Food Quality and Preference 40 (2015) 296–303

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Introduction of a quality index, based on Generalizability theory, as a measure of reliability for univariate- and multivariate sensory descriptive data

Alard Verhoef *, Gern Huijberts, Wim Vaessen

Essensor B.V., Rubensstraat 215A, 6717 VE Ede, The Netherlands

ARTICLE INFO

Article history:
Received 18 December 2013
Received in revised form 6 June 2014
Accepted 9 June 2014
Available online 24 June 2014

Keywords:
Generalizability (G) theory
Descriptive sensory analysis
Assessor and panel performance
Multivariate- and univariate panel reliability
Overall quality measure

ABSTRACT

In the sensory industry there is a need to quantify the overall usability or reliability of the final results of a sensory study. The sensory field has developed various statistical methods and techniques resulting in graphs or plots aimed at identifying assessors that contribute high error variance. Most of these methods are focused on estimating levels of repeatability, discriminability and agreement. Yet not much attention has gone to estimating the general level of quality of the dataset as a whole. Quality indicators could be helpful to quickly determine which attributes still require panel training, when panel training has been done in a sufficient manner and to determine the usability of the final results.

This contribution proposes to use a statistical method, Generalizability theory (G theory), to monitor reliability (or generalizability) (Cronbach, Rajaratnam, & Gleser, 1963; Gleser, Cronbach, & Rajaratnam, 1965; Rajaratnam, Cronbach, & Gleser, 1965) and to study the univariate and multivariate reliability of the dataset. Generalizability theory (G theory) can characterize, disentangle, and estimate all different sources of variation as defined by the investigator. This makes G theory a useful tool for characterizing products, when dealing with multiple different sources of variance in a set of product evaluations. Univariate generalizability coefficients (or univariate quality index) could serve as quality indicators to determine which attributes are sufficiently reliable and in which cases panel training is beneficial for the discriminative ability of the panel or when panel training is redundant. Multivariate generalizability coefficients (or multivariate quality index) could capture reliability of an entire study expressed in one easily interpretable measure and could answer questions related to the overall quality of the experimental design of the study.

In this study, a demonstration will be given on how G theory is used to test attribute reliability and to assess panel performance.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

In the sensory industry there is a need to quantify the overall usability or reliability of the final results of a sensory study. The sensory industry uses quantitative descriptive techniques to describe the sensory characteristics of a set of products, in order to give more specific insights about the effects of ingredients and processes to the product developers. For that, sensory research makes use of a panel of assessors, who use their human senses as instruments of measurement to evaluate a set of products according to specific sensory descriptors or attributes (e.g. appearance, taste, smell, mouthfeel attributes). Every assessor in the panel aims to (1) discriminate between products (discriminability), (2) repeat the assessments spread over different occasions (repeatability), and (3) agree with all other assessors on the sensory sensation of the attributes (agreement, includes similar use of scale and absence of cross-over effects) (Derndorfer, Baierl, Nimmervoll, & Sinkovits, 2005; Kermit & Lengard, 2005). Error variance (unwanted variation) increases when assessors do not meet one or more of these criteria. Usually panel training is given to decrease this error variance and improve panel performance. The sensory field has developed various statistical methods and techniques resulting in graphs or plots aimed at quickly and efficiently identifying aspects of the measurement that contribute high error variance (Dahl & Naes, 2009; Dahl, Tomic, Wold, & Naes, 2008; Dijksterhuis, 1995; Kermit & Lengard, 2005; Latreille et al., 2006;

Abbreviations: G theory, Generalizability theory; G study, generalizability study; D study, decision study; uQI, univariate quality index; mQI, multivariate quality index.

* Corresponding author. Tel.: +31 (0) 318 452222/611751779.
E-mail address: a.verhoef@essensor.nl (A. Verhoef).

http://dx.doi.org/10.1016/j.foodqual.2014.06.007
0950-3293/© 2014 Elsevier Ltd. All rights reserved.

Lea, Rodbotten, & Naes, 1995; Næs, Brockhoff, & Tomic, 2010). Some of these tools are available as open source sensory packages such as PanelCheck (PanelCheck, 2010).

Usually multiple analyses are necessary in order to evaluate all criteria that define a good panel result. To overcome this problem, several methods were developed that aimed to capture panel performance expressed in coefficients captured in collective frameworks. The "Scor-corr RV framework" (Schlich, Pineau, Brajon, & Qannai, 2004), based on the RV coefficient, and extended by Tomic, Forde, Delahunty, and Naes (2013) proposed using five indices (values between 0 and 100), for each assessor separately, representing a assessors agreement (amongst products and attributes), repeatability (amongst products and attributes) and discriminability. Yet this method has paid little attention to the estimation of overall usability or general levels of reliability of the dataset as a whole. Another study included reliability coefficients (Bi, 2003), to evaluate performance of descriptive panels. They define an agreement coefficient, describing the degree of consensus of the panel and panellist for each product, and a reliability coefficient, an intraclass correlation coefficient (ICC) and describing the discriminative ability of the panel or panellist for a set of products. They extended their work to a collective framework for monitoring and assessing panel performance (Bi & Kuesten, 2012) in which they use accuracy, validity and reliability indices. This framework is explained with a very general approach and it includes a multivariate intraclass correlation coefficient (MICC). Furthermore the paper emphasizes the similarities between ICC and Cronbach's alpha, a reliability coefficient of internal consistency. Although the authors did not make this comparison with Cronbach's alpha (Cronbach & Shavelson, 2004) with the larger context of Generalizability theory.

This paper proposes to use Generalizability theory (Cronbach, Rajaratnam, & Gleser, 1963; Gleser, Cronbach, & Rajaratnam, 1965; Rajaratnam, Cronbach, & Gleser, 1965), based on Cronbach's alpha, to study the univariate and multivariate reliability of a sensory dataset. Univariate generalizability coefficients could then serve as quality indicators to determine which attributes are sufficiently reliable and in which cases panel training is beneficial for the discriminative ability of the panel or when panel training is redundant. Multivariate generalizability coefficients could capture reliability of an entire study expressed in one easily interpretable measure and could answer questions related to the overall quality of the experimental design of the study. Generalizability theory uses analysis of variance to determine generalizability coefficients, that estimate the extent to which observations are generalizable beyond the scope of the sensory dataset at hand. Generalizability theory (G theory) has been applied in other research such as educational research (Nussbaum, 1984) and marketing research (Rentz, 1987).

The current paper aims to illustrate how G theory can be applied to estimate reliability and measurement error in sensory descriptive studies. First the basic principles of G theory are explained. A framework is presented consisting of equations for both univariate and multivariate G theory specified for sensory descriptive analysis. The framework is closely following the denotation used by Brennan (2001b). For demonstration purposes, this framework will also be applied in practice on a sensory descriptive study performed on different varieties of an Asian condiment sauce. The univariate generalizability coefficients and multivariate generalizability coefficients will be referred to as the *univariate quality index* (*uQI*) and *multivariate quality index* (*mQI*) respectively, in this paper.

Calculation

Theoretical basis of Generalizability theory

Regardless the quality of a panel, unwanted variance will never be fully eliminated and is an inherent part of measurement. Therefore any observed score is always composed of a true score (wanted variation) and an error score (unwanted variation), like shown in (1).

$$X = T + E \tag{1}$$

where *X*, *T*, and *E* are observed, true, and error score. Here a person's true score refers to the score a person would have when being measured an indefinite number of times in the same experimental setup (Cronbach & Shavelson, 2004). *Reliability* involves quantifying the consistencies and inconsistencies between true score (universe score) and observed score (Nunnally, 1967) and is therefore usually defined as the squared correlation ($\rho^2{}_{X,T}$) between observed – and true score, as shown in (2)

$$\rho^2_{(X,T)} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2} \tag{2}$$

Cronbach et al. developed a theory which became known as the Generalizability theory (Cronbach et al., 1963; Gleser et al., 1965; Rajaratnam et al., 1965). Generalizability theory (G theory) is a method used to investigate the univariate and multivariate generalizability (reliability) of data.

According to Cronbach and others, G theory offers a comprehensive framework for the study of reliability. G theory can characterize, disentangle, and estimate all different sources of variation as defined by the investigator (Brennan, 2011; Cardinet, Johnson, & Pini, 2012; Shavelson & Webb, 1991). This makes G theory applicable to principally any empirical study or other studies where estimating measurement error is important.

Cronbach introduced the concept of a *facet* to indicate aspects of a study that can vary from one study to the next and thus constitute variance. A facet usually is composed of a set of similar conditions (e.g. persons, items, or occasions), where the investigator decides what similar means. The *face of generalization* consists of all facets that are considered not to vary. Nevertheless, these facets contribute unwanted variation and contain random and systematic errors. In sensory analysis the face of generalization usually includes an assessor (*a*) facet and an occasion (*o*) facet. On the contrary, the *face of differentiation* consists the facet that is compared in the study and is the object of measurement. In most of the examples presented by Cronbach the face of differentiation contains the persons tested, but principally any facet could represent the face of differentiation (Cardinet, Tourneur, & Allal, 1976). In this study we will use product (*p*) as the facet of differentiation since the objective of the Asian condiment sauce study is to discriminate and characterize products.

G theory differentiates *generalizability studies* (G study) and *decision studies* (D study). A G study is not concerned with defining facets as facets of differentiation or generalization, nor is it focussed on estimating reliability. Instead, a G study is a purely objective study designed to identify as many sources of variance (facets) as possible and to estimate these errors. A G study is usually followed up by a D study which is focussed on estimating reliability for a specific purpose or decision. In this Asian condiment sauce study, the D study is to discriminate and characterize products. Depending on the results of the G study, the number of conditions in a D study can be increased or decreased depending on the magnitude of the errors.

A framework of G theory: univariate G theory

Linear model

The univariate application is used when there is only one dimension of generalization for each facet. This entails that univariate G theory is applied when reliability is estimated for each attribute separately. A typical sensory profiling study, includes a *product* (*p*) facet, *assessor* (*a*) facet, and *occasion* (*o*) facet. The corresponding univari-

Download English Version:

https://daneshyari.com/en/article/4317042

Download Persian Version:

https://daneshyari.com/article/4317042

Daneshyari.com