Appetite 40 (2003) 163–173

Appetite

www.elsevier.com/locate/appet

Research Report

Validation and cross-national comparison of the food neophobia scale (FNS) using confirmatory factor analysis

Phillip N. Ritchey[a], Robert A. Frank[b,*], Ulla-Kaisa Hursti[c], Hely Tuorila[d]

[a]*Department of Sociology, University of Cincinnati, Cincinnati, OH, USA*
[b]*Department of Psychology, Research and Advanced Studies, University of Cincinnati, ML 627, Cincinnati, OH 45221-0627, USA*
[c]*Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden*
[d]*Department of Food Technology, University of Helsinki, Helsinki, Finland*

Received 3 March 2002; accepted 7 August 2002

Abstract

The food neophobia scale [FNS; Appetite 19 (1992) 105] has been used to assess willingness to try new foods in studies conducted around the world. Although it is tempting to compare FNS scores across these studies, appropriate psychometric analyses are required to validate the scale and allow cross-cultural comparisons. These analyses were pursued in the current study using confirmatory factor analysis in conjunction with a data analysis strategy described by Steenkamp and Baumgartner [J. Consumer Res. 25 (1998) 78] and random, representative samples drawn from the United States, Sweden and Finland. A unidimensional scale was constructed using eight of the original 10 items from the FNS, and this model provided an excellent fit to the data from the US and Swedish samples. An acceptable fit was achieved for six items when data from the US, Sweden and Finland were used. Based on these analyses, we recommend that two items from the original FNS be dropped (items 5 and 9). Elimination of additional items may be premature given the potential contributions of difference in sampling and testing methodology associated with data collection from the three samples. Future efforts to develop a FNS for cross-national comparisons should target the development of a scale with 14–16 items so that dropping several items from a model (due to translation or other problems) allows retention of a sufficient number of items to insure a robust test. However, even with only six items, our results supported the conclusion that people from Sweden are generally more willing to try novel foods as compared to people from the US and Finland. Future studies should focus on the source of this enhanced willingness to try novel foods among the Swedes and the potential use of this information in the development of programs aimed at facilitating dietary change.

Keywords: Neophobia scale; Factor analysis; Standardized residuals

Problem

Both the importance of variety in the diet and recent increases in the availability of new foods have stimulated interest in the responses of humans to novel foods. Research on this topic has shown that people vary dramatically in their willingness to try new foods, with some people showing a strong bias toward novel food avoidance. This tendency is referred to as food neophobia. Recent studies have shown that food neophobia affects both the quality and variety of foods in the diet (Falciglia, Couch, Gribble, Pabst, & Frank, 2000), and therefore may be an important variable to consider in studies of dietary change.

The Food Neophobia Scale (FNS) is a psychometric instrument developed by Pliner and Hobden (1992) to measure food neophobia. People complete the FNS by indicating their degree of agreement/disagreement with 10 statements about foods or eating situations (Table 1). The internal consistency of the FNS has been verified in multiple laboratories using responses of diverse groups of people (Cronbach's α is typically in the 0.8–0.9 range; Frank, Reilley, Schroth, Werk, & Wehner, 1997; Hursti & Sjödén, 1997; Pliner & Hobden, 1992; Tuorila, Meiselman, Bell, Cardello, & Johnson, 1994). A number of studies have shown that the FNS accurately predicts responses to novel foods (Arvola, Lähteenmäki, & Tuorila, 1999; Hobden & Pliner, 1995; Hursti & Sjödén, 1997; Pliner, Eng, & Krishnan, 1995; Pliner & Hobden, 1992; Raudenbush & Frank, 1999; Raudenbush, Schroth, Reilley, & Frank, 1998;

* Corresponding author.
E-mail address: robert.frank@uc.edu (R.A. Frank).

0195-6663/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0195-6663(02)00134-4

Table 1
Items of the FNS

1. I am constantly sampling new and different foods
2. I don't trust new foods
3. If I don't know what a food is, I won't try it
4. I like foods from different cultures
5. Ethnic food looks too weird to eat
6. At dinner parties, I will try new foods
7. I am afraid to eat things I have never had before
8. I am very particular about the foods I eat
9. I will eat almost anything
10. I like to try new ethnic restaurants

Each item has a seven-point Likert response set: disagree strongly, disagree moderately, disagree slightly, neither disagree nor agree slightly, agree moderately, and agree strongly.

Tuorila et al., 1994). In addition, it has been used to assess the impact of a wide variety of independent variables on willingness to try novel foods (Loewen & Pliner, 1999; McFarlane & Pliner, 1997).

The FNS has been used with diverse samples collected in North America and elsewhere. A comparison of FNS scores across different national samples suggests that people from Canada, the United States and Finland are more neophobic than Swedes (Hursti & Sjödén, 1997; Pliner & Hobden, 1992; Raudenbush et al., 1998; Tuorila, Lähteenmäki, Pohjalainen, & Lotti, 2001). However, the original psychometric development of the scale was based on a convenience sample of Canadian college students. While studies report respectable reliability coefficients for the scale, high reliability does *not* insure the scale's unidimensionality, nor legitimize its use for cross-national comparisons. The psychometric properties of the scale need to be properly investigated with representative samples and quantitative tools that verify scale unidimensionality and comparable scale structure across national samples.

The current study used confirmatory factor analysis to assess the measurement model underlying the FNS using a representative sample drawn in the United States. This model was then evaluated for its fit to data collected in Sweden and Finland.

Scaling issues

Here we define our orientation to the issues of the study. Underlying the FNS is a *reflective* measurement model (reflective model; Diamantopoulos & Siguaw, 2000: p. 21). A single underlying factor (alternatively called construct or latent variable) – hopefully what we measure with the scale – is viewed as the common cause of each of several items. The reflective model expresses that the scale is multi-itemed and unidimensional. In a reflective model, saying that each item measures the same underlying factor is equivalent to saying that each is caused by (reflects) the same factor. Responses to an item at least partially measure the factor; thus, the underlying factor at least partially causes these observed scores.

Two important advantages of multiple items are that they: (1) can improve a measure's reliability and (2) permit assessments in a study's measurement phase that can contribute to a scale's construct validity. The concept of reliability focuses on *random* (as opposed to systematic) error in a measure. Improvement in reliability occurs because the scale is like an average of item scores; random error is partially offset or washed out in the averaging process. When a composite is formed from the items for use in research, high reliability (low random variability) in the scale is important. It is well recognized that validity of a scale is conditional on its reliability. Less generally recognized is that statistical parameters measuring relationships between the scale and other variables, e.g. correlation coefficients, partial correlation coefficients, and regression coefficients, are biased due to the lack of reliability in measurement. In general, the more reliable the variables, the less the distortion in parameter estimates. Additionally, the more reliable the variables, the greater the power of statistical tests.

As noted, multiple items permit assessments in a study's measurement phase that can contribute to a scale's construct validity. With four or more items, a researcher can use confirmatory factor analysis to corroborate that the expected reflective model fits the data. If the model fits, the findings contribute to the scale's construct validity; we are more confident that we are measuring what we think because the model conforms to our expectations.

A reflective model is unidimensional; each of its items measures the same phenomenon, although they may do so to varying degrees. Individual items may measure other phenomena too, referred to as unique factors or causes. However, the only *common* or shared influence of any of the items is the single underlying factor. The other phenomena tapped by items are uncorrelated with the common factor. When using confirmatory factor analysis to test the fit of a reflective model, we are testing the idea that our scale is unidimensional. Confirmatory factor analysis can corroborate that the postulated reflective model *fits* the observed data. As the name confirmatory factor analysis implies, the procedure tests to confirm if a particular factor model is consistent with the data. The reflective model asserts that the covariance among the member items are totally due to their having a common cause, the single underlying factor. The test of this model is based on this idea. The idea is expressed as a set of equations (following path analysis principles) in which observed item covariances equal the products of factor loadings, and observed item variances are partially accounted for the underlying factor and partially unaccounted for. As in regression, this unaccounted variance is termed *error* variance and is due to other (unique) factors not in the model and to random error.

The confirmatory factor analysis solution for a reflective model evolves through an iterative procedure. The variance of the underlying factor, factor loadings (which are slope