Four Dietary Items of the School Physical Activity and Nutrition (SPAN) Questionnaire Form a Robust Latent Variable Measuring Healthy Eating Patterns

Andrew L. Larsen, MA[1]; John J. McArdle, PhD[1]; Trina Robertson, RD, MS[2]; Genevieve Dunton, PhD[3]

ABSTRACT

Objective: Evaluate the factor structure and stability of 4 dietary items (fruit, fruit juice, vegetables, and milk) from the School Physical Activity and Nutrition questionnaire–elementary school version.

Methods: Secondary analysis of intervention data from third graders measured at pre-intervention, post-intervention (10 weeks), and 3-month follow-up. The researchers conducted structural equation modeling invariance analysis to test the stability of the factor structure of the 4 items.

Results: Data from 1,147 students. Fit indices revealed good fit for a single factor remaining stable across time (χ^2/degrees of freedom [DF] = 59.75/59, $P = .45$), gender (χ^2/DF = 149.72/128, $P = .09$), and study groups (χ^2/DF = 143.04/128, $P = .17$).

Conclusions and Implications: A healthy food factor consisting of the 4 items can be used in future data analysis. This offers several advantages in analysis, including the use of latent change scores that are more powerful, more informative, and more easily interpreted than traditional approaches.

Key Words: child, diet, nutrition assessment (*J Nutr Educ Behav*. 2015;47:253-258.)

Accepted December 18, 2014. Published online January 30, 2015.

Childhood obesity has risen substantially during the past few decades.[1] Monitoring children's health behaviors can help to determine the risk factors and causes of weight gain.[2] This is particularly important for diet because dietary patterns often extend into adolescence and adulthood and dietary intake has been linked to childhood obesity.[3,4] Various methods exist for measuring diet, such as 24-hour dietary recall, food record methods, and direct observation; however, they are costly and can be burdensome for children.[5,6] Child-administered measures of recalled food intake frequency such as the Block Health Habits and History Questionnaire have also been used in research,[7] but this instrument is time-consuming to complete in a classroom setting. The School Physical Activity and Nutrition (SPAN) questionnaire is a brief self-administered measure of recalled dietary intake over the previous day, used to evaluate school-based nutrition education programs.[8-10] Three versions of the questionnaire exist: high school (11th-grade level), middle school (eighth-grade level), and elementary school (fourth-grade level). The questionnaires cover the same content but are matched to age groups via reading level and cognitive demand. For example, the elementary school version contains fewer items and includes pictures of foods to facilitate recall and identification of items by children.[9]

However, the factor structure of the SPAN questionnaire has not been tested. Items that form a robust factor can be combined as a latent variable. A latent variable is an unobserved variable calculated from the interrelationships of measured variables. Creating latent variables in structural equation modeling (SEM) offers numerous advantages over traditional statistical methods.[11,12] For example, (1) SEM is a more theoretically sound method for testing mean differences than ANOVA or MANOVA, (2) latent variables statistically separate true score variance from error variance, (3) SEM allows researchers to specify and test specific a priori hypotheses and relationships beyond mean differences (eg, variance, covariance), and (4) SEM offers a more comprehensive analysis of relationships and changes than traditional methods (eg, ANOVA).[12] In practical terms, latent variables in SEM provide more power for statistical tests, give researchers more flexibility in analysis, and create clearer interpretations of results.

However, steps should be taken to ensure a latent variable is measured

[1]Department of Psychology, University of Southern California, Los Angeles, CA
[2]Dairy Council of California, Sacramento, CA
[3]Department of Preventive Medicine, University of Southern California, Los Angeles, CA
Conflict of Interest Disclosure: The authors' conflict of interest disclosures can be found online with this article on www.jneb.org.
Address for correspondence: Andrew L. Larsen, MA, Department of Psychology, University of Southern California, 3620 S McClintock Ave, SGM 501, Los Angeles, CA 90089-1061; Phone: (734) 395-6624; E-mail: allarsen@usc.edu

http://dx.doi.org/10.1016/j.jneb.2014.12.005

correctly before it can be used in analysis. Primarily, the measurement of the latent variable should remain stable (referred to as measurement invariance), meaning that relationships among observed variables forming the latent variable (ie, the loadings and error of each observed variable) should remain stable across contexts and subgroups and over time.[12,13] Simply measuring the same variables across time does not insure that the relationships among them will remain stable.[9] If the relationships change between groups or over time, scientific inferences become inaccurate because the same latent variable is not being measured across contexts.[13]

The purpose of the current research was to test for the existence (and stability) of a healthy foods latent variable among 4 food consumption items from the SPAN questionnaire (fruit, vegetables, milk, and 100% fruit juice). The stability of a latent variable can be tested by systematically holding certain model parameters to be invariant over time and over groups in SEM, and testing model fit and change in fit after each subsequent constraint. The first test was to see whether calculating a latent variable is a better representation of the data than simply summing the items. The stability of the healthy foods latent variable was then tested by constraining factor loadings, variable means, and error terms to be equal across 3 time points, and then again across the subgroups of gender and study groups (ie, intervention vs control groups). It is hypothesized that the 4 dietary items specified form a stable latent variable that can be used by researchers in future analysis.

METHODS

This project was a secondary analysis of data from a previously published report on a school-based nutrition intervention in third-grade classrooms.[14] The study evaluated the *Shaping Up My Choices* program developed by the Dairy Council of California and tested during the 2010–2011 school year. The study had a sample size of 1,147 students (675 complete cases) and featured a randomized controlled design with pre-assessment (0 weeks), post-assessment (10 weeks), and 3-month follow-up (~22 weeks) of variables such as nutrition knowledge, social-cognitive predictors, and dietary intake. The sample had a higher proportion of Asian students (19% vs 9%), a lower proportion of Hispanic students (35% vs 51%), and similar proportions of white/Caucasian (23% vs 27%) and African American students (9% vs 8%) compared with third graders attending public elementary schools across the state of California. The Ethics Committee of the Institutional Review Board at Independent Review Consulting, Inc approved the study. The researchers of the original study obtained written informed assent from all minors and used a passive parental consent procedure (ie, if the parent did not decline consent, the minor was approached for consent). More detail of the program can be found in the previously published report.[14]

The SPAN questionnaire was used to measure dietary intake. Trained research staff employed a standardized scripted procedure to administer the questionnaire. Staff members read each item out loud, pausing between items and answering students' questions along the way, and instructed children how to respond to the items (eg, *If you ate a burrito or taco, then you have eaten a tortilla*). Only the fruit, vegetables, 100% fruit juice, and milk items were included in the current analyses reported here. The items were as follows:

Yesterday, how many times did you eat fruit? Do not count 100% fruit juice.

Yesterday, how many times did you eat vegetables? Include all cooked and uncooked vegetables and potatoes. Do not count french fries or chips.

Yesterday, how many times did you drink fruit juice? Fruit juice is a 100% juice drink like orange juice, apple juice, or grape juice. Do not count punch, Kool-Aid, sports drinks, and other fruit flavored drinks.

Yesterday, how many times did you drink any kind of milk? Include chocolate or other flavored milk and drinks made with milk.

Answer options were *none*, *1 time*, *2 times*, or *≥ 3 times*. The questionnaire has been tested and found to have adequate reliability and validity in fourth graders, with Spearman correlations of 0.82 (milk), 0.79 (vegetables), 0.78 (fruit), and 0.69 (fruit juice) for 2-hour test-retest reliability and correlations of 0.56 (milk), 0.34 (vegetables), 0.40 (fruit), and 0.47 (fruit juice) with a 24-hour dietary recall that tested criterion-related validity.[8-10]

The majority of data analyses were done in R (R version 2.15.2, R Core Team, Vienna, Austria, 2012) using the Lavaan package (Lavaan version 0.5-10, Rosseel Y, 2012). The effect of classroom clustering was tested using Stata (StataSE release 12, StataCorp LP, College Station, TX, 2011). It exhibited a negligible effect (<5% of the total variance) and is not discussed further. The analysis tested for a stable factor structure (ie, measurement invariance) for a healthy foods latent variable by analyzing adequacy of model fit after progressively constraining the factor loadings, means, and errors of the observed variables over time and across subgroups of participants.[13] The logic was that if the model continues to fit the data well after adding constraints, the latent variable would be stable across the constrained parameters. It can be concluded that a stable latent variable exists and can be used in future analysis if the model continues to fit the data well after adding constraints over time and adding constraints across groups. Model fit was determined using a combination of the minimum function chi-square and degrees of freedom (DF) (χ^2/DF), the comparative fit index (CFI), the Tucker–Lewis index (TLI), and the root mean square error of approximation (RMSEA). Conventional standards for good fit are values above 0.95 for the CFI and TLI and values below 0.06 for the RMSEA.[15] Change in model fit was evaluated by changes in chi-square and DF ($\Delta\chi^2/\Delta$DF) and change in CFI (ΔCFI). Whereas P can be calculated for $\Delta\chi^2/\Delta$DF, ΔCFI > 0.01 is considered significant.[16] Missing data were handled using full information maximum likelihood, an estimation procedure that uses all available data to reduce bias resulting from missing data.