Psychometric Evaluation of Behavioral Intention Item Functioning Across Tobacco Product Categories

Stacey McCaffrey, Ryan Black, Stephanie Plunkett
Altria Client Services LLC, Richmond, VA, USA

SRNT 26th Annual Meeting
March 11-14, 2020, New Orleans, LA, USA
http://sciences.altria.com

This scientific research is presented by Altria Client Services LLC (ALCS). ALCS affiliate companies are tobacco product manufacturers.

ABSTRACT

As part of an FDA tobacco product application, FDA guidance recommends that applicants evaluate adults' behavioral intentions toward the candidate tobacco product, including trial, use, dual use and switching intentions. Altria Client Services previously developed and validated behavioral intention (BI) scales to support future FDA filling for an e-vapor product. However, the psychometric properties of these scales when modified to reference other tobacco product categories have not been evaluated. Therefore, the purpose of the current study was to determine whether the BI scales are valid when modified to reference an oral tobacco-derived nicotine (TDN) containing product and a moist smokeless tobacco (MST) product.

Data were extracted from two previously conducted studies, whereby the BI scales were modified to specify an oral TDN ("Study 1"; N=4,118) and an MST product ("Study 2"; N=5,871). These studies included current, never, and former tobacco product users. Rasch modeling and classical test theory approaches were utilized to evaluate rating scale functioning, unidimensionality, reliability, validity, and bias via differential item function (DIF). Additional DIF analyses were conducted to determine whether item functioning was substantially different across tobacco product categories (i.e., e-vapor, oral TDN, MST).

For both Study 1 and Study 2, Rasch analyses revealed that the BI items' Likert-type rating scales were functioning appropriately. Results provided support for unidimensionality, excellent internal consistency reliability, and convergent validity. Rasch-based DIF analyses did not suggest substantial bias based on age, race, gender, or tobacco use status. Finally, DIF analyses revealed that the BI items functioned similarly across tobacco products (i.e., e-vapor, oral TDN, MST).

These results provide strong evidence that the BI scales continue to exhibit strong psychometric properties when modified to reference other tobacco products, namely an oral TDN and an MST product. Future research could evaluate the predictive validity of these scales.

BACKGROUND

- As part of an FDA tobacco product application, FDA guidance recommends that applicants evaluate adult tobacco users' and nonusers' behavioral intentions toward the candidate tobacco product, including trial, use, dual use and switching intentions (PMTA Proposed Rule, FDA, 2019).
- Altria Client Services (ALCS) previously developed and validated behavioral intention scales for use with tobacco users and nonusers to support future FDA fillings for an e-vapor product (Parker Zdinak et al., 2018). The scales included intentions to try, use, dual use, and switch.
- ALCS BI scales were developed and validated in accordance with guidance and best practices (FDA Patient-Reported Outcome [PRO] Guidance for Industry, 2009). Specifically, these items were iteratively revised through cognitive testing with end-users before being subject to empirical evaluations, which included evaluation of rating scale functioning, unidimensionality, reliability (internal consistency, stability, and Rasch-derived reliability), validity, ability to detect change, and bias via differential item functioning (DIF).
- Although it is reasonable to expect that these scales will function adequately when modified to reference tobacco product categories other than e-vapor, it is worthwhile to explicitly evaluate functioning of modified items. FDA and International Society for Pharmacoeconomics and Outcomes Research (ISPOR) guidance recommend that sponsors provide evidence to confirm an instrument's adequacy once it has been modified (FDA PRO Guidance, 2009; Rothman et al., 2009).
- The purpose of the current study was to determine whether the BI scales are valid when modified to reference other tobacco product categories, namely, an oral TDN containing product and a MST product.

CONCLUSIONS

The ALCS Behavioral Intention scales appear to be reliable and valid tools for capturing intentions toward tobacco product use among diverse groups of adult tobacco users and nonusers. Specifically, these items exhibit similar psychometric functioning across various tobacco user and nonuser groups, and do not appear to function substantially differently based on respondent gender, age, or race. Further, the current study provides evidence that the Behavioral Intention scales' psychometric properties do not substantially differ when specifying an e-vapor product, an oral TDN product, or an MST product.

STRENGTHS AND LIMITATIONS

- This study utilized secondary analyses of previously collected data to evaluate psychometric functioning and invariance of the ALCS Behavioral Intention scales when modified to reference an oral TDN and MST product. The large sample sizes permitted us to split the sample into validation and cross-validation samples for purposes of confirming the stability of reliability and validity coefficients across sampling.
- Given the nature of the studies, it was not possible to evaluate test-retest reliability (stability). This could be evaluated in future research. Future research might also explore the predictive validity of the ALCS Behavioral Intention scales.
- The ALCS Behavioral Intention scales were developed and validated with adult tobacco users and nonusers; therefore, the psychometric properties of these scales for use with youth is unknown.

REFERENCES

-FDA. (2009). Guidance of industry: Patient-reported outcome measures: Use in medical product development to support labeling claims. Retrieved from https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf.
-FDA. (2019). Premarket Tobacco Product Applications and Recordkeeping Requirements. Federal Registrar, 84 (186). Retrieved from https://www.federalregister.gov/documents/2019/09/25/2019-20315/premarket-tobacco-product-applications-and-recordkeeping-requirements
-IBM. (2017). IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.
-Linacre, J. M. (2017). Winsteps® Rasch measurement computer program. Beaverton, Oregon: Winsteps.com.
-Linacre, J. M. (2019). Winsteps Rasch measurement computer program user's guide. Beaverton, Oregon: Winsteps.com.
-Masters, G. N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47, 149-174.
-Parker Zdinak, P., Black, R. A., Plunkett, S., McCaffrey, S., & Chow, S. (2018, February). Validation of Perceptions and Behavioral Intentions Survey: Psychometric evaluation of tobacco-related behavioral intentions to try, use, dual use, and switch. Poster presented at the Society for Research on Nicotine and Tobacco 24th Annual Meeting.
-Rothman, M., Burke, L., Erickson, P., Leidy, N. K., Patrick, D. L., & Petrie, C. D. (2009). Use of existing patient-reported outcome (PRO) instruments and their modifications: The ISPOR good research practices for evaluating and documenting content validity for the use of existing instruments and their modification PRO task force report. Value in Health, 12(8), 1075-1083.

METHODS

BEHAVIORAL INTENTION SCALES

Table 1: ALCS Behavioral Intention Scales

Scale	Item #	Item Content
Intention to Try	Try1	I am open to trying an on!® Nicotine Pouch product in the next 30 days.
	Try2*	Based on what you know about [product], how likely or unlikely are you...? to try [product]
	Try3*	Based on what you know about [product], how likely or unlikely are you...? to try [product] if one of your best friends were to offer [product] to you
Intention to Use	Use1	I would consider using [product] more than once
	Use2	I expect to use [product]
	Use3	It is likely that I will regularly use [product] in the next 6 months
	Use4	[Product] will be my regular brand of [product category] in the next 30 days
Intention to Dual Use	DualUse1	I plan to use [product] in addition to regular cigarettes
Intention to Switch	Switch1	I plan to gradually switch from regular cigarettes to [product]
	Switch2	I plan on using [product] as a complete replacement for regular cigarettes
	Switch3	I intend on switching from cigarettes to [product] in the next 6 months

Rating scale: 1=Strongly disagree, 2=Disagree, 3=Somewhat disagree, 4=Somewhat agree, 5=Agree, 6=Strongly agree,
*These items utilized a different rating scale: 1=Definitely not, 2=Very unlikely, 3=Somewhat unlikely, 4=Somewhat likely, 5=Very likely, 6=Definitely

PHASE 1 DATA SOURCES

Data for Phase 1 were extracted from two previously conducted studies, whereby the BI scales were modified to specify an oral TDN ("Study 1"; N=4,118) and an MST product ("Study 2"; N=5,871).

Study 1 included behavioral intention data (N=4,118) from the following five groups:
(1) adult smokers planning to quit in the next 30 days (ASPQ)
(2) adult smokers not planning to quit in the next 30 days (ASNPQ)
(3) adult other tobacco product users (Other)
(4) adult former tobacco users (Former)
(5) adult never tobacco users (Never)

Study 2 included behavioral intention data (N=5,871) from the following six groups:
(1) adult smokers planning to quit in the next 30 days (ASPQ)
(2) adult smokers not planning to quit in the next 30 days (ASNPQ)
(3) adult dual users of cigarettes and moist smokeless tobacco (Dual)
(4) adult moist smokeless tobacco users (MST)
(5) adult former tobacco users (Former)
(6) adult never tobacco users (Never)

PHASE 2 DATA SOURCES

Data for Phase 2 included the 2 data sources for Phase 1, as well as data from the original validation study, whereby the ALCS Behavioral Intention scales were developed and validated specifying an e-vapor product.
This study ("Study 3") included behavioral intention data (N=2,943) from the following five groups:

1) adult smokers planning to quit in the next 30 days (ASPQ)
(2) adult smokers not planning to quit in the next 30 days (ASNPQ)
(3) adult e-vapor users (E-Vapor)
(4) adult former tobacco users (Former)
(5) adult never tobacco users (Never)

ANALYTIC PLAN

Phase 1: Evaluation of the scale's psycho-metric properties when modified to reference oral TDN and MST

- To determine whether the Behavioral Intention scales are valid to specify these tobacco products, Rasch modeling and classical test theory approaches were utilized to evaluate:
  - rating scale functioning – evaluation of response option thresholds through a partial credit model (PCM; Masters, 1982)
  - reliability – internal consistency reliability (Cronbach's alpha) and Rasch-derived reliability (person reliability)
  - validity – Pearson correlations between the Behavioral Intention scales and a purchase intent item
  - bias via DIF – evaluation of bias for age (legal age to 24 years vs. >24 years), race (White/Caucasian vs. non-White/Caucasian), gender, and study group membership
- Additionally, assumptions of the Rasch measurement model, such as unidimensionality and equal item discriminations, were also evaluated.
  - the assumption of unidimensionality was evaluated through (1) Monte Carlo simulation studies ("parallel analysis") with 5,000 randomly generated parallel datasets, conducted on data from the validation sample, and (2) confirmatory factor analyses, conducted using data from the cross-validation sample.
- For classical test theory analyses (i.e., Cronbach's alpha, convergent validity), data were randomly split into validation and cross-validation samples to confirm that findings were stable across sampling.
- Analyses were conducted with SPSS version 25 (IBM, 2017), Amos version 25 (Arbuckle, 2017), and Winsteps version 4.0.0 (Linacre, 2017).

Phase 2: Evaluation of item invariance across tobacco products

DIF analyses were conducted to determine whether item functioning was substantially different across tobacco products (i.e., when the items referenced an e-vapor product, an oral TDN product, or an MST product). These analyses were conducted in Winsteps.

RESULTS

PHASE 1

Rating Scale Functioning

For both Study 1 and Study 2, the response category thresholds were ordered, indicating that the Behavioral Intention items' Likert-type rating scales were functioning appropriately. That is, it required a higher level of intention to endorse a higher level of agreement (e.g., strongly agree vs. agree) on the items' rating scales.

Unidimensionality, item fit, and discrimination

- For both Study 1 and 2, parallel analyses were conducted on data from the validation sample. Eigenvalues from the principal components analyses (PCA) were compared against the 95th percentile of the distribution of eigenvalues from the parallel datasets, and only the first eigenvalue was significant for each scale.
- Across both studies, results from CFAs with the cross-validation samples confirmed unidimensionality for each scale.
- For both Study 1 and 2, mean square infit and outfit chi-square fit statistics were all below 1.50, suggesting that the items fit the Rasch model (Linacre, 2019), and discrimination values were similar across items.

Reliability

- When modified to reference an oral TDN product (Study 1), Rasch-derived reliability (also called person reliability) was excellent for the Intention to Try (.90), Use (.93), and Switch scales (.94). This provides empirical support that the items accurately quantify persons with different levels of intention. Similarly, person reliability was excellent for the Intention to Try (.88), Use (.91), and Switch (.92) scales in Study 2.
- For both Study 1 and 2, internal consistency reliability, captured by Cronbach's alpha, was excellent (Table 2). This finding was confirmed with the cross-validation sample, suggesting that findings are stable over sampling.

Validity

As evidence of convergent validity, the Behavioral Intention scales and the Purchase Intent item were positively correlated in both Study 1 and 2 (Table 3). These findings were confirmed with the cross-validation sample.

Bias

For both Study 1 and 2, none of the items exhibited substantial DIF by gender (male/female), race (White/non-White), age (legal age to 24 years/ >24 years), or study group membership.

Table 2: Reliability Coefficients for the Behavioral Intention Scales When Modified to Reference Oral TDN and MST

Scale	Study Group	Study 1 (Oral TDN) Validation Sample n	α	Study 1 Cross-Validation Sample n	α	Study 2 (MST) Validation Sample n	α	Study 2 Cross-Validation Sample n	α
Try	All Study Groups	2070	.951	2048	.951	2823	.964	2683	.965
	ASPQ	378	.894	345	.904	426	.945	427	.950
	ASNPQ	454	.915	475	.906	508	.939	479	.941
	Other	356	.909	340	.929	-	-	-	-
	Dual	-	-	-	-	356	.903	360	.878
	MST	-	-	-	-	349	.916	368	.919
	Former	328	.938	330	.939	433	.924	402	.935
	Never	554	.941	558	.941	751	.910	647	.896
Use	All Study Groups	2070	.972	2048	.965	3009	.977	2862	.978
	ASPQ	378	.953	345	.949	426	.960	427	.966
	ASNPQ	454	.949	475	.937	508	.975	479	.968
	Other	356	.955	340	.940	-	-	-	-
	Dual	-	-	-	-	449	.917	444	.937
	MST	-	-	-	-	442	.948	463	.953
	Former	328	.963	330	.958	433	.970	402	.915
	Never	554	.959	558	.960	751	.958	647	.954
Switch	All Study Groups	832	.963	820	.960	1383	.968	1350	.966
	ASPQ	378	.961	345	.954	426	.963	427	.971
	ASNPQ	454	.963	475	.962	508	.969	479	.960
	Dual	-	-	-	-	449	.952	444	.949

Table 3: Convergent Validity Coefficients for the Behavioral Intention Scales When Modified to Reference Oral TDN and MST: Pearson Correlations Between Purchase Intent and Behavioral Intention Scales

	Study Group	Study 1 Validation Sample n	r (p)	Study 1 Cross-Validation Sample n	r (p)	Study 2 Validation Sample n	r (p)	Study 2 Cross-Validation Sample n	r (p)
Try	All Study Groups	2070	.660 (<.001)	2048	.647 (<.001)	2823	.753 (<.001)	2683	.772 (<.001)
	ASPQ	378	.529 (<.001)	345	.566 (<.001)	426	.662 (<.001)	427	.733 (<.001)
	ASNPQ	454	.600 (<.001)	475	.520 (<.002)	508	.685 (<.001)	479	.589 (<.001)
	Other	356	.576 (<.001)	340	.601 (<.001)	-	-	-	-
	Dual	-	-	-	-	356	.453 (<.001)	360	.487 (<.001)
	MST	-	-	-	-	349	.583 (<.001)	368	.656 (<.001)
	Former	328	.565 (<.001)	330	.544 (<.001)	433	.554 (<.001)	402	.575 (<.001)
	Never	554	.608 (<.001)	558	.597 (<.001)	751	.568 (<.001)	647	.626 (<.001)
Use	All Study Groups	2070	.712 (<.001)	2048	.697 (<.001)	2823	.793 (<.001)	2683	.800 (<.001)
	ASPQ	378	.620 (<.001)	345	.624 (<.001)	426	.705 (<.001)	427	.744 (<.001)
	ASNPQ	454	.649 (<.001)	475	.616 (<.001)	508	.733 (<.001)	479	.622 (<.001)
	Other	356	.657 (<.001)	340	.640 (<.001)	-	-	-	-
	Dual	-	-	-	-	356	.571 (<.001)	360	.581 (<.001)
	MST	-	-	-	-	349	.683 (<.001)	368	.714 (<.001)
	Former	328	.639 (<.001)	330	.585 (<.001)	433	.645 (<.001)	402	.558 (<.001)
	Never	554	.670 (<.001)	558	.646 (<.001)	751	.594 (<.001)	647	.737 (<.001)
Switch	All Study Groups	832	.610 (<.001)	820	.625 (<.001)	1290	.672 (<.001)	1266	.675 (<.001)
	ASPQ	378	.636 (<.001)	345	.654 (<.001)	426	.714 (<.001)	427	.762 (<.001)
	ASNPQ	454	.587 (<.001)	475	.601 (<.001)	508	.697 (<.001)	479	.598 (<.001)
	Dual	-	-	-	-	356	.384 (<.001)	360	.454 (<.001)
Dual Use	All Study Groups	832	.574 (<.001)	820	.582 (<.001)	1290	.751 (<.001)	1266	.723 (<.001)
	ASPQ	378	.535 (<.001)	345	.530 (<.001)	426	.726 (<.001)	427	.719 (<.001)
	ASNPQ	454	.609 (<.001)	475	.621 (<.001)	508	.738 (<.001)	479	.620 (<.001)
	Dual	-	-	-	-	356	.528 (<.001)	360	.584 (<.001)

PHASE 2

DIF analyses were conducted with the original e-vapor product data, the oral TDN product data, and MST data. Results did not reveal substantial DIF, suggesting that the Behavioral Intention scales function similarly across these tobacco products.