatherton et al., 1989]); the NDSS (Shiffman et al., 2004); the WISDM(Piper et al., 2004); the Questionnaire of Smoking Urges (Tiffany & Drobes, 1991); and the Cigarette Dependence Scale (Etter et al., 2003). In most cases, the wording of items in the item bank was modified from its original source in order to conform to PROMIS item banking standards (Cella et al., 2007) or in other cases, item text was modified based on feedback from the cognitive interviews used to refine the items (see Edelen et al., 2012). Thus, the text of items in the Nicotine Dependence item bank may not correspond exactly to the text of items from the original scales; however, the concept represented by a PROMIS item links that item to an original scaled item. The same 55 items were also considered for nondaily smokers, along with 8 items deemed to have content relevant for nondaily smokers (e.g., My attitude about my smoking is that I can "take it or leave it" at any time).

Using the exploratory subsample of daily smokers (N = 3,021) and the full sample of nondaily smokers (N = 1,183), we examined the underlying factor structures of the 55- and 63-item sets with the software IRTPRO (Cai, du Toit, & Thissen, 2011). Local dependence (LD) diagnostic indices (Chen & Thissen, 1997) and high-dimensional exploratory item factor analyses (Cai, 2010) were used to identify clusters of related items or LD departures from unidimensionality. Item bifactor models (Cai, Yang, & Hansen, 2011; Gibbons & Hedeker, 1992) were then specified to account for these LD clusters.

Examining model results for each smoker type, study team members evaluated items within each specific factor in order to select subsets of items that would collectively be more unidimensional than the initial sets of 55 and 63 items. We considered each item's loading on the nicotine dependence factor, the percentage of common variance accounted for by the nicotine dependence factor (i.e., item explained common variance or I-ECV; Stucky, Thissen, & Edelen, 2013) and substantive content. Small numbers of items were selected from each item cluster (or specific factor). The two resulting item subsets for daily and nondaily smokers were selected to more closely conform to the unidimensional structure assumed in the final item response theory (IRT) models.

After selecting items for inclusion and removal in this way, the dimensionality of the two resultant item sets was reevaluated by testing the fit of a 1-factor model using the Mplus software (Muthén & Muthén, 1998–2010) with WLSMV estimation for categorical response items and standard model fit indices and criteria (root mean squared error of approximation [RMSEA] ≤ 0.08, Tucker–Lewis index [TLI] ≥ 0.95, comparative fit index [CFI] ≥ 0.95; Browne & Cudeck, 1993; Hu & Bentler, 1999). For daily smokers, model fit was assessed first in the exploratory subsample (N = 3,021) and then confirmed using the validation subsample (N = 1,180); the analysis for nondaily smokers used the full nondaily sample (N = 1,183).

Differential Item Functioning

After identifying and confirming two sufficiently unidimensional item sets to represent nicotine dependence, the item sets were further evaluated for differential item functioning (DIF). These evaluations were conducted using the full daily (N = 4,201) and nondaily (N = 1,183) smoker samples with IRTPRO (Cai et al., 2011). DIF indicates a response bias for a particular subgroup that is not accounted for by group-level mean and variance differences. DIF was evaluated for significance according to gender, race/ethnicity (White, Black, and Hispanic), and age (18–30, 31–50, and 51+) using established procedures (Edelen, Thissen, Teresi, Kleinman, & Ocepek-Welikson, 2006; Orlando & Marshall, 2002). Items with significant DIF were further evaluated for "impact" by considering the *w*eighted *a*rea *b*etween the expected score *c*urves (wABC) and the expected *d*ifference in *e*xpected *a p*osteriori score (dEAP), indices described in more detail in Hansen et al. Items with wABC values greater than 0.30 were screened for potential removal by evaluating graphical illustrations of the subgroups' expected scores curves, along with the values of the wABC and dEAP indices. Items judged to have nonignorable DIF were removed from further consideration in their respective item banks (i.e., daily or nondaily).

Calibration of Item Banks

The Nicotine Dependence item banks for daily and nondaily smokers were concurrently calibrated using data from the full combined sample ($N = 5{,}384$, $N_{(\text{daily})} = 4{,}201$, $N_{(\text{nondaily})} = 1{,}183$). We estimated a two-group IRT model with groups distinguishing daily and nondaily smokers. This calibration, which specified the daily smokers as the reference group, fixed the daily nicotine dependence mean to 0 and the standard deviation to 1, and estimated unique nondaily mean and standard deviation. Following PROMIS standards, IRT scores were subsequently rescaled using the T-score metric to have a mean of 50 and a standard deviation of 10 for daily smokers. The scale for the daily—nondaily group difference was set based on preidentified anchor items whose parameter estimates were constrained to be equal across the groups. Item parameters for nonanchor items were estimated separately for the two groups (see Hansen et al. for more details). The utility of the item banks was determined using IRT-based test information, score precision, and marginal reliability (MR).

Short Form Development

Item parameters from the final calibration were used in the development of nicotine dependence fixed-item SFs. In order to simplify the administration and scoring of these forms, only those items with equal parameters for daily and nondaily smokers (i.e., anchor items in the two-group calibration) were considered for SF inclusion. Among all the possible combinations of eligible items, candidate SFs were identified using selection criteria related to overall content balance, inclusion of items favored by the study team, and the reliability of score estimates across a broad range of nicotine dependence. Following PROMIS procedures, SF scoring was based on a transformation of the sum of responses to SF items. The use of summed scores has the particular advantage of allowing for the creation of translation tables by which researchers may convert an observed sum into an IRT-scaled score (Thissen, Nelson, Rosa, & McLeod, 2001). The performance of the SFs was evaluated using simulated data. For both the daily and nondaily item banks, we examined the reliability of each SF and obtained correlations of SF scores with scores based on the patterns of responses to the full sets of items.

CAT Simulation

Computerized adaptive tests utilize item selection algorithms to administer items that are tailored to the respondent's estimated standing on the measured construct, often resulting in reductions in test length and respondent burden. We conducted CAT simulations using Firestar (Choi, 2009) to evaluate the utility of computer adaptive administration of the daily and nondaily smoker Nicotine Dependence item banks. These simulations (a) provide an indication of the average number of items from the Nicotine Dependence item banks that would be administered under typical CAT conditions, (b) indicate which items would be most routinely selected for CAT administration, and (c) characterize the expected CAT-based score reliability.

RESULTS

Item Factor Analyses

Bifactor models, each with 11 specific factors, were selected to characterize the structure of both the 55 daily smoker items (using the exploratory daily smoker sample) and the 63 nondaily smoker items. In both cases, these models were selected based on their interpretability, comparisons of fit indices, and LD chi-squares. The specific factors identified in the bifactor model represent the content "clusters" in the nicotine dependence item sets. For illustrative purposes, Table 2 includes the general and specific factor loadings for items comprising the first three specific factors in the daily smoker solution (the full list of candidate items is available from the first author upon request). These three clusters represent content associated with needing to smoke upon waking in the morning, automatic smoking or smoking without being aware, and emotional distress at the thought of not being able to smoke.

The study team reviewed the bifactor model results for all 55 daily smoker and 63 nondaily smoker items and selected at least one item per specific factor to retain for further consideration in the item banks. Item selection was based on the strength of the general factor loading and item content. For the first three specific factors in Table 2, retained items are indicated with an asterisk. For some specific factor item clusters, the I-ECV indicated a strongly unidimensional item loading on the general factor. In these cases, additional items per specific factor were selected.

Table 2. Example Items and Loadings From the 55-Item Bifactor Model for Daily Smokers

		Specific factors			
Item stem	General	1	2	3	I-ECV
I smoke even when I am so ill that I am in bed most of the day.*	0.60	0.24			0.86
I smoke more frequently during the first hours after waking than during the rest of the day.	0.32	0.30			0.53
After I wake up I typically smoke my first cigarette of the day.	0.47	0.67			0.33
I find myself reaching for cigarettes without thinking about it.*	0.63		0.65		0.48
I light cigarettes without thinking about it.	0.56		0.71		0.39
I smoke automatically without being aware of it.	0.58		0.67		0.42
I have found a cigarette in my mouth and did not remember putting it there.	0.51		0.50		0.51
When I run out of cigarettes, I find it almost unbearable.*	0.83			0.20	0.94
The idea of not having any cigarettes causes me stress.*	0.81			0.35	0.84
The thought of never smoking again is overwhelming.*	0.66			0.40	0.73
I would go crazy if I couldn't smoke.	0.79			0.34	0.84

Note. For illustrative purposes, only the first three specific factors are included. The full model had 11 specific factors and 1 general nicotine dependence factor. Asterisks indicate items selected to form the preliminary Nicotine Dependence item bank for daily smokers.

This process led to the selection of 33 daily smoker items and 34 nondaily smoker items that balanced item content and closely represented the nicotine dependence dimension. Next, 1-factor models were fit to the selected item sets to confirm that they were sufficiently unidimensional. Relative to the original 55 daily smoker items (CFI = 0.907, TLI = 0.903, RMSEA = 0.060), the reduced set of 33 daily smoker items showed improved fit in both the exploratory and confirmatory subsamples (Exploratory: CFI = 0.977, TLI = 0.976, RMSEA = 0.043; Confirmatory: CFI = 0.982, TLI = 0.980, RMSEA = 0.040) with only a trivial reduction in reliability (MR went from 0.98 to 0.97). Furthermore, in the exploratory subsample, the test-level explained common variance (ECV; Reise, 2012) associated with the nicotine dependence (general) factor increased substantially from 0.72 to 0.86 indicating a more strongly unidimensional model. Fit indices for the nondaily smokers also suggest a strongly unidimensional item set (CFI = 0.987, TLI = 0.986, RMSEA = 0.038), with improvement in fit compared with the 63-item set (CFI = 0.956, TLI = 0.954, RMSEA = 0.044) and minimal loss in precision (MR went from 0.98 to 0.97). Similar to daily smoker results, the ECV associated with the nicotine dependence (general) factor in the nondaily sample solution increased from 0.81 to 0.89.

Differential Item Functioning

Next, the 33 daily and 34 nondaily smoker items underwent DIF testing according to gender, race/ethnicity (White, Black, and Hispanic), and age (18–30, 31–50, and 51+). For the daily smokers, across all comparisons, seven items met the wABC criteria for consideration of removal (i.e., wABC > 0.30), and six items were ultimately removed because of DIF. For the nondaily smokers, seven items were considered for removal, and all seven were ultimately removed. Information about the items removed due to DIF is summarized in Table 3. Notably, four items were identified as having DIF in both the daily and nondaily smoker samples.

Figure 1 illustrates two items that were removed from the daily smoker Nicotine Dependence item bank. The first item ("I am tempted to smoke when I am on the phone") was removed due to differences in functioning across males and females (wABC = 0.47; dEAP = 0.18). The second item ("I'm hooked on cigarettes") was removed due to differences across the

Table 3. Nicotine Dependence Items Removed Because of Differential Item Functioning (DIF)

Item stem	Number of comparisons with wABC > 0.3	DIF variable	wABC	dEAP
Daily smokers				
I am very much aware of when I am not smoking.	1	White vs. Black	0.62	0.23
I feel like I smoke all the time.	1	White vs. Black	0.58	0.23
I am tempted to smoke when I am on the phone.	1	Female vs. male	0.47	0.18
Compared to when I first started smoking, I need to smoke a lot more now.	1	White vs. Black	0.45	0.17
If I quit smoking, I will experience intense cravings for a cigarette.	1	White vs. Black	0.41	–0.19
I'm hooked on cigarettes.	1	Age 18–30 vs. age 51+	0.41	0.24
Nondaily smokers				
My attitude about my smoking is that I can 'take it or leave it' at any time.	6	Age 18–30 vs. age 51+	0.86	0.25
		Hispanic vs. White	0.55	0.17
		Age 18–30 vs. age 31–50	0.51	0.14
		Age 31–50 vs. age 51+	0.36	0.16
		Hispanic vs. Black	0.34	0.12
		Female vs. Male	0.33	0.02
I am very much aware of when I am not smoking.	2	White vs. Black	0.63	0.20
		Hispanic vs. Black	0.41	0.14
Compared to when I first started smoking, I need to smoke a lot more now.	1	Age 18–30 vs. age 51+	0.55	–0.35
I am tempted to smoke when I am watching TV.	2	Hispanic vs. Black	0.52	0.26
		White vs. Black	0.43	0.25
I'm hooked on cigarettes.	3	Age 18–30 vs. age 51+	0.47	0.26
		Hispanic vs. White	0.36	0.20
		Hispanic vs. Black	0.33	–0.18
After eating I want a cigarette.	4	Age 18–30 vs. age 51+	0.45	0.20
		Hispanic vs. Black	0.41	0.20
		Age 31–50 vs. age 51+	0.31	0.17
		Hispanic vs. White	0.30	0.03
I am tempted to smoke when I am on the phone.	2	Age 18–30 vs. age 51+	0.40	0.19
		Female vs. male	0.39	0.20

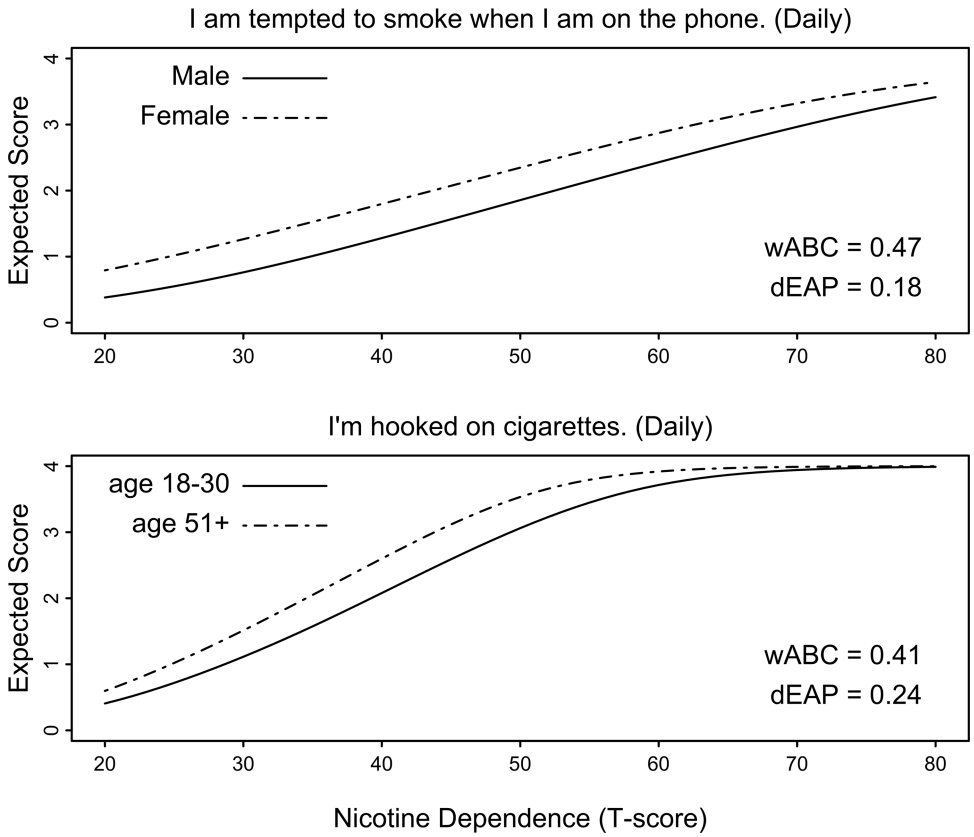

Figure 1. Two example items with differential item functioning that were removed from the daily smoker Nicotine Dependence item bank.

youngest and oldest age groups (wABC = 0.41; dEAP = 0.24). The *x*-axes in Figure 1 use the standard PROMIS *T*-score scale with mean 50 and standard deviation 10. The nonoverlapping curves for the two items indicate that the expected scores at a given level of dependence depend on the grouping variable (even after accounting for differences in the groups' means and variances). Specifically, women provide consistently higher ratings than men on the "…on the phone" item, given equal levels of nicotine dependence. Similarly, daily smokers who are 51 and older provide higher ratings on the "hooked on cigarettes" item than those who are 30 and younger, except at the highest levels of overall nicotine dependence.

Calibration of Item Banks

Using the two-group IRT model with daily smokers as the reference group, 32 total items were calibrated. Within this set, 20 were anchor items (identical item parameters for daily and nondaily smokers) and 2 had unique item parameters for daily and nondaily smokers. In addition, there were five items per bank that were nonoverlapping (i.e., items that only occur for that particular smoker group). This process resulted in two Nicotine Dependence item banks (one for daily and one for nondaily smokers) each with a total of 27 items. As can be seen in Table 4, the final items tended to be strongly related to the underlying nicotine dependence construct (*a* parameters for items in both banks ranged from 1.41 to 3.66) and covered a wide range of the nicotine dependence continuum (*b* parameters ranged from –2.33 to 2.23) that is fairly symmetric around the nicotine dependence mean.

Figure 2 illustrates the score reliability for the daily and nondaily smoker Nicotine Dependence item banks (and SFs) on a standard *T*-score scale. Full bank scores have reliability values greater than 0.9 from nearly three standard deviations below the mean to three standard deviations above the mean (i.e., from 20 to 80, in the *T*-score scale). Nondaily smokers had a mean value of 39.3, 1.07 standard deviations below the daily smoker mean of 50. In addition, the nondaily smoker sample had slightly more nicotine dependence variability (*SD* = 11.3) compared with daily smokers (*SD* = 10).

Nicotine Dependence SFs

Combinations of the 20 anchor items were examined, and 4- and 8-item SFs were ultimately selected, following the procedures described in Hansen et al. The items comprising these SFs are indicated in Table 4, and the summed score to IRT score translation table for both SFs is contained in Table 5. Figure 2 shows the reduction in score reliability when going from the complete item banks (of 27 items each) to the SFs. Despite this reduction, the marginal reliabilities of the SF scores remain quite good (0.91 for the eight-item SF, 0.81 for the four-item SF). In addition, these scores correlate strongly (0.90–0.96) with those obtained from the complete banks. The results suggest that the four- and eight-item SFs provide efficient and reliable measures of nicotine dependence. The particular choice of SF may be guided by considerations related to respondent burden, the desired levels of score precision, and the context for form administration and scoring.

Table 4. Nicotine Dependence Item Banks for Daily and Nondaily Smokers

Item	D/ND	CAT		Item parameters				
		D	ND	*a*	b_1	b_2	b_3	b_4
When I haven't been able to smoke for a few hours, the craving gets intolerable. (4,8 SF)[a]	Both	0.01	0.14	3.13	−1.25	−0.31	0.64	1.38
I drop everything to go out and buy cigarettes. (4,8 SF)[a]	Both	0.11	0.02	1.84	−1.23	−0.03	1.10	2.03
I smoke more before going into a situation where smoking is not allowed. (4,8 SF)[a]	Both	0.03	0.04	1.57	−1.96	−0.99	0.17	1.19
I find myself reaching for cigarettes without thinking about it. (4,8 SF)[a]	Both	0.56	0.34	1.41	−1.54	−0.48	0.80	2.16
My urges to smoke keep getting stronger if I don't smoke. (8 SF)	Both	0.09	0.01	2.56	−1.46	−0.57	0.20	0.97
After not smoking for a while, I need to smoke in order to avoid feeling any discomfort. (8 SF)[a]	Both	0.01	0.10	2.33	−1.35	−0.53	0.48	1.34
When I'm really craving a cigarette, it feels like I'm in the grip of some unknown force that I cannot control. (8 SF)[a]	Both	0.02	0.00	2.27	−1.20	−0.36	0.59	1.44
I crave cigarettes at certain times of day. (8 SF)	Both	0.06	0.01	1.62	−2.13	−1.11	−0.04	0.94
When I go without a cigarette for a few hours, I experience craving.	Both	1.00	1.00	3.66	−1.35	−0.53	0.09	0.73
I frequently crave cigarettes.	Both	0.87	0.85	3.18	−1.75	−0.69	0.08	0.90
Cravings for a cigarette make it difficult for me to quit.	Both	0.77	0.90	3.17	−1.76	−0.95	−0.31	0.40
My desire to smoke seems overpowering.	Both	0.55	0.39	3.02	−1.19	−0.46	0.28	0.98
It is hard to ignore urges to smoke.	Both	0.15	0.52	2.99	−1.86	−0.96	−0.19	0.62
When I go too long without a cigarette I feel impatient.[a]	Both	0.15	0.19	2.82	−1.64	−0.81	0.22	1.09
When I go too long without a cigarette I get strong urges that are hard to get rid of.[a]	Both	0.14	0.20	2.81	−1.65	−0.77	0.23	1.02
I get a real gnawing hunger for a cigarette when l haven't smoked in a while.[a]	Both	0.04	0.28	2.80	−2.03	−1.04	0.04	0.90
When I run out of cigarettes, I find it almost unbearable.[a]	Both	0.01	0.03	2.56	−1.56	−0.67	0.21	0.93
The idea of not having any cigarettes causes me stress.	Both	0.00	0.00	2.49	−1.40	−0.44	0.22	0.90
It is hard for me to go without smoking for a whole day.	Both	0.03	0.17	2.41	−2.00	−1.21	−0.63	0.02
I am tempted to smoke when I realize I haven't smoked for a while.[a]	Both	0.01	0.00	1.64	−2.33	−1.26	0.09	1.23
Smoking is a large part of my daily life.	Both	0.03	0.00	1.52	−2.19	−1.04	−0.06	0.90
I smoke even when I am so ill that I am in bed most of the day.[a]	Both	0.04	0.01	1.46	−1.21	0.05	1.14	2.23
The thought of never smoking again is overwhelming.	D	0.03		1.56	−0.97	−0.08	0.74	1.44
After eating I want a cigarette.[a]	D	0.00		1.21	−4.14	−3.04	−1.49	−0.12
I am tempted to smoke when I am happy.[a]	D	0.00		0.98	−4.02	−2.14	0.68	2.43
I smoke when I am alone.[a]	D	0.00		0.81	−6.63	−5.01	−2.02	0.62
I am tempted to smoke when I am driving.[a]	D	0.00		0.74	−3.37	−2.48	−1.08	0.81
I would go crazy if I couldn't smoke.	ND		0.02	2.40	−1.13	−0.13	0.59	1.28
I feel like I smoke all the time.	ND		0.00	2.09	−0.78	−0.18	0.53	1.11
If I quit smoking, I will experience intense cravings for a cigarette.	ND		0.19	1.80	−2.30	−1.25	−0.38	0.44
My life is full of reminders to smoke.	ND		0.08	1.61	−1.72	−0.59	0.41	1.31
I become more addicted the more I smoke.	ND		0.04	1.53	−1.35	−0.50	0.50	1.38

Note. SF = short form. D/ND column indicates if the item parameters were identical in daily and nondaily groups (both), unique to the daily group (D), or unique to the nondaily group (ND). CAT column indicates the rate of item administration for the 10-item maximum condition. Item slope and threshold parameters were obtained through calibrations of the full item banks.
[a]Indicates items that used the following response options: 0 = never, 1 = rarely, 2 = sometimes, 3 = often, 4 = always. All other items used the following response options: 0 = not at all, 1 = a little bit, 2 = somewhat, 3 = quite a bit, 4 = very much.

CAT Simulations

CAT simulations were conducted on the daily and nondaily smoker Nicotine Dependence item banks. Table 6 provides the results of simulations that used a standard error of 3.0 (in the *T*-score metric) as the CAT stopping criterion, which corresponds to a reliability of slightly greater than 0.90, and a range of limits on the maximum number of items allowed to be administered (4, 6, 8, 10, and 12). To summarize these results, the correlation between CAT and full bank scores is greater than 0.95, and the average CAT will terminate with a standard error of 3 when about four to five items have been administered (or about 5 or 6 for nondaily smokers). Item administration rates for the 10-item CAT simulation condition are displayed in Table 4. For the daily bank, only five

Table 5. Nicotine Dependence Summed Score to Scaled Score Translation Table for the 4- and 8-Item Short Forms

4-item Short Form			8-item Short Form		
Summed score	Scaled score (*T*)	Standard error	Summed score	Scaled score (*T*)	Standard error
0	27	6.3	0	23	5.7
1	32	5.4	1	27	4.6
2	35	5.2	2	30	4.3
3	38	5.0	3	32	4.0
4	41	4.8	4	34	3.8
5	43	4.8	5	35	3.6
6	46	4.7	6	37	3.5
7	48	4.7	7	38	3.4
8	50	4.6	8	40	3.3
9	53	4.6	9	41	3.3
10	55	4.6	10	42	3.2
11	57	4.6	11	44	3.2
12	60	4.7	12	45	3.2
13	63	4.7	13	46	3.2
14	65	4.9	14	47	3.2
15	69	5.2	15	48	3.2
16	73	6.0	16	50	3.2
			17	51	3.2
			18	52	3.2
			19	53	3.2
			20	54	3.2
			21	55	3.2
			22	57	3.2
			23	58	3.2
			24	59	3.2
			25	61	3.3
			26	62	3.3
			27	63	3.4
			28	65	3.6
			29	67	3.8
			30	69	4.1
			31	72	4.5
			32	75	5.4

items are administered at a rate of 0.5 or higher, and only four are administered at that high rate for the nondaily bank. Thus, after administration of a handful of items, the subsequent item selection rates are fairly similar. The items that are administered at high frequency tend to correspond across the daily and nondaily banks but do not correspond very closely with the four- and eight-item SFs. This is because the SF selection emphasized content coverage, whereas the CAT selection algorithm in our simulations relied exclusively on empirical information.

DISCUSSION

The purpose of this article was to present the development and refinement of two new item banks to assess nicotine dependence among daily and nondaily cigarette smokers. These item banks were constructed as part of the PROMIS Smoking Initiative, a comprehensive effort designed to advance a more unified framework for cigarette smoking assessment. In this study, a core set of 32 items was calibrated for daily and nondaily smokers; 22 items were identical for daily and nondaily smokers, with 5 items unique for daily smokers and 5 items unique for nondaily smokers. For both daily and nondaily smokers, scores from items comprising the Nicotine Dependence item banks are highly reliable and strongly unidimensional. Our evaluations of DIF suggest that the item banks perform similarly for men and women, adult smokers of different ages, and for smokers of different racial and ethnic backgrounds.

Item content covers a range of conceptually relevant domains that represent the different expressions of cigarette nicotine dependence (Rosenthal et al., 2011); for example, items relating to craving ("I crave cigarettes at certain times of the day"), withdrawal symptoms ("When I go too long without a cigarette I feel impatient"), and compulsive use ("I drop everything to go out and buy cigarettes") are featured prominently in the daily and nondaily smoker item banks. The conceptually oriented feature of the Nicotine Dependence item banks is not surprising, given that many of the items were drawn from existing conceptually rich measures of cigarette nicotine dependence, like the NDSS

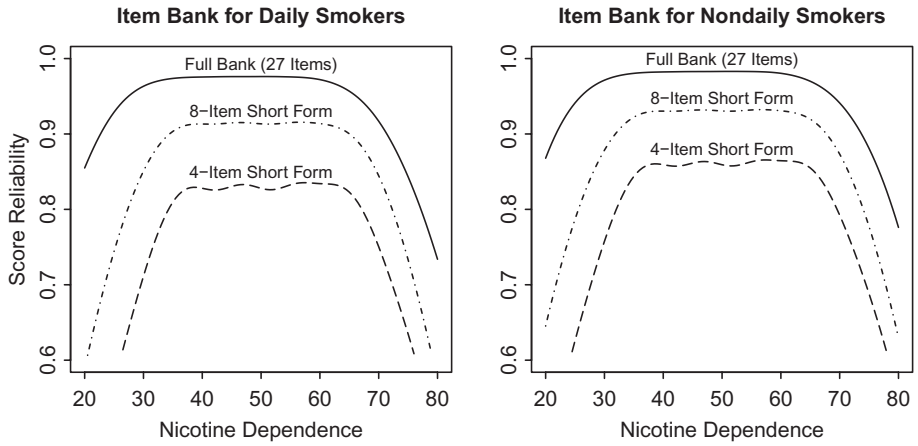

Figure 2. Score reliability for the daily and nondaily smoker Nicotine Dependence item banks.

Table 6. Simulated Adaptive Tests for the Nicotine Dependence Item Banks

	Daily smokers						Nondaily smokers					
	Maximum no. of items						Maximum no. of items					
	4	6	8	10	12	All items	4	6	8	10	12	All items
Average items administered	3.82	4.36	4.58	4.72	4.83	27	3.90	4.75	5.16	5.52	5.81	27
Proportion receiving maximum items	0.82	0.18	0.09	0.07	0.05	1	0.90	0.32	0.19	0.17	0.14	1
Marginal reliability	0.89	0.90	0.91	0.91	0.91	0.97	0.89	0.91	0.92	0.92	0.92	0.97
r (T_{CAT},T_{full})	0.95`	0.95	0.96	0.96	0.96	0.98	0.95	0.96	0.96	0.96	0.96	0.98

(Shiffman et al., 2004) and WISDM (Piper et al., 2004), that were developed from a strong theoretical frame that viewed nicotine dependence as a syndrome composed of multiple features (Shadel et al., 2000). The PROMIS Nicotine Dependence item banks differ from these other scales in that they use a smaller, more efficient set of conceptually diverse items, have been specifically tested and found appropriate for both daily and nondaily smokers, and offer a variety of practical assessment options.

Indeed, although longer forms can be generated based on the item banks (up to 27 items each for daily and nondaily smokers), key features of this investigation were the development of SFs and the simulation of CAT (computer adaptive testing). SFs and CAT allow for the administration of a relatively small number of items in research settings where assessment using the full item bank is not desired or feasible. Two highly reliable SFs (four and eight items) were developed, and the scores obtained from these forms were strongly correlated with assessments based on the full item banks. These SFs are identical for daily and nondaily smokers and in fact are shorter than the Fagerström Tolerance Questionnaire and its revision (Fagerström, 1978; Heatherton et al., 1991). Similarly, CAT simulations showed that an average of fewer than six items could be administered with only minimal loss of information or reliability. Thus, the PROMIS Nicotine Dependence item banks present researchers and clinicians with a host of administration options—all of which are psychometrically strong. A key point is that the SFs are designed to balance efficiency and breadth of coverage of concepts relating to the (cigarette) nicotine dependence syndrome (see Shadel et al., 2000). In contrast, the CAT optimizes test length only, often leading to more narrow content coverage. Thus, users who desire more conceptual diversity in a SF assessment of nicotine dependence would likely opt to use the SF versions.

The full Nicotine Dependence item banks (and consequently the SFs) did not include the item "time to first cigarette in the morning." This is the first item of the Fagerström Tolerance Scale (and is also one of two items in the Heaviness of Smoking Index; Heatherton et al., 1989) and has been shown to be a strong, single item predictor of smoking cessation outcomes (Baker et al., 2007). This item was included in earlier iterations of the Nicotine Dependence item banks (see Edelen et al., 2012), but it failed to reach the psychometric threshold of other items in the bank (as presented in this article) and was therefore discarded. In other words, this item is not calibrated with the other items in the Nicotine Dependence item bank and thus is not included in the bank itself. However, due to the interest in this item and its seeming clinical importance, it is available via the project Web site (http://www.rand.org/health/projects/promis-smoking-initiative.html) as part of a supplementary item set reflecting smoking rates and quitting history that can be administered outside of the items that comprise the banks. Contents of the banks themselves are also available for download from the project Web site as well as via Assessment Center.

The approach taken by the PROMIS Smoking Initiative to develop the Nicotine Dependence item banks is unique in several respects. First, unlike most nicotine dependence scale development efforts that have sought to fill gaps in or improve upon measures already in use (e.g., the NDSS was developed to measure nicotine dependence as a syndrome, filling a gap because the FTQ and FTND were not developed within a syndromal framework; Shiffman et al., 2004), the PROMIS Smoking Initiative used these measures, and others, as conceptual and empirical starting points, forming its item banks from items that have already been evaluated and tested (and supplemented with new items generated from focus groups of current smokers; see Edelen et al., 2012). Second, the PROMIS Smoking Initiative is among the first to use contemporary methods in psychometrics, like IRT, to evaluate and refine the assessment items. These methods have only rarely been adopted in tobacco control research and usually on a more limited capacity. For example, IRT has been used in a more focused way to examine nicotine dependence in single scales (Courvoisier & Etter, 2008). An advantage of using IRT is that the items are calibrated on the same scale, so that the selection of any subset, SF, or brief administration of the items will result in IRT-scaled scores comparable to the results from this calibration sample. For example, in a study where the outcome of interest is meeting criteria for a nicotine dependence diagnosis at a cut point along the nicotine dependence continuum, a subset of these items could be "handpicked" by the researcher that best discriminate at that cut point, thus avoiding the burden of having to administer the complete item bank. Third, the item bank is remarkable for its psychometric strength, as well as for its flexibility in administration options (e.g., SFs, CAT, computerized, or paper and pencil). Finally, to the extent that the definition of the construct of nicotine dependence evolves or changes, new items can be easily added to the item bank for further evaluation and testing.

Several future research directions are planned for this item bank and for the PROMIS Smoking Initiative more generally. The final article in this supplement presents preliminary validity information for the six banks including estimated

correlations among the six item bank scores and between the six item banks and other PROMIS measures (e.g., alcohol consumption, physical functioning, and emotional distress). This article also characterizes differences in item bank scores according to select discrete demographic (e.g., gender, age, and ethnicity) and smoking behavior subgroups (e.g., motivation to quit, quantity of smoking). Results are compared for daily and nondaily item bank scores. Further data collection activities are underway to replicate these findings in an independent sample, establish test–retest reliability, and evaluate the bank scores' sensitivity to change over time. This work also includes development of scoring crosswalks from traditional smoking measures to the new item bank scores. For example, research will assess items from other legacy measures of cigarette nicotine dependence (e.g., NDSS, WISDM), allowing scores from established scales to be "crosswalked" to the Nicotine Dependence item banks using simultaneous IRT calibration. Important directions for future research include studying the utility of these items for predicting treatment outcome and response to different smoking cessation treatments as well as examining the relationship between these items and indices of nicotine dependence such as withdrawal and cotinine levels. These latter steps are particularly important in order to provide additional evidence for the validity and clinical utility of these item banks.

FUNDING

This work was supported by the National Institute on Drug Abuse (R01DA026943 to MOE). Additionally, MH was supported by the Institute for Education Sciences with a predoctoral training grant (R305B080016), and LC was supported by the National Institute on Drug Abuse (R01DA030466).

DECLARATION OF INTERESTS

None declared.

ACKNOWLEDGMENTS

The authors would like to thank the PROMIS Smoking Initiative Advisory Group: Ronald D. Hays and Michael Ong, UCLA; David Cella, Feinberg School of Medicine, Northwestern University; Daniel McCaffrey, Educational Testing Service; Raymond Niaura, American Legacy Foundation, Brown University; Paul Pilkonis, University of Pittsburgh; and David Thissen, University of North Carolina at Chapel Hill.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychological Association.

Baker, T., Piper, M. E., McCarthy, D. E., Bolt, D. M., Smith, S., Kim, S., … Toll, B. (2007). Time to first cigarette in the morning as an index of ability to quit smoking: Implications for nicotine dependence. *Nicotine & Tobacco Research*, *9*, S555–S570. doi:10.1080/14622200701673480

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Cai, L. (2010). High-dimensional exploratory item factor analysis by a Metropolis-Hastings Robbins-Monro algorithm. *Psychometrika*, *75*, 33–57. doi:10.1007/s11336-009-9136-x

Cai, L., du Toit, S. H. C., & Thissen, D. (2011). IRTPRO: Flexible, multidimensional, multiple categorical IRT modeling [Computer software]. Chicago, IL: Scientific Software International.

Cai, L., Yang, J. S., & Hansen, M. (2011). Generalized full-information item bifactor analysis. *Psychological Methods*, *16*, 221–248. doi:10.1037/a0023350

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., … Rose, M.; PROMIS Cooperative Group. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Medical Care*, 45(5 Suppl. 1), S3–S11. doi.org/:10.1097/01.mlr.0000258615.42478.55

Chen, W.-H., & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, *22*, 265–289. doi:10.3102/10769986022003265

Choi, S. W. (2009). Firestar: Computerized adaptive testing simulation program for polytomous item response theory models. *Applied Psychological Measurement*, *33*, 644–645. doi:10.1177/0146621608329892

Colby, S. M., Tiffany, S., Shiffman, S., & Niaura, R. S. (2000). Are adolescent smokers dependent on nicotine? A review of the evidence. *Drug and Alcohol Dependence*, *59*, S83–S95. doi:10.1016/S0376-8716(99)00166-0

Courvoisier, D. S., & Etter, J. F. (2018). Using item response theory to study the convergent and discriminant validity of three questionnaires measuring cigarette dependence. *Psychology of Addictive Behaviors*, *22*, 391–401. doi:10.1037/0893-164X.22.3.391

Edelen, M. O., Thissen, D., Teresi, J., Kleinman, M., & Ocepek-Welikson, K. (2006). Identification of differential item functioning using Item Response Theory and the likelihood-based model comparison approach: Application to the Mini-Mental Status Examination. *Medical Care*, *44*(Suppl. 3), S134–S142. doi:10.1097/01.mlr.0000245251.83359.8c

Edelen, M. O., Tucker, J. S., Shadel, W. G., Stucky, B. D., & Cai, L. (2012). Toward a more systematic assessment of smoking: Development of a smoking module for PROMIS®. *Addictive Behaviors*, *37*, 1278–1284. doi:10.1016/j.addbeh.2012.06.016

Etter, J. F., Le Houezec, J., & Perneger, T. V. (2003). A self-administered questionnaire to measure dependence on cigarettes: The cigarette dependence scale. *Neuropsychopharmacology*, *28*, 359–370. doi:10.1038/sj.npp.1300030

Fagerström, K. O. (1978). Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addictive Behaviors*, *3*, 235–241. doi:10.1016/0306-4603(78)90024-2

Fiore, M. C., Jaen, C. R., Baker, T. B., Bailey, W. C., Benowitz, N. L., Curry, S. J., … Wewers, M. E. (2008). *Clinical practice guideline—Treating tobacco use and dependence: 2008 update*. Rockville, MD: Public Health Service.

Gibbons, R. D., & Hedeker, D. R. (1992). Full-information item bifactor analysis. *Psychometrika*, *57*, 423–436. doi:10.1007/BF02295430

Hansen, M., Cai, L., Stucky, B., Tucker, J., Shadel, W. G., & Edelen, M. O. (2014) Methodology for developing and evaluating the PROMIS® smoking item banks. *Nicotine & Tobacco Research*, *16*, S174–S188.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerström, K. O. (1991). The Fagerström test for nicotine dependence: A revision of the Fagerström Tolerance Questionnaire. *British Journal of Addiction*, *86*, 1119–1127. doi:10.1111/j.1360-0443.1991.tb01879.x

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., Rickert, W., & Robinson, J. (1989). Measuring the heaviness of smoking: Using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. *British Journal of Addiction*, *84*, 791–799. doi:10.1111/j.1360-0443.1989.tb03059.x

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118

Hughes, J. R. (2011). The hardening hypothesis: Is the ability to quit decreasing due to increasing nicotine dependence? A review and commentary. *Drug and Alcohol Dependence*, *117*, 111–117. doi:10.1016/j.drugalcdep.2011.02.009

Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

Niaura, R., & Shadel, W. G. (2003). Screening and assessment. In D. B. Abrams, R. Niaura, R. Brown, K. Emmons, M. G. Goldstein, & P. M. Monti (Eds.), *The tobacco dependence treatment handbook* (pp. 27–72). New York, NY: Guilford.

Orlando, M., & Marshall, G. N. (2002). Differential item functioning in a Spanish translation of the PTSD checklist: Detection and evaluation of impact. *Psychological Assessment*, *14*, 50–59. doi:10.1037/1040-3590.14.1.50

Piper, M. E., McCarthy, D. E., & Baker, T. B. (2006). Assessing tobacco dependence: A guide to measure evaluation and selection. *Nicotine & Tobacco Research*, *8*, 339–351. doi:10.1037/1040-3590.14.1.50

Piper, M. E., Piasecki, T. M., Federman, E. B., Bolt, D. M., Smith, S. S., Fiore, M. C., & Backer, T. B. (2004). A multiple motive approach to tobacco dependence: The Wisconsin Inventory of Smoking Dependence Motives (WISDM-68). *Journal of Consulting and Clinical Psychology*, *72*, 139–154. doi:10.1037/0022-006X.72.2.139

Pomerleau, C. S., Carton, S. M., Lutzke, M. L., Flessland, K. A., & Pomerleau, O. F. (1994). Reliability of the Fagerström Tolerance Questionnaire and the Fagerström test for nicotine dependence. *Addictive Behaviors*, *19*, 33–39. doi:10.1016/0306-4603(94)90049-3

Reeve, B., Hays, R. D., Bjorner, J., Cook, K., Crane, P. K., Teresi, J. A., & Cella, D. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the Patient-Reported Outcome Measurement Information System (PROMIS). *Medical Care*, *45*, S22–S31. doi:10.1097/01.mlr.0000250483.85507.04

Reise, S. P. (2012). Invited paper: The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696. doi:10.1080/00273171.2012.715555

Rosenthal, D. G., Weitzman, M., & Benowitz, N. L. (2011). Nicotine addiction: Mechanisms and consequences. *International Journal of Mental Health*, *40*, 22–38. doi:10.2753/IMH0020-7411400102

Shadel, W. G., Shiffman, S., Niaura, R., Nichter, M., & Abrams, D. B. (2000). Current models of nicotine dependence: What is known and what is needed to advance understanding of tobacco etiology among youth. *Drug and Alcohol Dependence*, *59*(Suppl. 1), S9–22. doi:10.1016/S0376-8716(99)00162-3

Shiffman, S., Ferguson, S. G., Dunbar, M. S., & Scholl, S. M. (2012). Tobacco dependence among intermittent smokers. *Nicotine & Tobacco Research*, *14*, 1372–1381. doi:10.1093/ntr/nts097

Shiffman, S., Ferguson, S., Gwaltney, C., Balabanas, M., & Shadel, W. G. (2006). Reduction of abstinence induced withdrawal and craving using nicotine replacement therapy. *Psychopharmacology*, *184*, 637–644. doi:10.1007/s00213-005-0184-3

Shiffman, S., & Paty, J. (2006). Smoking patterns and dependence: Contrasting chippers and heavy smokers. *Journal of Abnormal Psychology*, *115*, 509–523. doi:10.1037/0021-843X.115.3.509

Shiffman, S., Waters, A., & Hickcox, M. (2004). The nicotine dependence syndrome scale: A multidimensional measure of nicotine dependence. *Nicotine & Tobacco Research*, *6*, 327–348. doi:10.1080/1462220042000202481

Stolerman, I., & Jarvis, M. (1995). The scientific case that nicotine is addictive. *Psychopharmacology*, *117*, 2–10. doi:10.1007/BF02245088

Stucky, B. D., Thissen, D., & Edelen, M. O. (2013). Using logistic approximations of marginal trace lines to develop short assessments. *Applied Psychological Measurement*, *37*, 41–57. doi:10.1177/0146621612462759

Thissen, D., Nelson, L., Rosa, K., & McLeod, L. D. (2001). Item response theory for items scored in more than two categories. In D. Thissen & H. Wainer (Eds.), *Test scoring* (pp. 141–186). Mahwah, NJ: Lawrence Erlbaum & Associates.

Tiffany, S. T., & Drobes, D. J. (1991). The development and initial validation of a questionnaire on smoking urges. *British Journal of Addiction*, *86*, 1467–1476. doi:10.1111/j.1360-0443.1991.tb01732.x