

# Item reduction of the Wisconsin Upper Respiratory Symptom Survey (WURSS-21) leads to the WURSS-11

Chidi N. Obasi · Roger L. Brown · Bruce P. Barrett

Accepted: 10 October 2013  
© Springer Science+Business Media Dordrecht 2013

## Abstract

**Purpose** To develop a shorter version of the Wisconsin Upper Respiratory Symptom Survey (WURSS-21), a self-report questionnaire for evaluating daily symptoms and functional impairments during acute respiratory illness (ARI).

**Methods** WURSS-21 data were retrieved from 4 studies ( $n = 1167$ ) spanning the years 2002–2010. Similar methodologies were employed among these studies. Degree of missingness, ceiling/floor effects, and exploratory (EFA) and confirmatory (CFA) factor analyses were investigated and used to guide item retention. Stability of the reduced WURSS was evaluated across the first 3 days of ARI.

**Results** Degree of missingness was  $<1\%$  and appeared to be completely at random. Seven WURSS items with  $>30\%$  of ratings of zero (floor effects) were eliminated. Cross-loading items (head congestion, sleep well and breathe easily) were excluded following EFA on subset-1. Subsequent CFA using subset-2 showed satisfactory indices of fit. The reduced WURSS-11 instrument demonstrated 3 dimensions of 3 items each and was stable across 3 days of illness. The indicated dimensions (items) include nasal (runny nose, plugged nose, sneezing), throat (cough,

sore throat, scratchy throat), and quality of life (feeling tired, think clearly, accomplish daily activities).

**Conclusion** The WURSS-11 has similar dimensional structure as the WURSS-21. This shorter version may reduce the time and burden required for completing the survey.

**Keywords** Common cold · Quality of life · Factor analysis · Acute respiratory infection · Symptom assessment

## Introduction

A shorter yet valid version of the Wisconsin Upper Respiratory Symptom Survey (WURSS) is desirable for reducing the burden involved in filling out the questionnaire frequently during a time of illness. The current WURSS-21 has been shown to be a reliable and responsive research tool useful in evaluating the severity of acute respiratory illness (ARI) [1].

It was developed from the WURSS-44 using item-domain reliability, responsiveness, and importance to participants as main criteria [2, 3]. This work suggested that general symptoms and quality-of-life functions were valued more highly than specific symptoms [1, 4]. These original WURSS versions showed 3 factors for WURSS-21 and 8 factors for WURSS-44. Both versions have been validated [4] against surveys and laboratory-assessed biomarkers [5, 6].

At least one version of the WURSS instrument has been utilized by investigators in 200 institutions in 40 countries and translated into 7 languages (<http://www.fammed.wisc.edu/research/external-funded/wurss>). With the increasing interest and use of the WURSS, the aim of this study was to

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s11136-013-0561-z) contains supplementary material, which is available to authorized users.

---

C. N. Obasi (✉) · B. P. Barrett  
Department of Family Medicine, University of Wisconsin-Madison, 1100 Delaplaine Ct., Madison, WI 53715, USA  
e-mail: cobasi@wisc.edu

R. L. Brown  
Design and Statistics Unit, Schools of Nursing, Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA

derive a shorter version which could reduce survey completion time and perhaps increase response rates.

## Method

### Data sources

Data for this work came from 4 previous studies generating WURSS-21 information including (1) WURSS-21 development [4] and (2) WURSS-21 validation [1] which spanned the years 2002–2007. Additional data came from 2 clinical trials titled (3) “Placebo: Physician or Pill? RCT in a Common Cold Model” (2004–2008) [7, 8] and (4) “Meditation or Exercise for Preventing Acute Respiratory Infections” (2009–2010) [9].

Briefly,  $n = 1167$  people with ARI self-reported daily information on the WURSS-21, if they had symptoms of <48 h, a score of >2 on the Jackson scale, and felt they were having a cold. Two consecutive daily responses with “Yes” to the question “Do you think you have a cold?” and “No” to “Do you think you are still sick?” respectively confirmed the start and end of the illness episode.

### Statistical analysis

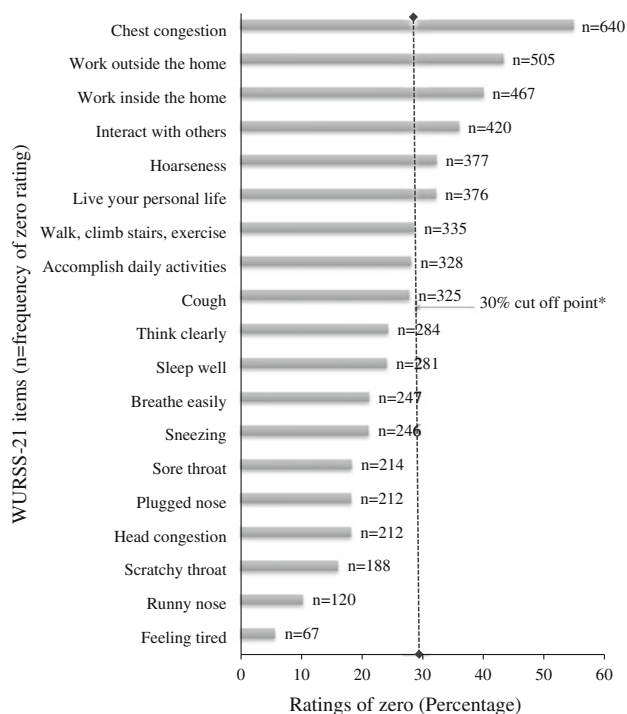
The WURSS introductory (“How sick do you feel today?”) and concluding items (“Compared to yesterday...”) were excluded from this analysis because they are designed to measure different time frames and are usually analyzed separately. Analysis was restricted to the first 3 days of illness to maximize sample size and because ARI severity tends to peak within this time [10].

Following the Kroonenberg and Lewis approach [11], we randomly divided the combined dataset into 2 equal sized subsets. Approach-1 then employed exploratory factor analysis (EFA) on subset-1, while approach-2 tested these results using confirmatory factor analysis (CFA) on subset-2. Subsequently, the combined subsets were used to obtain final parameter estimates of the reduced WURSS model. Model invariance was assessed across the first 3 days of illness [12].

Approach-1 used EFA to identify the underlying dimensional/factor structure toward a reduced number of WURSS-21 items with significant loadings. An oblique solution accounted for correlations between the factors while weighted least squares with mean and variance adjustment (WLSMV) was used to model the ordered categorical nature of the WURSS items. Several models involving multiple dimensions were evaluated and the best dimensional model was selected. Model choice was based on dimensional cohesion and satisfactory applications of factor analysis [13].

Approach-2 employed CFA to examine the model fit and dimensional structure of the selected items. Assessing the measurement model validity occurs when the theoretical measurement model is compared with the reality model to see how well the data fits. Chi-square testing of goodness of fit statistics including comparative fit index (CFI), Tucker–Lewis index (TLI), and root mean square error of approximation (RMSEA) were used to check measurement model validity [14–16]. [17]

Longitudinal measurement invariance was conducted using multiple CFA models with WLSMV. Following procedures suggested by Vandenberg and Lance, progressively more stringent nested models were constructed and evaluated [18]. All models were constructed in Mplus version 7.11 [19]. Model 1 consisted of two correlated common factors corresponding to two time periods (day 1 and day 3) with correlated residual variances for the same items over the time periods. Model 2 was the same as model 1 but constrained the loadings to be the same or invariant over time. Model 3 extended model 2 by constraining the factor variances to be invariant over time. The DIFFTEST option in Mplus was used to obtain a correct chi-square difference test when the WLSMV estimators are



**Fig. 1** Frequency ratings of zero on WURSS-21 during acute respiratory illness. \*Any WURSS item with > 30 % its total count ( $n = 1167$ ) indicating ratings of zero were not retained for further investigation. These 6 items excluded from further investigation because they demonstrated substantive floor effects were live your personal life ( $n = 376$ ), hoarseness ( $n = 377$ ), interact with others ( $n = 420$ ), work inside the home ( $n = 467$ ), work outside the home ( $n = 505$ ), and chest congestion ( $n = 640$ )

used because the difference in chi-square values for two nested models using the WLSMV chi-square values is not distributed as chi square [20].

## Results

Of the  $n = 1167$  participants included in this analysis, most were female (66 %), non-smokers (65 %), and college graduates (49 %), with a mean age of 35 years ( $SD = 15$ ). Randomization (without replacement) yielded  $n = 584$  participants in subset-1 and  $n = 583$  in subset-2. Results showed a decrease in number of observations with each consecutive increase in days of illness, due primarily to early recovery. Data from illness day 1 ( $n = 1167$ ) was used for most analyses, as it contained more observations than day 2 ( $n = 1155$ ) or day 3 ( $n = 1152$ ).

The proportion of missing data was less than 1 %, with most from head congestion 0.9 % ( $n = 11$ ) and least from cough 0.2 % ( $n = 2$ ). Little's missing completely at random test failed to reject the null hypothesis that no identifiable pattern exists to the missing data [21]. The combined random pattern of missing values and estimated

low degrees of missingness may reflect ease of WURSS use among participants.

Frequency of rating items as "0" (no symptom or no interference with daily function) varied among the WURSS items, with least for feeling tired (6 %) and most for chest congestion (55 %). Six items with  $>30\%$  of their responses having "0" rating were excluded from further investigation because they demonstrated substantive floor effects and therefore may be less able to assess severity of ARI over time (Fig. 1).

To minimize redundancy, live your personal life, interact with others, work outside, or inside the home, and walk climb stairs exercise were excluded, while accomplish daily activity was retained because it showed the highest correlation with most of these items and may reasonably encompass their individual meanings.

### Approach-1

Of the 19 WURSS items examined, EFA was performed on 12 while 7 were eliminated based on  $>30\%$  floor effects and redundancy. The EFA indicated 3 clinically meaningful dimensions/factors: 1-head/nasal; 2-chest/throat; and

**Table 1** Approach-1 (exploratory factor analysis)

WURSS items	Dimension/factor (F)		
	F1-loading (nasal) (standard error)	F2-loading (throat) (standard error)	F3-loading (quality of life) (standard error)
Runny nose	<b>0.76 (0.025)</b>	0.030 (0.044)	-0.0090 (0.020)
Plugged nose	<b>0.71 (0.036)</b>	-0.040 (0.030)	0.16 (0.054)
Sneezing	<b>0.70 (0.035)</b>	0.043 (0.044)	-0.021 (0.038)
Sore throat	-0.030 (0.017)	<b>0.64 (0.063)</b>	0.15 (0.055)
Scratchy throat	0.057 (0.072)	<b>1.035 (0.071)</b>	-0.003 (0.0040)
Cough	0.18 (0.054)	<b>0.25 (0.046)</b>	0.28 (0.055)
<i>Head congestion</i>	<i>0.46 (0.040)</i>	<i>0.040 (0.030)</i>	<i>0.40 (0.046)</i>
Feeling tired	-0.030 (0.041)	0.064 (0.043)	<b>0.74 (0.037)</b>
Think clearly	-0.0040 (0.010)	-0.025 (0.027)	<b>0.83 (0.022)</b>
<i>Sleep well</i>	<i>0.33 (0.045)</i>	<i>0.048 (0.033)</i>	<i>0.52 (0.043)</i>
<i>Breathe easily</i>	<i>0.52 (0.039)</i>	<i>-0.043 (0.029)</i>	<i>0.45 (0.047)</i>
Accomplish daily activities	0.040 (0.045)	-0.018 (0.024)	<b>0.85 (0.035)</b>

ARI day 1 data were explored using subset-1. These are unstandardized loadings

WURSS-21 excluding 2 items "How sick do you feel today?" and "Compared to yesterday, I feel...?"

Utilized applications include observed break point on scree plot, number of eigenvalues  $>1$ , cumulative percent of variance explained by model, factor loadings  $\geq 0.30$ , minimal number of cross-loading items, and each factor having  $\geq 3$  WURSS items[13]

Loading estimates in bold indicate domains where WURSS items belong

Italics indicate cross-loaders with loadings  $>0.3$ , to be excluded. Three significant cross-loading items (head congestion, breathe easily and sleep well) were excluded

Nine items were retained for confirmatory analysis: nasal (runny nose, plugged nose and sneezing), throat (sore throat, scratchy throat and cough), or QOL dimension (feeling tired, think clearly and accomplish daily activities)

3-quality of life (QOL)/functional impairment (3-factor model) with reasonably strong loadings of the WURSS items. The item assessing cough showed a somewhat weaker loading estimate but was retained because it is known to occur frequently during ARI, with substantive patient-oriented impact. Three significant cross-loading items were excluded while the remaining 9 items were retained for confirmatory analysis (Table 1).

#### Approach-2

CFA using the retained 9 items showed significant reliability and dimensional cohesion. Exclusion or inclusion of cough resulted in satisfactory indices of model fit with good coefficient estimates. The model excluding the cough item showed slightly better fit (CFI = 0.97, TLI = 0.96, RMSEA = 0.09) compared to the model including cough (CFI = 0.93, TLI = 0.89, RMSEA = 0.13). However, cough was retained because of the widespread recognition that cough is an important symptomatic and quality-of-life component of ARI.

Following satisfactory EFA–CFA, both subsets were combined and the item–dimension structure showed high internal consistency (Cronbach’s  $\alpha$ ) and adequate convergent validity (composite reliability and average-variance-extracted-AVE, Table 2). Finally, the 3-factor model with retained 9 items explained a significant proportion of variance in the original WURSS-21 ( $R^2 = 0.92$ ). It also showed significant factor correlations (Fig. 2). Discriminant validity was satisfactory because construct’s AVE was greater than the shared variance [22].

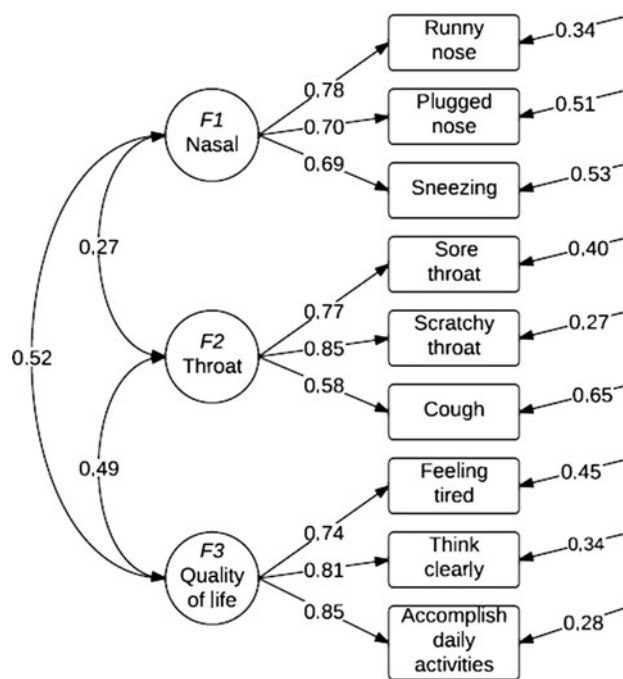
Invariance was based on non-significant chi-square difference tests ( $p > 0.05$ ), and minimal change in the goodness-of-fit indices of CFI and TLI ( $\leq 0.01$ ) between models [23]. Results from Table 3 indicate that the WURSS demonstrated evidence of longitudinal scalar invariance for factors 1 and 3, but not for factor variances. The reason factor 2 failed the invariance assessment was due to the incorporation of a fairly weak item (i.e., cough based on clinical rather than statistical justification).

**Table 2** Estimates of reliability and validity of the Wisconsin Upper Respiratory Symptom Survey (WURSS-11)

Dimensions	Nasal	Throat	Quality of life
Cronbach’s $\alpha$	0.74	0.71	0.81
Composite reliability	0.77	0.78	0.84
Average variance extracted	0.53	0.55	0.64

Adequate convergent validity = Composite reliability > average variance extracted or Average variance extracted > 0.5

Discriminant validity was satisfied because the average variances extracted for nasal (0.53), throat (0.55), and QOL (0.64) were larger than their shared variances (square of the factor correlation) [22]



**Fig. 2** Three-dimensional factor model of Wisconsin Upper Respiratory Symptom Survey (WURSS-11)\*. \*Standardized values. The Three dimensions are nasal (runny nose, plugged nose, sneezing), throat (cough, sore and scratchy throat), and QOL (feeling tired, think clearly, accomplish daily activities). Introductory “How sick do you feel today?” and concluding item “Compared to yesterday, I feel...?” which measure different time frames were not included during the analysis but will be added to the retained 9 items to form the WURSS-11

#### Discussion

This study demonstrates that the WURSS-21 can be reduced to WURSS-11, preserving reliability and domain structure. The introductory and concluding items will be added to the retained 9 items to form the WURSS-11.

Among the retained items, QOL measures appeared to perform somewhat better, with higher parameter estimates, compared to symptom severity. This supports previous studies which showed that QOL items perform slightly better and are more important to patients than specific symptoms [1, 4, 24].

The reduced WURSS-11 demonstrates satisfactory indices of fit with good reliability estimates. Most importantly, it retains and reflects similar dimensional construct as the parent WURSS-21. While not yet assessed, WURSS-11’s retention of roughly half of the items from the parent survey may suggest a reduction in WURSS completion time by as much as 50 %.

These results show stability of WURSS-11 across the first 3 days of ARI. It also suggests WURSS-11 retains content validity while construct validity is likely to be satisfactory, based on previous findings of significant convergent and discriminant validities [1, 4, 6]. However,

**Table 3** Longitudinal invariance of the WURSS over 3-day period

Factors	Models	$\chi^2$	df	p value	CFI	TLI	RMSEA	WRMR
Factor 1	Model 1	2.533	5	0.772	1.00	1.00	0.000	0.143
	Model 2	5.020	7	0.658	1.00	1.00	0.000	0.210
	Model 3	23.108	8	0.003	0.997	0.994	0.040	0.534
Factor 2	Model 1	21.779	5	<0.001	0.998	0.994	0.054	0.439
	Model 2	51.097	7	<0.001	0.995	0.988	0.073	0.722
	Model 3	137.23	8	<0.001	0.984	0.970	0.118	1.247
Factor 3	Model 1	5.527	5	0.354	1.00	1.00	0.010	0.190
	Model 2	8.587	7	0.283	1.00	1.00	0.014	0.261
	Model 3	73.833	8	<0.001	0.993	0.987	0.084	0.949
Factors	Model comparison <sup>a</sup>	$\chi^2_{diff}$	df <sub>diff</sub>	p value	$\Delta$ CFI	$\Delta$ TLI		
<i>Comparisons of the WURSS over 3-day period</i>								
Factor 1	Model 1 versus 2	3.017	2	0.22	0.000	0.000		
	Model 2 versus 3	9.730	1	0.01	0.003	0.006		
Factor 2	Model 1 versus 2	26.512	2	<0.001	0.003	0.006		
	Model 2 versus 3	65.976	1	<0.001	0.011	0.018		
Factor 3	Model 1 versus 2	3.490	2	0.17	0.000	0.000		
	Model 2 versus 3	33.013	1	<0.001	0.007	0.013		

$\chi^2$  = chi square; df = degrees of freedom; root mean square error of approximation (RMSEA <0.06 acceptable fit); comparative fit index (CFI >0.95 acceptable fit), or Tucker–Lewis index (TLI >0.95 acceptable fit); weighted root mean square residual (WRMR)

<sup>a</sup> Chi-square difference test based on Satorra and Bentler correction [17, 20]

prospective studies are needed to compare the performance of the WURSS-11 to its parent WURSS-21 and external comparators such as general and illness-specific self-report health measures, and viral nucleic acids.

Study limitations include limited diversity in the study population and lack of external replicated validation. However, generalizability is strengthened by the large sample size ( $n = 1167$ ), assorted age groups, large time span (2002–2010), different seasons of ARI with incidental viral microbes, and most importantly, the similarities between the inclusion and exclusion criteria among all 4 WURSS studies.

In conclusion, this study has successfully created the WURSS-11, which is a shorter version of the current WURSS-21. It suggests that the WURSS-11 has similar dimensional structure as the parent WURSS-21 and may be as reliable and responsive as the longer versions of WURSS. This shortened self-reporting survey may reduce WURSS completion time and increase response rate, but will need additional validation in future studies.

**Acknowledgments** Data analyzed here come from several separate studies supported by: (1) Patient-Oriented Career Development Grant (K23 AT00051-01) from the National Center for Complementary and Alternative Medicine (NCCAM) at the National Institutes of Health; (2) Career development grant from the Robert Wood Johnson Foundation Generalist Physician Scholars Program used for project analysis; (3) Clinical Research Feasibility Funds (CREFF) award from the NIH-funded University of Wisconsin-General Clinical Research Center (MO1 RR03186); (4) 1R01AT001428 Placebo: Physician or Pill? A randomized trial in a common cold model (NIH NCCAM); (5) 1R01AT004313 Meditation and Exercise for Prevention of Acute Respiratory Infection (NIH NCCAM); (6) K24AT006543 Midcareer Investigator Award (NIH NCCAM). Additional support include (7) National Research Service Award (T32HP10010) from the Health Resources and Services Administration; and (8) Grant 1UL1RR025011 from the Clinical and Translational Sciences Award program of the National Center for Research Resources, National Institutes of Health.

**Conflict of interest** The WURSS-11 will be free of charge for educational, non-profit, and public interest work, but will be licensed for commercial or for-profit use by the Wisconsin Alumni Research Foundation (WARF). Chidi Obasi, Roger Brown, and Bruce Barrett have transferred copyright of the WURSS-11 to WARF and may receive author share royalties from such commercial licensing, but have no other interest conflicts. See <http://www.fammed.wisc.edu/>

[research/external-funded/wurss](#) for registration and licensing information. All authors contributed significantly and approved the final version of this manuscript.

## References

- Barrett, B., Brown, R. L., Mundt, M. P., et al. (2009). Validation of a short form Wisconsin Upper Respiratory Symptom Survey (WURSS-21). *Health and Quality of Life Outcomes*, 7, 76.
- Barrett, B. P., Brown, R. L., Locken, K., et al. (2002). Treatment of the common cold with unrefined echinacea. A randomized, double-blind, placebo-controlled trial. *Annals of Internal Medicine*, 137(12), 939–946.
- Barrett, B., Locken, K., Maberry, R., et al. (2002). The Wisconsin Upper Respiratory Symptom Survey (WURSS): A new research instrument for assessing the common cold. *Journal of Family Practice*, 51(3), 265.
- Barrett, B., Brown, R., Mundt, M., et al. (2005). The Wisconsin Upper Respiratory Symptom Survey is responsive, reliable, and valid. *Journal of Clinical Epidemiology*, 58(6), 609–617.
- Jackson, G. G., Dowling, H. F., & Muldoon, R. L. (1962). Present concepts of the common cold. *American Journal of Public Health*, 52, 940–945.
- Barrett, B., Brown, R., Volland, R., et al. (2006). Relations among questionnaire and laboratory measures of rhinovirus infection. *European Respiratory Journal*, 28(2), 358–363.
- Barrett, B., Rakel, D., Chewning, B., et al. (2007). Rationale and methods for a trial assessing placebo, echinacea, and doctor–patient interaction in the common cold. *Explore (New York, N.Y.)*, 3(6), 561–572.
- Barrett, B., Brown, R., Rakel, D., et al. (2011). Placebo effects and the common cold: A randomized controlled trial. *Annals of Family Medicine*, 9(4), 312–322.
- Barrett, B., Hayney, M. S., Muller, D., et al. (2012). Meditation or exercise for preventing acute respiratory infection: A randomized controlled trial. *Annals of Family Medicine*, 10(4), 337–346.
- Heikkinen, T., & Jarvinen, A. (2003). The common cold. *Lancet*, 361(9351), 51–59.
- Kroonenberg, P. M., & Lewis, C. (1982). Methodological issues in the search for a factor model: Exploration through confirmation. *Journal of Educational Statistics*, 7, 69–89.
- Cudeck, R., & Browne, M. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, 18, 147–167.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, 10, 7. <http://pareonline.net/getvn.asp?v=10&n=7>. 2005.
- Kline, R. B. (2010). *Principles and practice of structural equation modeling*. NY: Guilford Press.
- Jöreskog, K. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183–202.
- Holtzman, S., & Vezzu, S. (2011). Confirmatory factor analysis and structural equation modeling of noncognitive assessments using PROC CALIS. In *NorthEast SAS Users Group (NESUG), 2011 proceedings*. September 11–14, 2011, Portland, Maine.
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4–69.
- Muthén, L., & Muthén, B. (1998–2013). *Mplus user's guide* (6th edn). Los Angeles, CA: Muthén & Muthén.
- Bryant, F. B., & Satorra, A. (2012). Principles and practice of scaled difference Chi square testing. *Journal of Structural Equation Modeling*, 19(3), 372–398.
- Little, R. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Journal of Structural Equation Modeling*, 14(3), 464–504.
- Obasi, C. N., Brown, R., Ewers, T., et al. (2012). Advantage of meditation over exercise in reducing cold and flu illness is related to improved function and quality of life. *Influenza and Other Respiratory Viruses*. doi:10.1111/irv.12053.