# Inter-rater reliability of measures to characterize the tobacco retail environment in Mexico

Marissa G Hall, MSPH,[1] Christy Kollath-Cattano, PhD,[2] Luz Myriam Reynales-Shigematsu, PhD,[3] James F Thrasher, PhD, MA, MS.[2,3]

**Hall MG, Kollath-Cattano C, Reynales-Shigematsu LM, Thrasher JF. Inter-rater reliability of measures to characterize the tobacco retail environment in Mexico. Salud Publica Mex 2015;57:514-518.**

## Abstract

**Objective.** To evaluate the inter-rater reliability of a data collection instrument to assess the tobacco retail environment in Mexico, after major marketing regulations were implemented. **Materials and methods.** In 2013, two data collectors independently evaluated 21 stores in two census tracts, through a data collection instrument that assessed the presence of price promotions, whether single cigarettes were sold, the number of visible advertisements, the presence of signage prohibiting the sale of cigarettes to minors, and characteristics of cigarette pack displays. We evaluated the inter-rater reliability of the collected data, through the calculation of metrics such as intraclass correlation coefficient, percent agreement, Cohen's kappa and Krippendorff's alpha. **Results.** Most measures demonstrated substantial or perfect inter-rater reliability. **Conclusions.** Our results indicate the potential utility of the data collection instrument for future point-of-sale research.

Keywords: tobacco;reproducibility of results; Mexico

**Hall MG, Kollath-Cattano C, Reynales-Shigematsu LM, Thrasher JF. Confiabilidad interobservador de las medidas para evaluar la publicidad del tabaco en los puntos de venta en México. Salud Publica Mex 2015;57:514-518.**

## Resumen

**Objetivo.** Evaluar la confiabilidad interobservador de un instrumento de recolección de datos para el entorno minorista de tabaco en México, después de que se implementaran fuertes regulaciones para su comercialización. **Material y métodos.** En 2013, se evaluaron de forma independiente 21 tiendas en dos áreas geoestadísticas básicas, a través de un instrumento de recolección de datos que evalúa la presencia de promociones, si se venden cigarrillos sueltos, el número de anuncios visibles, la presencia de la señalización que prohíbe la venta de cigarrillos a menores de edad, y las características de los exhibidores de cigarrillos. Se evaluó la fiabilidad interobservador de los datos recabados, a través del cálculo de métricas como el coeficiente de correlación intraclase, porcentaje de concordancia, kappa de Cohen y alfa de Krippendorff. **Resultados.** La mayoría de las variables observadas demostraron un nivel de confiabilidad interobservador sustancial. **Conclusiones.** Los resultados indican la utilidad potencial del instrumento de recolección de datos para la investigación futura.

Palabras clave: tabaco; reproducibilidad de resultados; México

(1) Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina. USA.
(2) Department of Health Promotion, Education and Behavior, Arnold School of Public Health, University of South Carolina. USA.
(3) Tobacco Control Research Department, National Institute of Public Health. Mexico.

**Received on:** March 12, 2015 • **Accepted on:** September 29, 2015
Corresponding author: MSPH. Marissa G. Hall. Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina. osenau Hall, CB #7440 Chapel Hill, NC 27599, USA.
E-mail: mghall@unc.edu

Globally, tobacco marketing through traditional mass media channels (e.g., radio, TV) has been increasingly banned.[1] As a result, the tobacco industry uses various strategies to market their products at the point of sale (PoS), including advertisements, product displays, and price promotions. Two systematic reviews have found positive associations between exposure to PoS tobacco marketing and tobacco use.[2,3] The World Health Organization's Framework Convention on Tobacco Control requires its Parties to enact a comprehensive ban on tobacco advertising, promotion, and sponsorship, including PoS advertising,[4] but data are needed to show the extent and impact of PoS marketing.

The tobacco retail environment remains understudied in Mexico, where 16% of adults smoke cigarettes.[5] The Mexican government has taken steps to regulate the tobacco retail environment and tobacco advertising and marketing. For example, the government banned the sale of single cigarettes in 1999. Moreover, in 2000, Mexico banned tobacco advertising on billboards near schools, as well as advertising on radio, internet, and television.[6] More recently, Mexico passed the 2008 General Law of Tobacco Control, which banned PoS marketing except inside of stores only accessible to adults.[7] The law also requires that pictorial warnings on cigarette packs be visible while on display.[7] However, recent studies in Mexico have found low compliance with many of these measures.[8,9,*]

Reliable measurement of the tobacco retail environment is critical for monitoring compliance to PoS laws, and for characterizing the impact of the tobacco retail environment on smoking behavior. A recent systematic review described 88 tobacco store audit studies, finding that only 11% of studies reported the reliability of the measures used.[10] These studies[11-21] generally found adequate reliability with a few exceptions (see Appendix A of Lee and colleagues 2014).[10] All studies with reliability assessments were conducted in high-income countries, and their findings may not generalize to low- and middle-income countries, which may include different types of retail environments and promotional strategies. For example, the illegal sale of single cigarettes is common in Mexico,[8,22,23] whereas single cigarettes are less frequently sold in higher-income countries. Neither of the two Mexican studies in the review reported reliability information,[6,24] highlighting the need for additional research on the reliability of measures used in this distinct environment. Furthermore, these studies took place before many of marketing regulations were implemented in 2009. The current study aimed to evaluate the inter-rater reliability of a data collection instrument that assessed the tobacco retail environment in Mexico, in the period after major marketing regulations were implemented.

## Materials and methods

The data collection instrument and protocol were adapted from two previously-developed PoS study protocols.[25,26] In 2013, researchers developed a paper-and-pencil data collection instrument that assessed the presence of price promotions, whether single cigarettes were sold, the number of visible advertisements, the presence of signage prohibiting the sale of cigarettes to minors, and characteristics of cigarette pack displays.* The instrument also assessed the area occupied by cigarette pack displays by counting the number of rows of packs in the display and the number of packs in each row, taking into consideration the pack side that was displayed (i.e., side, top, face).

The first author developed a brief training manual describing the coding protocol.*After reviewing the training manual, two data collectors visited one census tract in Puebla and another in Guadalajara in May, 2013. Data collectors canvassed the census tract and, pretending to be customers, visited all tobacco retailers that were open for business. Data collectors independently answered questions about each store. Here, we present data from 21 stores and 35 product displays (some stores had more than one display). There were no missing data for the variables of interest.

Analyses were conducted in Stata version 13.1. We measured product display size in $cm^2$, multiplying the number of rows by the number packs in each row by the area of the pack face. We then computed the intraclass correlation coefficient, representing the proportion of the total variance that is due to differences across displays rather than between raters, which is recommend for evaluating the inter-rater reliability of continuous variables.[27-29] For categorical variables, we first computed the percent agreement, representing the number of agreements divided by the total number of observations. We also calculated Cohen's kappa ($k$), which measures the level of agreement among raters while correcting for the expected agreement that would be expected by chance (i.e., higher expected agreement will reduce kappa).[30,31]

* Rodríguez-Bolaños RA, Reynales-Shigematsu LM, Barrientos-Gutiérrez T, Santos-Luna R, Ávila-Tang E, Sáenz de Miera Juárez B. Vigilancia y monitoreo de las estrategias para el control del tabaco: Un análisis a través del Sistema de Información Geográfica. In press.

* The manual is available online at: http://www.controltabaco.mx/proyectos/confiabilidad-inter-observador

Finally, we calculated Krippendorff's alpha (α), an alternative measure of inter-rater reliability that calculates reliability by examining the frequencies of disagreements rather than agreements and also adjusts for small sample sizes.[32] We chose to calculate multiple metrics in order to provide a more comprehensive, triangulated assessment of inter-rater reliability. We were unable to compute measures of inter-rater reliability for questions that did not have variability in responses, including presence of exterior publicity, sale of items with tobacco publicity, placement of tobacco products at children's eye-level, and sale of contraband cigarettes. We adopted commonly-used thresholds for determining acceptable reliability in which values less than 0.40 indicate poor or fair agreement, 0.41 to 0.60 indicate moderate agreement, 0.61 to 0.80 indicate substantial agreement, and 0.81 to 1.00 indicate near perfect or perfect agreement.[33]

## Results

The average size of tobacco product displays was 921 cm.$^2$ The intraclass correlation coefficient for product display size was 0.99, demonstrating high reliability. In other words, 99% of the variance in size was due to the product display, rather than the rater. Most categorical variables had high inter-rater reliability (table I). There was perfect agreement on four variables, including presence of interior publicity, presence of any promotions, presence of price promotions or coupon, and sale of single cigarettes (k=1.00). Six variables demonstrated substantial agreement ($\alpha > 0.60$). Three variables (i.e., presence of anti-smoking signage sponsored by the tobacco industry, shelf location, and whether the shelf was within one meter of the register) had poor or fair reliability.

## Discussion

Most measures of tobacco marketing of PoS had adequate or high inter-rater reliability, demonstrating that they are appropriate measures for future tobacco PoS research. These measures performed similarly to point-of-sale measures tested in high-income settings.[11,21] Our innovative measure for evaluating the size of tobacco product displays performed well; reliability for this measure has not been reported previously in the peer-reviewed literature. This variable may be particularly important because the tobacco industry increasingly relies on tobacco product displays for advertisement, especially other forms of advertising are banned.[1]Exposure to tobacco product displays may increase youth smoking susceptibility[34-36] and impulse cigarette purchases.[37,38]As such, evaluating characteristics of tobacco product displays (e.g., size, location) may be particularly important for future PoS research.

Three of our measures had low reliability and may benefit from slight changes. The question about the

**Table I**

**Inter-rater reliability of variables assessing the tobacco retail environment in Mexico, 2013**

| *Variable* | *% yes, rater 1* | *% yes, rater 2* | *% agreement* | *k* | α |
|---|---|---|---|---|---|
| General (*n* stores = 21) | | | | | |
| Interior publicity: yes/no | 19 | 19 | 100 | 1.00 | -- |
| Number of ads inside the establishment | -- | -- | 75 | .56 | .70 |
| Any promotions: yes/no | 14 | 14 | 100 | 1.00 | -- |
| Price promotions or coupons: yes/no | 14 | 14 | 100 | 1.00 | -- |
| Single cigarettes sold: yes/no | 62 | 62 | 100 | 1.00 | -- |
| Anti-smoking signage sponsored by the tobacco industry: yes/no | 14 | 0 | 86 | .00 | -.05 |
| Signage prohibiting sale of cigarettes to minors: yes/no | 29 | 19 | 90 | .74 | .74 |
| Product displays (*n* displays = 35) | | | | | |
| Shelf location: Above the register/behind the register/other | -- | -- | 71 | .57 | .56 |
| Shelf within one meter of the register: yes/no | 94 | 90 | 90 | .35 | .36 |
| Shelf exclusively for cigarettes: yes/no | 81 | 84 | 97 | .89 | .87 |
| Shelf also used to display sample packs: yes/no | 58 | 58 | 87 | .74 | .73 |
| Shelf contains color or brand of cigarettes: yes/no | 58 | 61 | 90 | .80 | .80 |
| Pictorial warnings visible on the shelf: yes/no | 65 | 65 | 94 | .86 | .86 |

Note: Krippendorff's alpha cannot be calculated when percent agreement is 100%

product display location, for example, lacked a response option for whether the display was located in front of the cash register, causing some confusion among data collectors. Inter-rater reliability may also be improved by using a more detailed coding protocol and with more intensive training of data collectors. For instance, the question about industry-sponsored signage performed poorly, perhaps because the training manual lacked a photograph example of this type of sign.

Overall, our measures performed well and thus hold promise for future PoS research. Moreover, the data collection instrument was feasible to use and yielded high-quality data when collecting data covertly. Future studies should test these measures in a larger sample of retailers, and in different environments, to see if the measures continue to perform well. Reliable measurement and a well-developed protocol are crucial for accurately capturing important elements of the tobacco PoS environment.

## Acknowledgements

The authors acknowledge Luis Javier Robles-Arellano, Rosibel Rodríguez-Bolaños, Rosaura Pérez-Hernández, Rene Santos, Ana Lidia Salgado, Moisés Ortiz-Vargas, Francisco Bernabé-Jiménez and Patrick Nerz for their assistance with study development and data collection.

*Declaration of conflict of interests.* The authors declare that they have no conflict of interests.

## References

1. Henriksen L. Comprehensive tobacco marketing restrictions: promotion, packaging, price and place. Tob Control 2012;21(2):147-153.
2. Paynter J, Edwards R. The impact of tobacco promotion at the point of sale: a systematic review. Nicotine Tob Res 2009;11(1):25-35.
3. Robertson L, McGee R, Marsh L, Hoek J. A Systematic Review on the Impact of Point-of-Sale Tobacco Promotion on Smoking. Nicotine Tob Res 2015;17(1):2-17. PubMed PMD: 25173775. Epub 2014/09/01.
4. World Health Organization. Framework Convention on Tobacco Control 2013 [cited 2013 April 20]. Available from: http://www.who.int/fctc/en/.
5. Global Adult Tobacco Survey. Mexico country report. Washington, DC: World Health Organization, 2009.
6. Hernández-Ávila JE, Tirado-Ramírez E, Santos-Luna R, Reynales-Shigematsu LM. Use of Geographical Information Systems for billboards and points-of-sale surveillance in two Mexico cities. Salud Publica Mex 2007;49:s241-s246.
7. Ley general para el control de tabaco, México: Diario Oficial de la Federación 06-01-2010, 2008.
8. Hall MG, Fleischer NL, Reynales-Shigematsu LM, Arillo-Santillan E, Thrasher JF. Increasing availability and consumption of single cigarettes: trends and implications for smoking cessation from the ITC Mexico Survey. Tob Control 2015;24 Suppl 3:iii64-iii70. PubMed 25192772. Pubmed Central PMCID: PMC 4465400. Epub 2014/09/07.
9. Perez-Hernandez R, Thrasher JF, Rodriguez-Bolanos R, Barrientos-Gutierrez I, Ibanez-Hernandez NA. Autorreporte de exposicion a publicidad y promocion de tabaco en una cohorte de fumadores mexicanos. Salud Publica Mex 2012;54(3):204-212.
10. Lee JG, Henriksen L, Myers AE, Dauphinee AL, Ribisl KM. A systematic review of store audit methods for assessing tobacco marketing and products at the point of sale. Tob Control 2014;23(2):98-106.
11. Shareck M, Dassa C, Frohlich KL. Improving the measurement of neighbourhood characteristics through systematic observation: inequalities in smoking as a case study. Health place 2012;18(3):671-682.
12. Seidenberg AB, Caughey RW, Rees VW, Connolly GN. Storefront cigarette advertising differs by community demographic profile. Am J Health Promot 2010 24(6):e26-31.
13. Bernstein SL, Cabral L, Maantay J, Peprah D, Lounsbury D, Maroko A, *et al*. Disparities in access to over-the-counter nicotine replacement products in New York City pharmacies. Am J Public Health 2009;99(9):1699-1704.
14. Devlin E, Anderson S, Borland R, MacKintosh AM, Hastings G. Development of a research tool to monitor point of sale promotions. Social Marketing Quarterly 2006;12(1):29-39.
15. Feighery EC, Ribisl KM, Schleicher N, Lee RE, Halvorson S. Cigarette advertising and promotional strategies in retail outlets: results of a statewide survey in California. Tob Control 2001;10(2):184-188.
16. Henriksen L, Feighery EC, Schleicher NC, Haladjian HH, Fortmann SP. Reaching youth at the point of sale: cigarette marketing is more prevalent in stores where adolescents shop frequently. Tob Control 2004;13(3):315-318.
17. Wakefield MA, Terry-McElrath YM, Chaloupka FJ, Barker DC, Slater SJ, Clark PI, *et al*. Tobacco industry marketing at point of purchase after the 1998 MSA billboard advertising ban. Am J Public Health 2002;92(6):937-940.
18. Hosler AS, Dharssi A. Reliability of a survey tool for measuring consumer nutrition environment in urban food stores. J Public Health Manag Pract 2011;17(5):E1-8.
19. Hosler AS, Kammer JR. Point-of-purchase tobacco access and advertisement in food stores. Tob Control 2012;21(4):451-452.
20. Jason LA, Pokorny SB, Mikulski K, Schoeny ME. Assessing storefront tobacco advertising after the billboard ban. Eval Health Prof 2004;27(1):22-33.
21. Quedley M, Ng B, Sapre N, Blakiston M, Crawford A, Devadas R, *et al*. In sight, in mind: retailer compliance with legislation on limiting retail tobacco displays. Nicotine Tob Res 2008;10(8):1347-1354.
22. Thrasher JF, Villalobos V, Barnoya J, Sansores R, O'Connor R. Consumption of single cigarettes and quitting behavior: a longitudinal analysis of Mexican smokers. BMC Public Health 2011;11:134.
23. Thrasher JF, Villalobos V, Dorantes-Alonso A, Arillo-Santillan E, Cummings KM, O'Connor R, *et al*. Does the availability of single cigarettes promote or inhibit cigarette consumption? Perceptions, prevalence and correlates of single cigarette use among adult Mexican smokers. Tob Control 2009;18(6):431-437.
24. Rodríguez-Bolaños R, Reynales-Shigematsu LM, Ibañez-Hernández NA, Santos-Luna R, Valdés-Salgado R, Ávila-Tang E, *et al.* Estrategia de vigilancia para el control del tabaco en México: publicidad, promoción y patrocinio, empaque y etiquetado. Salud Publica Mex 2010;52:S254-S266.
25. Tobacco Control Tracking Tools. Baltimore (MD): Johns Hopkins Bloomberg School of Public Health, 2005.
26. Johns Hopkins Bloomberg School of Public Health, International Union Against Tuberculosis and Lung Disease, Campaign for Tobacco-Free Kids. Assessing Compliance with Tobacco Advertising, Promotion, and Sponsorship (TAPS) Bans. Baltimore(MD): Johns Hopkins Bloomberg School of Public Health, 2013.
27. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(2):420.
28. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973;33:613-619.

29. Hallgren KA. Computing inter-rater reliability for observational data: An overview and tutorial. Tutorials in quantitative methods for psychology. 2012;8(1):23.
30. Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull 1968;70(4):213.
31. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20(1):37-46.
32. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Commun Methods Meas 2007;1(1):77-89.
33. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977:159-74.
34. Spanopoulos D, Britton J, McNeill A, Ratschen E, Szatkowski L. Tobacco display and brand communication at the point of sale: implications for adolescent smoking behaviour. Tob Control 2014;23(1):64-69.
35. Scheffels J, Lavik R. Out of sight, out of mind? Removal of point-of-sale tobacco displays in Norway. Tob Control 2013;22(e1):e37-42.
36. Paynter J, Edwards R, Schluter PJ, McDuff I. Point of sale tobacco displays and smoking among 14-15 year olds in New Zealand: a cross-sectional study. Tob Control 2009;18(4):268-274.
37. Li L, Borland R, Fong GT, Thrasher JF, Hammond D, Cummings KM. Impact of point-of-sale tobacco display bans: findings from the International Tobacco Control Four Country Survey. Health Educ Res 2013;28(5):898-910.
38. Hoek J, Gifford H, Pirikahu G, Thomson G, Edwards R. How do tobacco retail displays affect cessation attempts? Findings from a qualitative study. Tob Control 2010;19(4):334-337.