Situational appropriateness in food-oriented consumer research: Concept, method, and applications

6

Davide Giacalone
SDU Innovation & Design Engineering, Department of Technology and Innovation, Faculty of Engineering, University of Southern Denmark, Odense, Denmark

6.1 Appropriateness as a basic context construct

In the food and beverage industry, hedonic responses (expressed as degree of liking or preference for a set of test products) have traditionally been an important product performance indicator. With both everyday experience and empirical research telling us that the sensory acceptability of foods strongly predicts consumption, there is little doubt that this is a meaningful dimension to study in central location tests (CLT). Indeed, sensory and consumer scientists have long been interested in understanding the relationship between specific sensory properties and food acceptability, to identify those that maximize liking within a specific product category, as well as for specific consumer segments within that category.

Unfortunately, strong sensory and consumer performance does not necessarily predict product success: for example, it does not guarantee that a product will be purchased often or that it is appropriate for the usage situation that the product developers had in mind (Jaeger & Porcherot, 2017; Rosas-Nexticapa, Angulo, & O'Mahony, 2005). As Marshall (1993) reminds us, the final decision to buy or consume a particular food depends as much on the anticipated usage context as it does on the intrinsic qualities of the product (or the consumer). Highly liked and even preferred products may not be chosen simply because they are inappropriate. For example, a consumer may highly appreciate a very complex wine when fine dining, but the same individual would choose a less expensive one for a routine meal or a picnic. Therefore, understanding consumer food choices requires us to think about how consumers use products, as well as how they select them.

The concept of appropriateness focuses on the usage contexts of food products (as well as consumer products generally). Appropriateness is defined as "the quality of being especially suitable or fitting" (Merriam-Webster's Dictionary, 2018, online at merriam-webster.com). In consumer research, *situational appropriateness* is generally defined as the perceived fit between a product and a target usage context (Ratneshwar & Shocker, 1991; Schutz, 1988). The term "perceived" reflects the fact that situational appropriateness is a cognitive phenomenon that requires a subjective evaluation on the part of the consumer. However, unlike evaluations of preferences

Context. https://doi.org/10.1016/B978-0-12-814495-4.00006-4

and acceptability, where the emphasis is squarely on subjective experiences, the concept of appropriateness includes a plethora of rules that are culturally determined and learned through experience and socialization (e.g., what ingredients go together, what is a typical portion size, what foods one usually eats for breakfast, etc.). Indeed, the term "appropriateness" in and of itself suggests a normative aspect of how well a food fits the situation in which it is supposed to be consumed.

How relevant is perceived situational appropriateness for explaining consumers' food-related behavior? As a starting point, one can certainly expect that appropriateness interacts with aspects related to the product (both intrinsic—such as sensory properties or caloric content, and extrinsic—such as a packaging and brand) and to the consumer (personal preferences, physical states such as hunger, thirst, etc., and mental processes, such as specific goals and motives) to determine consumers' choices. However, while the latter two aspects are always in focus in sensory and consumer research, situational influences have received comparatively little attention. If we look at the general marketing and consumer psychology literature, the relevance of the appropriateness construct for explaining consumer behavior is supported by a substantial body of research showing that consumers differentiate products on the basis of the anticipated usage situations (Lai, 1991). The underlying theoretical premise for this line of work is that consumers' perceptions of products rarely occur in isolation, but, rather, relative to some frame of reference. The usage situation of a product is an ecological factor that can help define consumers' goals, and thus orient their choices toward "situationally appropriate" solutions (Giacalone & Jaeger, 2016; Ratneshwar & Shocker, 1991). In consumer psychology, situational effects on consumer choices are typically explained on the basis of the compatibility principle (Tversky, Sattah, & Slovic, 1988; Tversky & Simonson, 1993), according to which when individuals make choices they tend to select the options that are superior on a salient dimension. In this sense, the anticipated usage situation can orient consumers' attention to product attributes relevant to fulfilling goals associated with it. To illustrate the point, consider the example of a consumer who wants to purchase a bottle of wine out of n alternatives. In theory, the consumer could evaluate all relevant attributes (e.g., price, origin, vintage, grape variety, alcohol by volume, etc.), assign a subjective value to each, and choose the product with the best overall value. This hypothetical decision-making process is essentially the familiar theory of rational choice based on individual preferences. Let us now assume that the same person is now choosing between the same set of alternative products, keeping in mind, say, that they have just decided to go on a diet. In such a situation, the compatibility principle suggests that they will be more likely to focus on the dimension that is made salient by the anticipated usage situation, which then provides a cognitively efficient "metric" for comparing the wine alternatives. In this instance, this consumer may equate alcohol content with caloric content, and choose the wine with the lowest alcohol by volume as the most appropriate given their current health goal. This second model of decision-making process takes its point of departure into bounded rationality (the notion that consumers' information processing capabilities are limited and flawed) and assumes that consumer choices are often context-dependent (Tversky & Simonson, 1993; Warlop & Ratneshwar, 1993).

The important implication of this viewpoint is that consumers in many cases choose products to fulfill the goals associated with a particular consumption situation, rather than solely on the basis of individual preferences and product characteristics (Belk, 1975, Ratneshwar & Shocker, 1991; Giacalone, 2018). Informed by this understanding, a number of methodological approaches based on situational segmentation, that is, on the identification of perceived product benefits across different usage situations, have been developed.

Looking at the history of situational research in sensory and consumer science, we can distinguish between two macro approaches. The first one involves the use of natural or physically manipulated environments to vary contextual elements of interest, and then to record consumer responses in a target context (e.g., Bell, Meiselman, Pierson, & Reeve, 1994; Di Monaco, Giacalone, Pepe, Masi, & Cavella, 2014; Meiselman, Johnson, Reeve, & Crouch, 2000; Sester et al., 2013). Most studies adopting this line of work have focused on food acceptability or intake as primary outcome measures, but some have also provided indications that situational appropriateness is relevant to consumer choices. For example, a well cited paper by Bell and colleagues demonstrated that changing the décor and the names of items of a restaurant menu increased the frequency of choice for items that patrons perceived as more congruent with the environment, such as pasta with an Italian ethnic theme (Bell et al., 1994).

While natural or naturalistic approaches can provide a high degree of ecological validity, the practical planning and implementation of these studies may be burdensome and/or present pitfalls due to factors unforeseen by the researchers (Cardello & Meiselman, 2018; Jaeger & Porcherot, 2017). This has limited the number of such studies, though efforts to measure consumer responses in natural environments have been called upon by several authors (Jaeger & Porcherot, 2017; Meiselman, 2013), and may be expected to increase siginificantly in the near future, thanks to the increased availability of immersive virtual reality (VR) technologies (Jaeger, Hort, Porcherot, et al., 2017; see also Hehn et al., and Hartmann & Siegrist, this volume).

The second major stream of research has focused on situational effects in central location tests (CLT), where the majority of sensory and consumer studies takes place. This has consisted in the modification of standard CLT protocols to include a usage context of interest, for example, by presenting consumers with a scenario as a means to evoke a focal situation to consumers, and ask them to imagine this situation while carrying out the product evaluations (Hein, Hamid, Jaeger, & Delahunty, 2010, 2012; Jaeger & Porcherot, 2017). Other CLT-based approaches have involved explicit evaluation of situational appropriateness of products, using a wide variety of methods, ranging from qualitative methods, such as focus groups (Elzerman, van Boekel, & Luning, 2013), personal interviews (Hartwell, Shepherd, Edwards, & Johns, 2016), and word associations (de Andrade, de Aguiar Sobral, Ares, & Deliza, 2016), to quantitative methods such as Free Choice Profiling (Jack & Piggott, 1991; Piggott, Sheen, & Apostolidou, 1990) and repertory grid interviews (Jack, Piggott, & Patterson, 1994; McEwan & Thomson, 1988; Jaeger, Rossiter, & Lau, 2005; Scriven, Gains, Green, & Thomson, 1989; Raats & Shepherd, 1991/1992).

Among CLT-based methods, one of the earliest and best-known is the "item-by-use" (IBU) method, developed and popularized by Schutz (1988, 1994). In this approach, which is sometimes referred to as "substitution in use" in the marketing literature (e.g., Ratneshwar & Shocker, 1991), consumers are presented with a product set and a list of possible usage contexts, and are asked to indicate how well a product fits each of them (Schutz, 1988, 1994). The IBU approach enables a comprehensive analysis of the appropriateness of products with reference to a specific usage context, thus explicitly considering products as means to reach an end defined by a particular usage situation. Situational appropriateness ratings may have important behavioral correlates. In particular, recent research has shown that appropriateness evaluations (1) are predictive of consumer food choices, willingness to use, and reported intake (Elzerman, Hoek, van Boekel, & Luning, 2015; Giacalone & Jaeger, 2016, 2017; Sosa, Martínez, Arruiz, Hough, & Mucci, 2005) and (2) underpin other product-related variables, such as emotional responses and perceived product uniqueness (Jaeger, Cardello, Chheang, et al., 2017; Jaeger, Cardello, Jin, et al., 2017; Piqueras-Fiszman & Jaeger, 2014a, 2014b, 2015a, 2015b). Taken collectively, these studies indicated that situational appropriateness should be considered as an important product performance criterion in CLT tests. A thorough discussion of the IBU approach constitutes the main focus for the rest of the chapter.

6.2 The "Item-by-use" (IBU) approach to measuring product appropriateness

6.2.1 *Origin and historical use in sensory and consumer science*

The introduction of the IBU approach in sensory and consumer research is owed to the pioneering work of Howard G. Schutz, who first became exposed to the situational appropriateness construct while working with anthropologist Volney Stefflre, and later adapted a basic anthropological technique (Stefflre, 1971) for application in consumer product testing in the 1970s, during his tenure at the University of California at Davis.

Schutz's reasoning in developing the IBU appropriateness method was that the common food classification systems based on nominal categories (e.g., dairy, bakery, etc.) or nutritional content did not necessarily reflect the way consumers would classify different food items (Schutz, Rucker, & Russell, 1975), and may therefore be of limited value for developing and marketing new products.

As originally proposed, the IBU method consists of presenting a consumer with a list of possible consumption situations, and asking him/her to rate how well a product fit each of them, on a 7-point scale ranging from "'Never appropriate" to "Always appropriate." In principle, the IBU method can be applied to any product category, including non-food products, such as personal and household care products, and even durable products, as long as they have more than one possible usage context.

Table 6.1 provides a comprehensive (yet possibly non-exhaustive) list of journal publications featuring the IBU approach, starting from the early work of Schutz

Table 6.1 List of research publications employing the IBU approach, sorted chronologically

Reference	Year	Product category	Subjects (N)	Contexts (N)	Stimuli (N)	Type of stimulus	Type of scale
Schutz & Ortega	1974	Wine and other alcoholic drinks	52	48	56	Names	7-pt scale
Schutz, Fridgen, & Damrell	1975	Rice and related products	200	48	52	Names	7-pt scale
Schutz, Rucker, & Russell	1975	Misc. food and beverages	200	48	56	Names	7-pt scale
Schutz & Rucker	1975	Misc. food and beverages	60	10	10	Names	2-,3-,6-,7-pt scales
Baird & Schutz	1976	Misc. food and beverages	135	25	20	Names	2-pt scale
Schutz & Phillips	1976	Textiles	50	48	46	Names	7-pt scale
Bruhn & Schutz	1986	Dairy products	51	45	46	Names	7-pt scale
Resurreccion	1986	Vegetables	67	10	43	Names	7-pt scale
Martens et al.	1987	Vegetables	135	34	30	Names	7-pt scale
White, Resurreccion, & Lillard	1988	Meat products	203	12	10	Names	7-pt scale
Raats & Shepherd	1991	Milk	40	49	7	Images	7-pt scale
Nantachai, Petty, & Scriven	1991	Meat products	30/30	21/19	18/18	Names	100 mm line scale
Lähteenmäki & Tuorila	1995	Ice-cream	40	10	3	Samples (Blind)	9-pt scale
Cardello & Schutz	1996	Misc. food and beverages (29 studies)	27 to 38	10	1 to 4	Samples (Blind)	7-pt scale
Lähteenmäki & Tuorila	1997	Drinks	243	18	8	Samples (Blind)	7-pt scale
Cardello et al.	2000	Misc. food and beverages	115	10	22	Names	7-pt scale
Jaeger	2000	Apple products	124	15	4	Images	7-pt scale
Sosa et al.	2005	Seasoning sauces	240	16	5	Names	7-pt scale
Schutz, Cardello & Winterhalter	2005	Textiles	100	30	16	Names	7-pt scale
Mejlholm & Martens	2006	Beer	38	23	10	Samples (Blind)	7-pt scale

Continued

Table 6.1 Continued

Elzermann et al.	2011	Meat substitutes in dishes	93	4	5	Dishes	100 mm line scale
Hersleth et al.	2011	Ham	81	2	8	Samples (Informed)	9-pt scale
Bach et al.	2013	Jerusalem artichoke	49	2	5	Samples (Blind)	5-pt scale
Elzermann et al.	2015	Meat substitutes in dishes (online survey)	251	6	5	Images	100 mm line scale
Piqueras-Fiszman & Jaeger	2014a	Apple/ Brownie	76/81	3	1/1	Images	9-pt scale
Piqueras-Fiszman & Jaeger	2014b	Brownie/ Kiwifruit/ Brownie or crisps	115/302/ 188	3/3/3	3/3/1	Images	9-pt scale
Piqueras-Fiszman & Jaeger	2015	Misc. food and beverages (7 studies)	1336 (96 to 417)	4 to 16	1	Names	7-pt scale
Giacalone et al.	2015	Beer	76/97/ 93/145	15/15/ 15/9	9/9/9/9	Images	CATA
Bach et al.	2015	Beetroot	49	3	5	Samples (Blind)	5-pt scale
Giacalone & Jaeger	2016	Fruit/Wine/ chocolate/ kiwifruit	246/112/ 192/302	16/16/ 15/15	19/12/ 13/12	Names or images	CATA
Geertsen, Allesen-Holm, & Giacalone	2016	Sea-buckthorn beverages	200	8	7	Samples (Blind)	CATA
Stolzenbach et al.	2016	Apple juice	196	6	4	Samples (Blind)	7-pt scale
Cardello et al.	2016	Beer	203	15	8	Samples (Blind)	CATA
Jaeger et al.	2017a	Beer	128	14	9	Samples (Blind)	CATA
Jaeger et al.	2017b	Dark chocolate/ white chocolate	139/128	12/13	8/8	Samples (Blind)	CATA

and collaborators in the 1970s, up to the time of writing. Most of the publications involve food and beverage products, but applications in non-food products have also been proposed (Schutz, Cardello, & Winterhalter, 2005; Schutz & Phillips, 1976).

Early applications (1970s–1990s) of the IBU approach involved evaluations of food names, and are characterized by a very large number of both stimuli and usage

contexts, sometimes resulting in >2500 food use combinations, and a completion time of more than two hours (Schutz & Ortega, 1974). These studies were conducted as self-administered questionnaires, and respondents reportedly spread the task over different time intervals. A key takeaway of these early studies on appropriateness is that they consistently show, through the application of factor analysis, that consumers grouped foods according to common usage contexts such as "high-calorie foods," "specialty meal items," "common meal items," "inexpensive filling foods," "healthy foods," and so forth (Marshall, 1993; Schutz, 1988, 1994), but that these factors were not synonymous with any objectively defined food category. Schutz (1988) has reviewed most of these early studies utilizing the IBU method, and argued that appropriateness is an important cognitive-contextual aspect for both theoretical and practical purposes.

The IBU approach gained momentum in the mid-90s, following some influential papers that extended its application range significantly. Raats and Shepherd (1991) were the first to use the IBU approach to differentiate products within the same category, in a study comparing different types of milk (fresh *vs*. UHT and with different fat content levels), demonstrating the usefulness of this technique with a product range close to that of a typical CLT. This study was also the first to use product images and accompanying information as test stimuli, as opposed to food names.

Shortly afterward, studies by Lähteenmäki and Tuorila (1995, 1997) and by Cardello and Schutz (1996) extended the IBU method to measuring responses to actual products in blind tests, demonstrating that (1) consumers can differentiate both *within* and *between* products on the basis of sensory differences, and, (2) that products that do not differ in terms of acceptability can nonetheless be widely different in terms of appropriateness for specific usage contexts. Fig. 6.1 shows an example of this incidence, both when considering (a) widely different products, as well as (b) products in the same category.

Cardello and Schutz (1996) also reported that the presence of appropriateness ratings does not significantly alter hedonic ratings, suggesting that collecting appropriateness data in conjunction with standard sensory and consumer tests can be done without hampering the validity of the test results.

Taken overall, these studies provided a strong rationale for using the IBU method as an adjunct to hedonic testing of food products in CLT. Cardello and Schutz (1996) called for the routine collection of IBU data in CLT studies, in order to evaluate not only whether the test products have high acceptability, but also high appropriateness for the usage context(s) they are intended for.

Let us note here that this call is as relevant as ever some 20 years later, and that the IBU method fits quite naturally with common new product development practices, where products are generally developed with an end goal in mind, for example, to fit a specific consumer segment (e.g., women, elderly, etc.), or to provide a specific benefit (e.g., health, convenience, etc.).

Later studies have indicated that the situational appropriateness construct may have very important behavioral correlates for predicting consumer choice and consumption. Cardello, Schutz, Snow, and Lesher (2000) found significant associations between IBU appropriateness ratings and *expected* liking for food (with correlations

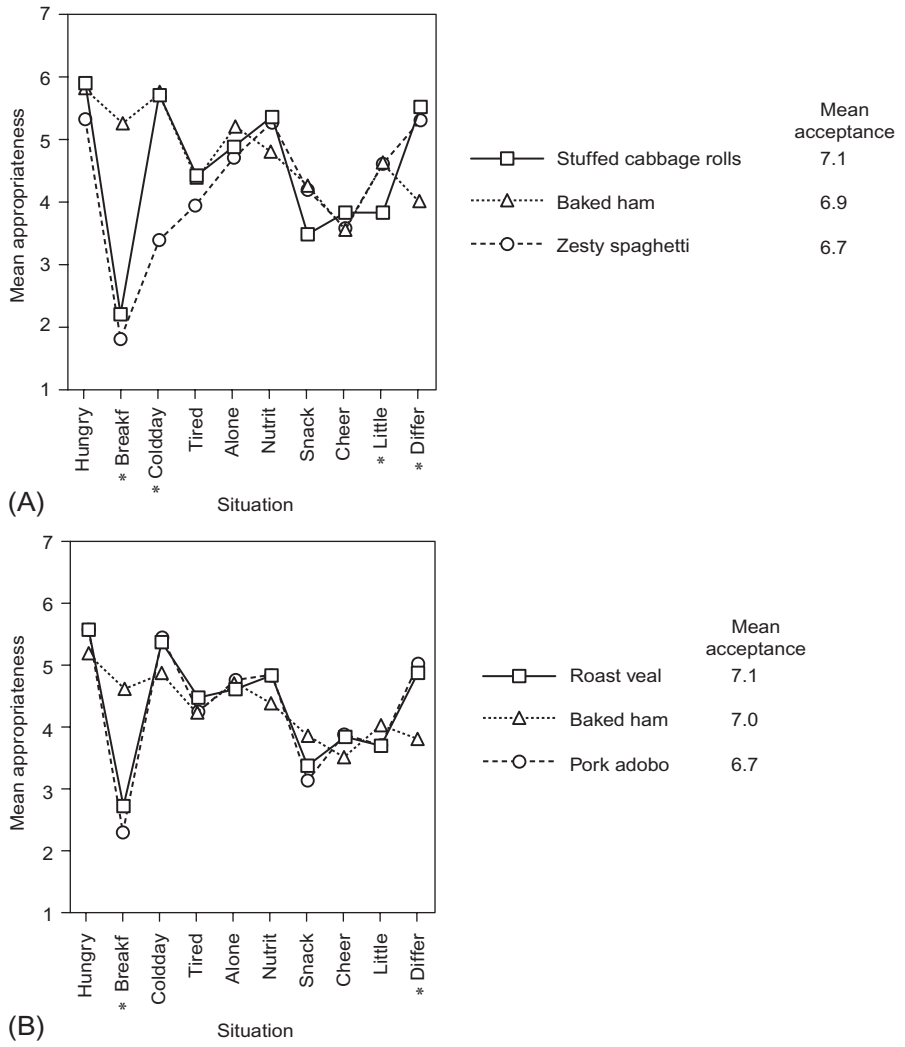

Fig. 6.1 Products having equal acceptability can nonetheless differ greatly in appropriateness for use, as exemplified by this plot showing mean liking and mean appropriateness ratings for three widely different products (A), and for three products in the same category (B) (*, $P < .05$; "cheer", when I want cheering up; "little", when I have little time to eat; "differ", when I want something different).

Reproduced with permission from Schutz, H. G. (1995). Eating situations, food appropriateness, and consumption. In Marriott, B. M. (Ed.), *Not eating enough: Overcoming underconsumption of military operational rations* (pp. 341–359), Washington, DC: National Academies Press.

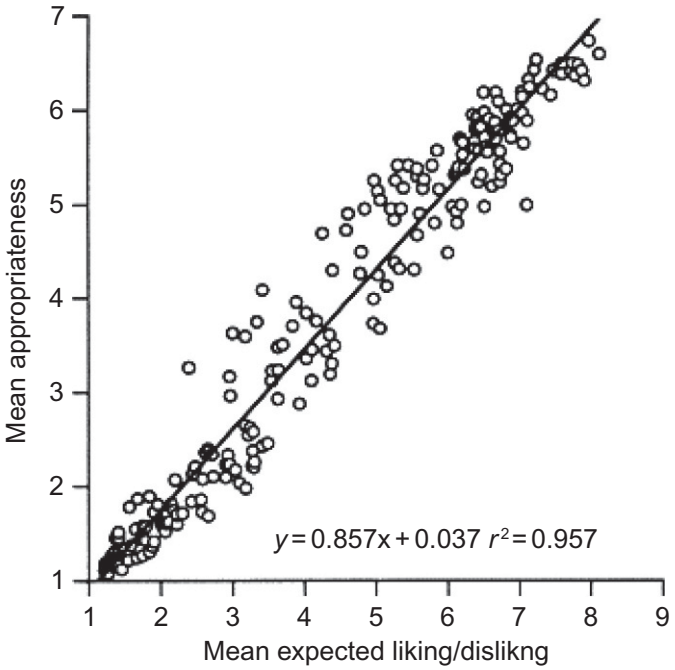

Fig. 6.2 Correlation between mean ($N = 115$) appropriateness ratings (7-pt scale) and expected liking (9-pt scale) for 220 food/uses combination, showing a high degree of associations between these two measures.
Reprinted with permission from Elsevier from Cardello, A. V., Schutz, H., Snow, C., Lesher, L. (2000). Predictors of food acceptance, consumption and satisfaction in specific eating situations. *Food Quality and Preference, 11,* 201–216.

ranging between 0.53 and 0.83, depending on context), and a very high degree of linear dependence when considering mean ratings (Fig. 6.2).

In a paper that has probably flown under the radar, Sosa and colleagues (Sosa et al., 2005) studied the correlation between appropriateness and self-reported frequency of intake in a series of usage contexts, reporting a near perfect correlation (all correlations $\geq$0.96) between these two variables (Fig. 6.3). Albeit based on a single product category (seasoning sauces), these results are indicative of appropriateness being a strong predictor of food consumption, especially when compared with hedonic ratings, which, at best, have been found to account for only about 50% of the variance in consumption (Cardello et al., 2000; Lau, Hanada, Kaminskyj, & Krondl, 1979; Tuorila, Hyvönen, & Vainio, 1994).

In recent years, the IBU method has been adopted for a wide range of applications, such as relating appropriateness for use to sensory attributes (e.g., Hersleth, Berggren, Westad, & Martens, 2005; Mejlholm & Martens, 2006), product packaging and information (e.g., Hersleth, Lengard, Verbeke, Guerrero, & Næs, 2011; Jaeger, 2000), and has been increasingly used in conjunction with hedonic and other perceptual product

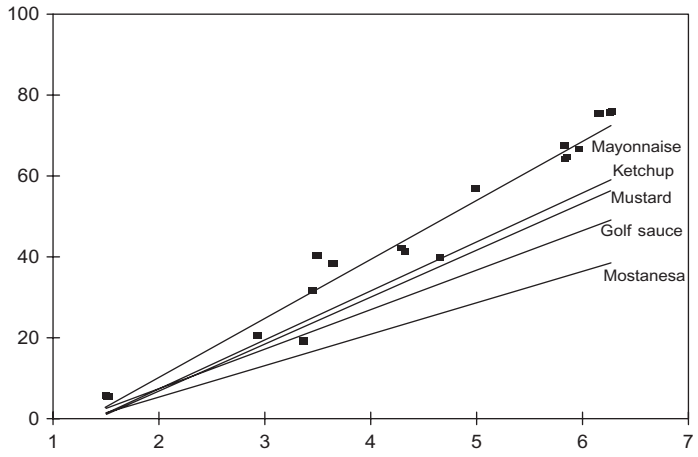

Fig. 6.3 Correlation between mean ($N = 240$) appropriateness ratings (7-pt scale) and self-reported frequency of consumption (100 mm scale from "never used" to "always used") for five seasoning sauces (Sosa et al., 2005). The scatterplot only shows the points for one of the sauces (mayonnaise), but according to the authors all products had a similar dispersion and all correlation coefficients were ≥ 0.96.
Reprinted with permission from Elsevier from Sosa et al. (2005).

responses, such as expectations, emotions, and well-being (Cardello et al., 2016; Jaeger, Cardello, Chheang, et al., 2017; Jaeger, Cardello, Jin, et al., 2017; Jaeger, Hort, Porcherot, et al., 2017). The value of IBU data in new product development has also been highlighted (e.g., Geertsen, Allesen-Holm, & Giacalone, 2016; Stolzenbach, Bredie, Christensen, & Byrne, 2016). An interesting application niche has focused on appropriateness in the context of culinary preparations, that is, by using meals and dishes as usage contexts, to drive gastronomic food development (Bach, Kidmose, Thybo, & Edelenbos, 2013; Bach, Mikkelsen, Kidmose, & Edelenbos, 2015), or to evaluate the possibility of ingredient substitutions (Elzerman et al., 2015; Elzerman, Hoek, van Boekel, & Luning, 2011).

Summarizing, the IBU approach is a flexible instrument for evaluating appropriateness in CLT tests. As a natural supplement to hedonic testing, it can help ensure that the test products not only have high acceptability, but also a high fit with the usage context(s) they are intended for. Although the total number of publications listed ($N = 34$) indicates a certain popularity of this method, the bulk of the work on appropriateness comes from few authors. Additionally, one may argue that this number is very low if compared with the number of studies that have focused on liking or preference in that same period. There may be different reasons for this. Low awareness may be one: for example, in my experience, many colleagues and reviewers (wrongly) assume that the IBU method is only applicable to differentiate between widely different products, but that it is not useful for products within the same category. This is likely a reminiscence of the early appropriateness papers; however, there are several examples of successful applications of discriminations within a range of product

differences that is very typical for CLTs (see Section 6.2.2.2). Perhaps there is a need for developing standard tools that can facilitate the adoption of the IBU method, in a similar way as methodological development in emotional measurement has paved the way for research in this area (Meiselman, 2015). Or maybe Schutz was simply way ahead of his time, and the field is still catching up. If that is the case, in light of the growing interest for context in consumer research, we can expect an increase in the utilization of situational appropriateness in sensory and consumer research.

With this in mind, the reminder of the chapter covers the main methodological aspects to consider when using the IBU method, and discusses recent advances and venues for future research in the area.

6.2.2 Methodological considerations for appropriateness evaluations using the "Item-by-use" (IBU) approach

6.2.2.1 Ballot format

Table 6.1 shows that the IBU method is quite flexible with respect to key methodological choices.

In early applications, the basic questionnaire format has consisted of a matrix crossing foods and uses, such as the one shown in Fig. 6.4. The task for the respondent is to fill in the matrix with values corresponding to perceived appropriateness in each use situation (typically on a 7-pt scale).

With respect to measurement, the 7-pt category scale originally proposed by Schutz has been the most common option, but other types of scales (e.g., 5-pt, 9-pt, line scales), as well as the check-all-that-apply (CATA) format, have also been used to collect appropriateness responses. The question of whether scale length and format affect the data has received little attention in the literature, with the exception of an old paper by Schutz and Rucker (1975) comparing data from 2-, 3-, 6-, and 7-point rating scales. This research concluded that the number of scale points did not affect the cognitive structures derived from appropriateness responses, indicating a high degree of robustness for this type of measurement (Schutz & Martens, 2001; Schutz & Rucker, 1975). Interestingly, this was the case also for a 2-point scale version (1 = Appropriate, 0 = Inappropriate) whose binary nature is equivalent to what one would today call a CATA question in a forced choice format, some 35 years before this method became popular in sensory and consumer science. The same binary scale was also employed in a later study comparing cognitive food structures in different ethnic groups (Baird & Schutz, 1976).

Generally, shorter scales (and especially the CATA format) are easier to process for consumers (Jaeger et al., 2013; Schutz and Rucker, 1975), so the choice of scale should ultimately take into consideration practical aspects such as the total testing time, the number of stimuli to evaluate, the co-presence of an acceptability and/or other questions in the same ballot, etc. Fig. 6.5 shows an example of a CATA ballot using for evaluation of visual stimuli.

The food-use matrix is more commonly associated with evaluation of food names in a self-administered questionnaires. When visual and taste stimuli are used—which

	When watching TV	When you are depressed	When eating out	When you are sick	In the summer	With coffee	When riding in a car	Late at night	When you are really hungry	For Sunday dinner
Jello										
Potato chips										
Chicken										
Orange juice										
Celery										
Soup										
Pizza										
Cereal										
Pie										
Grape										

Fig. 6.4 Example of a IBU matrix for appropriateness evaluations of food names across different usage situations. The respondents are instructed to fill each cell with a value ranging from 1 (Never appropriate) to 7 (Always appropriate).
Modified from Schutz, H. G., & Rucker, M. H. (1975). A comparison of variable configurations across scale lengths: An empirical study. *Educational and Psychological Assessment*, *35*, 319–324.

is the common case when appropriateness evaluations are conducted in the context of standard CLT tests—stimuli are normally evaluated monadically in separate ballots. It is customary to ask about overall acceptance first, and then have them rate the product for use appropriateness after, in order to minimize the risk of potential biases.

6.2.2.2 Selection of test stimuli

As already mentioned, the IBU is a flexible approach that can be used with a variety of stimuli, including names, actual products, and images (Table 6.1). Depending on the aims of the study, stimuli can either include items from different product categories, or be restricted to a specific product category.

Including a broad number of stimuli spanning across product categories is typically the case when the goal is to understand how consumers classify foods, as in earlier applications of the IBU approach (Schutz, 1988). The number of items for this type of study has ranged from 10 to as many as 56 (Table 6.1). There are no strict rules for

Please look at the picture on the right, and then answer the questions below.

DB Export Dry
(DB Breweries, NZ)

Which of the following situation(s) do you think this beer would be appropriate for?

Please tick *all* that apply.

- ☐ When I want something different
- ☐ To drink alone
- ☐ When I want something refreshing
- ☐ With dinner
- ☐ For a special occasion
- ☐ At a casual dining restaurant
- ☐ At home
- ☐ When I want to relax
- ☐ At a pub
- ☐ Anytime
- ☐ As an alternative to wine
- ☐ As a drink for women
- ☐ At a sport event
- ☐ To impress someone
- ☐ At parties
- ☐ I don't know

How familiar are you with the beer shown in the picture?

☐ 1 Not at all familiar	☐ 2	☐ 3	☐ 4	☐ 5 Extremely familiar

How often do you drink the beer shown in the picture?

☐ Never	☐ Rarely (1–2 times a year, or less frequently)	☐ Sometimes (3–12 times a year)	☐ Often (1–3 times a month)	☐ Regularly (once a week or more)

How familiar are you with the DB EXPORT brand in general?

☐ 1 Not at all familiar	☐ 2	☐ 3	☐ 4	☐ 5 Extremely familiar

Fig. 6.5 Example of a ballot for evaluation of IBU appropriateness in a CATA format using visual stimuli from a study on beer (Giacalone et al., 2015).

the choice of items, other than they should, of course, be representative of the foods available in that particular market, and cover different sensory intensity levels and nutritional groups (for examples of such lists, see Schutz, 1988, 1994; Schutz & Martens, 2001). An important thing to consider is the level of specificity to provide

to the consumers, which depends entirely on the research aims. For example, whether to provide a generic or a branded food name (e.g., "cola" or "Coca-Cola"), or whether to present food items separately or in combination (e.g., eggs *vs*. eggs and bacon), with or without additional qualifiers (e.g., "eggs" vs "fried eggs"), etc. These are very important choices, as they will almost certainly affect perceived appropriateness.

In other situations, stimuli may be restricted to a specific product category, where physical products or pictorial representations are used to elicit appropriateness evaluation from consumers. This would often be the case when appropriateness data is collected in CLT tests. There is still somewhat of a misconception to this day that the IBU appropriateness context has relevance only for differentiating between different food products, but that it cannot differentiate between products within the same category. However, this is not the case. For example, Lähteenmäki and Tuorila showed that appropriateness ratings obtained from blind taste tests differentiated between brands of vanilla ice-creams (1995), and between blueberry-raspberry juices varying in sweetener content (1997). Several other examples can be found in the literature (see e.g., Geertsen et al., 2016; Hersleth et al., 2005; Mejlholm & Martens, 2006; Stolzenbach et al., 2016), showing that appropriateness as a concept is relevant to differentiating product variants within the same category. Naturally, the magnitude of difference in perceived appropriateness is related to the heterogeneity in the stimuli set, and will generally increase manifolds as soon as extrinsic factors are included (Giacalone et al., 2015; Giacalone & Jaeger, 2016).

The number and choice of stimuli for studies restricted to one specific product category is generally a lot lower, with the majority of studies including 6–8 stimuli (Table 6.1). This is because the majority of these studies include other types of evaluations in the same test ballot, such as acceptability, sensory attributes, and product-related emotions (e.g., Cardello et al., 2016; Cardello & Schutz, 1996), so the number of stimuli to evaluate needs to be kept low, as to not interfere with the normal progress of the test. In this case, selecting an appropriate number of stimuli is generally based on the usual considerations for consumer tests (ballot length, the type of stimuli and the type of evaluation requested, avoiding fatigue, etc.).

Regarding the type of stimuli, blind samples are a common choice in CLT tests (Table 6.1), which makes the most sense when one wants to link appropriateness to sensory variation and acceptability. The use of pictorial images as stimuli is also increasingly seen in the literature, and with good reason: because vision is the most important sensory modality at the point of purchase, product appearance is often a very important cue for assessing the perceived usage appropriateness of products. Accordingly, this stimulus format has been effectively employed in studies investigating the influence of extrinsic product aspects, such as the packaging, in categories that depend strongly on visual inspection (Giacalone & Jaeger, 2016; Jaeger, Hedderley, & MacFie, 2001).

6.2.2.3 Usage contexts

Selection and number of usage contexts, as for the stimuli, will depend on the study objectives. The generation of usage contexts is a crucial aspect of the IBU method, although as noted by Jaeger and Porcherot (2017), the process is seldom described in much detail in the literature.

As a general rule, the usage contexts should cover all those that are representative and frequent for that product category, as well as any additional target usage that may be important to testers. A theoretical and practical challenge is obviously in defining and operationalizing usage contexts. Thus, a good way to develop a list of usage contexts is to take as one's point of departure the existing classifications of situational variables (Belk, 1975; Bisogni et al., 2007; Blake, Bisogni, Sobal, Devine, & Jastran, 2007; Meiselman, 2006, Meiselman, 2008; Wansink, 2004), and start developing relevant options for each class.

For example, in his landmark paper on situational influences on consumer behavior, Belk (1975) proposed a broad taxonomy of situational variables that include (1) *physical* and (2) *social surroundings*, (3) *temporal perspectives*, (4) *task definition*, and (5) *antecedent states*. Table 6.2 shows an example of usage classes from Belk's taxonomy with examples relevant for the product category "beer" (from Giacalone et al., 2015). Several lists of usage contexts for specific product categories can be found in the papers listed in Table 6.1.

Although existing literature is a good starting point, attribute generation could also be done with a bottom-up approach, for example, by using a small group of consumers to generate a list of usage situations. In general, some degree of consumer involvement in developing the list of items is highly recommended to ensure that the items included in the ballot are understood, and that no attributes that may be important to consumers are left out.

An important decision concerns the number of usage contexts to include. The key thing to keep in mind is that the number of food-usage context combinations will grow geometrically, so that a 10 by 10 matrix, such as the one in Fig. 6.4, will entail 100 evaluations, whereas a 50 by 50 matrix similar to those used in early applications will require 2500 evaluations on the part of each consumer. Obviously, the test situation (home, CLT, online), the stimuli type, and the possible presence of other tasks need to be considered, as it is crucial not to overburden the cognitive and affective system of

Table 6.2 Belk's (1975) taxonomy of usage situations

Physical surroundings • On a camping trip • Watching TV at home • At a pub	**Task definition** • To celebrate an achievement • As a gift for someone • For a special occasion
Social surroundings • To serve to guests • To impress someone • For women	**Antecedent states** • As a thirst quencher • When I want to relax • When I want something different
Temporal perspective • At a BBQ in the summer • For lunch • As an alternative to wine for dinner	

Examples of contextual attributes within each class are given for beer (from Giacalone et al., 2015).

the consumers. Thorough pilot testing will serve as the best indication of the expected elapsed time for completion, and therefore what is reasonable to include the ballot.

One final crucial consideration is the way the usage contexts are to be presented to the consumers. In the vast majority of studies, IBU contexts are presented as very short phrases (see examples in Table 6.2 and Fig. 6.4). While this way is efficient and easy to implement, the downside is that the corresponding mental image may not be very compelling and vivid in the mind of the respondents (Jaeger & Porcherot, 2017). Additionally, because consumers are asked to imagine a usage context in response to a written depiction, there is a risk that specific aspects of the context may vary quite substantially across consumers (Hein et al., 2010, 2012), and/or not match the intention of the researcher.

This is in fact not necessarily a negative feature. For example, Köster and Mojet (2015) suggest that subjectively interpreting usage contexts may often be more relevant for the consumers than evaluating a more objective definition of context, because it enables them to draw on their own previous personal experiences. However, from the researcher's perspective, a more objective definition of the usage context is often desirable, as it allows for a clearer interpretation of the results. If that is the case, more detailed descriptions and pictorial representations may be used to better convey the target usage context to consumers (Elzerman et al., 2015; Giacalone et al., 2015). For example, Fig. 6.6 shows two usage contexts from a study on situational appropriateness of beer (Giacalone et al., 2015) in the form of specific eating situations with an accompanying caption.

Formal comparisons of different stimuli formats in the context of IBU studies have not been carried out; however, visual stimuli are assumed to enhance the ecological validity of CLT-based studies of contexts (Cardello & Meiselman, 2018; Jaeger & Porcherot, 2017; Köster, 2003), and have the added benefit of reducing the ambiguity regarding what consumers have imagined during their appropriateness evaluations.

6.2.2.4 Respondents

Regarding how many consumers to include, Table 6.1 shows a considerable variation in the literature, with sample sizes ranging from as few as 30 consumers (Nantachai et al., 1991), to >400 (Piqueras-Fiszman & Jaeger, 2015a, 2015b). The *N* of respondents seem to have progressively increased in recent years, following a general trend in sensory and consumer toward larger and more representative sample sizes, as well as more attention to diversity and consumer segmentation (Meiselman, 2013).

It has been noted that IBU data present a low amount of inter-individual variability relative to other types of consumer evaluations. For example, Schutz (1994) reports that it is possible to achieve stable results, that is, reliable mean ratings and component structures, with as few as 25 consumers (with the 7-pt scale, but in my experience, that it is the case even with CATA). This is remarkable if compared with for example, hedonic data, where 100+ consumers are typically considered as the minimum sample size (Hough et al., 2006). This discrepancy can be attributed to the different nature of affective and cognitive evaluations; while hedonic data are affected by a multitude of factors that provide a large amount of individual variation, appropriateness data are

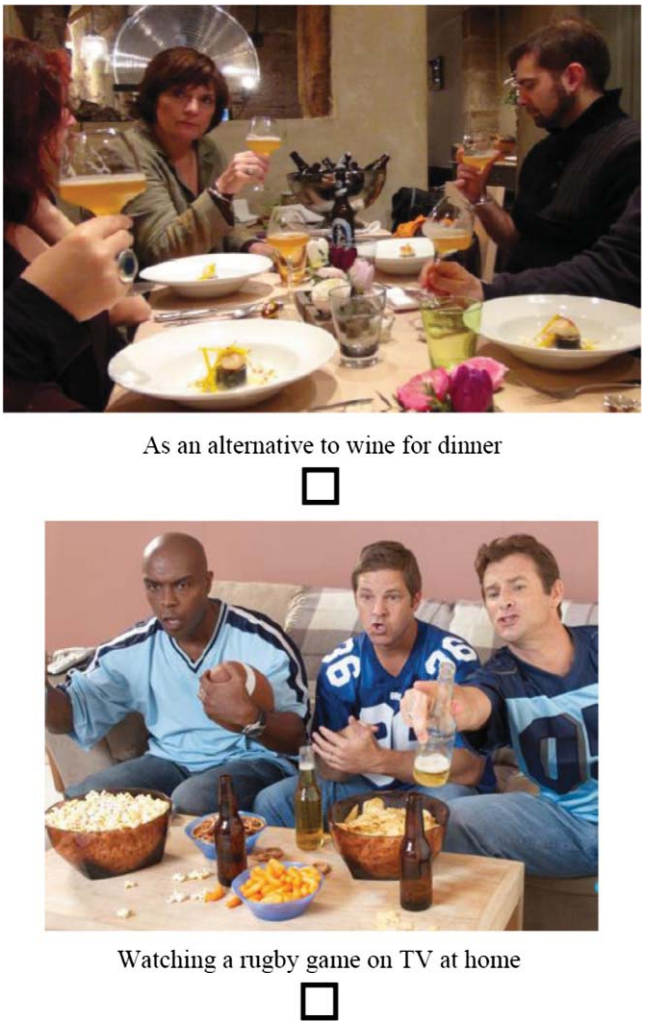

Fig. 6.6 Example of visual depictions of usage contexts in a IBU appropriateness task (in this case with a CATA format).
Reprinted with permission from Elsevier from Giacalone, D., Frøst, M. B., Bredie, W. L., Pineau, B., Hunter, D. C., Paisley, A. G., Beresford, M. K., & Jaeger, S. R. (2015). Situational appropriateness of beer is influenced by product familiarity. *Food Quality and Preference*, *39*, 16–27.

more likely to reflect the culturally agreed place that a food occupies within a certain consumer group. For example, it seems safe to assume that asking consumers how much they like beer would result in a higher variability in response than if the same consumers were asked to rate how appropriate they thought beer was "for breakfast." Accordingly, Giacalone and colleagues report that IBU data have very high

reproducibility also when replicating studies within different consumer groups, confirming the primarily coenotropic nature of appropriateness evaluations (Giacalone et al., 2015).

6.2.2.5 Analysis of appropriateness data

The analysis of IBU appropriateness data involves both univariate and multivariate techniques. The first step is generally to compute descriptive statistics for each product for each individual usage context, and to test for statistical differences between the products by an appropriate statistical test (e.g., ANOVA on product means when rating scales are used, Cochran's Q test when the CATA format is used). Univariate analyses such as these can be presented visually, for example, using bar or line charts, or numerically in a table where products are ranked by appropriateness supplemented by post-hoc testing comparing the means or proportions presented. These simple results are already of great practical value in a product development context, as they give a ranking of products enabling a quick evaluation of how each product performs in each usage situation.

Cardello and Schutz (1996) also proposed to supplement product means analyses with an overall distance metric (D) computed for each product pair. This can be done using the standardized euclidean distance formula:

$$D_{ij} = \frac{\sum_{k=1}^{n} \left(1 - \left[\left(X_{ik} - X_{jk}\right)^2 / X_{Max}^2\right]\right)}{n}$$

where D_{ij} is the distance between product i and product j, X_{ik} and X_{jk} are the mean appropriateness ratings for usage context k, X_{Max} is the largest possible appropriateness value (corresponding to the highest point on the scale or to 1 when proportions are used), and n is the number of usage contexts. The advantage of using this scaled distance measure is that it provides a global assessment of absolute similarity between products, not just a similarity of pattern (as, e.g., a correlation coefficient would). As observed by Schutz (1994), this may be particularly relevant to determine the viability of substituting one product for another; for example, when introducing a novel food from one culture to another. In practice, one should always supplement the overall analysis with an analysis at the individual usage context. Especially if the list of usage contexts is long, D_{ij} may often tend toward high value that may overestimate the similarity between two items.

Multivariate data analyses of IBU data are very useful to derive perceptual maps showing how consumers cognitively organize different food items and usages. Principal Component Analysis (PCA) on a table containing mean ratings across usage contexts has been the most common approach for rating data (e.g., Schutz, 1988, 1994; Schutz & Martens, 2001); whereas for the CATA format, the same is typically accomplished using correspondence analyses on a contingency table crossing products and usage contexts (e.g., Giacalone et al., 2015). Regardless of the analytical approach used, it is interesting to notice that multivariate analyses of appropriateness data

generally result in efficient representations and interpretable component structures. The proportion of variance explained in the first two components is generally high (70%–90%)—especially compared with what one usually gets from hedonic data—suggesting that patterns of product-usage associations are robust and widely held by consumers.

In addition to PCA and CA, multiblock approaches such as Partial Least Squares Regression (PLSR) can be used to relate appropriateness data to other types of data. For example, Hersleth et al. (2005) used a PLSR model to predict appropriateness of bread from sensory data, and Mejlholm and Martens (2006) used a three-block PLSR approach to predict consumer preferences for beer from appropriateness data, sensory profiles, and consumer background data simultaneously.

6.2.3 Other approaches to evaluating product appropriateness

As previously mentioned, the IBU approach is not the only CLT-based approach to evaluating situational appropriateness. For example, the Repertory Grid Method (RGM) (Kelly, 1955) has been used by several authors to investigate contextual aspects of food choice and acceptance (Jack et al., 1994; Jaeger et al., 2005; McEwan & Thomson, 1988; Monteleone, Raats, & Mela, 1997; Raats & Shepherd, 1991/1992; Scriven et al., 1989). Briefly, the RGM is an interview in which a consumer is presented with groups of three products (triads), two of which are arbitrarily associated with each other and dissociated from the third. The consumer is then asked to describe how (s) he thinks that the two associated products are similar, and how they are different from the third. The specific question may be tweaked to meet the specific intention of the researcher. For example, if situational appropriateness is in focus, consumers may be asked to describe in what usage contexts they would use the two associated products, but not the third (Gains, 1994). Once the consumer has exhausted all possible constructs, the task is repeated for the other two combinations in the triad, as well for any other triads, resulting in a large number of usage contexts that discriminate between the product under test. Then, consumers typically provide ratings for each product for each of the constructs they had previously individually generated. The data are then submitted to Generalized Procrustes Analysis (Gower, 1975) to derive a common perceptual map of the products from the individual ratings (or "repertory grids").

In a sense, RGM can be considered a more structured version of the Free Choice Profile (Jack & Piggott, 1991), which also has been used for eliciting situational appropriateness constructs from consumers (e.g., Piggott et al., 1990).

A lengthy comparison of all possible approaches is beyond the scope of this chapter. However, one point that is important to recognize is that approaches such as RGM and FCP are ideographic in nature, that is, they let each consumer develop their own individual set of usage contexts, whereas the IBU approach is nomothetic, where one set of usage contexts is evaluated by all consumers (Schutz, 1994). The IBU approach thus assumes that all usage contexts are relevant to all consumers, which may not necessarily be the case. To what degree this difference is important in practice is unclear, as studies that have compared the two approaches (Nantachai et al., 1991; Raats & Shepherd, 1991/1992) have generally concluded that they provide similar results.

Based on practical consideration, the IBU approach is certainly faster, more flexible, and overall better suited for large-scale CLT tests in a product development context, where the aim is to generalize to a population of interest. By contrast, RGM is an excellent method when one wants to really elucidate the cognitive structures consumers hold within a certain product category, but can quickly become impractical as soon as the number of consumers (and of products) increases. It might be best to think in terms of complementarities. Since the RGM is a process of contrast and similarities, it seems especially suited as a method for generating usage contexts to be further studied with the IBU approach, thus ensuring that usage contexts are as relevant as possible to the consumer population. For example, Nantachai et al. (1991) used repertory grid interviews to develop usage contexts for testing in a larger IBU study. Geertsen et al. (2016) used a version of projective mapping to the same effect.

Another methodological approach that has been proposed to measuring appropriateness is conjoint analysis, where usage contexts can be specified as design factors in a conjoint design, together with other factors of interest, such as sensory, packaging, information, and so forth (Almli et al., 2011; Jaeger, 2000; Jager & Rose, 2008). The significant advantage of the conjoint approach is that it allows one to estimate interaction effects between usage contexts and other experimental factors, which is of course of great interest. The downside is that, compared with the IBU approach, a much lower number of usage contexts can be included in a single study. The application context should be considered in this case. For example, in a product development context, the IBU approach may be more suited to early stage product development, when the focus is often interested in testing different product formulations. Conjoint methods appear advantageous when closer to market launch, where a few target usage contexts have been identified, and where blending intrinsic and extrinsic product elements is most important.

6.3 Current directions in appropriateness research

This section reviews recent advances in appropriateness research, as well as open venues for future research. Attention is given to three areas that have substantive interest for sensory and consumer researchers: (1) behavioral and attitudinal correlates of situational appropriateness, (2) consumer segmentation based on appropriateness data, and (3) the potential of immersive technology in appropriateness research.

6.3.1 Attitudinal and behavioral correlates of appropriateness

6.3.1.1 Choice, consumption, willingness to pay

Having already established that appropriateness is not identical to liking (Cardello et al., 2000; Cardello & Schutz, 1996), one important research question is the relevance of appropriateness for other behavioral and attitudinal outcomes of interest. Can appropriateness be used to predict consumer choice, frequency of consumption, and willingness to pay?

Regarding frequency of purchase, the already mentioned study by Sosa and collaborators (Sosa et al., 2005) suggested a strong link between IBU ratings of situational appropriateness and self-reported frequency of consumption of different seasoning sauces. This is in agreement with an earlier study by Shepherd, Schutz, and Sparks (1993), in which appropriateness evaluations of a 50×50 item-by-use matrix was found to account for 44% of the variance in frequency of use. It is interesting to note that in the same dataset, the authors report that preference (operationalized as the attribute "when I want something I really like" in the IBU matrix) could explain 12% of variance in frequency of use. Although both studies are based only on self-reported estimates of consumption frequency, these data are very indicative that appropriateness evaluations are predictive of repeat purchase and frequency of consumption. Studies on different product categories, and possibly using actual consumption frequency (e.g., using food recall diaries) would be very useful to confirm and extend these results.

A few studies have investigated the relationships between appropriateness and willingness to try and purchase. Elzerman and colleagues (Elzerman et al., 2015) reported that willingness to try meat substitutes were highly related to the appropriateness of the usage context (in that study, meal combinations). Another similarly motivated study was published by Lai (1991), who investigated consumers' willingness to adopt different variants of a canned beverage (Wulong tea) across different usage contexts. This study also found that consumers were more willingly to adopt products that could be seen as more situationally appropriate.

What about choice itself? As laid out in the introduction, the general marketing and consumer behavior literature supports the notion that consumers' anticipated usage is a strong determinant of consumer choice (Belk, 1975; Lai, 1991; Jager & Rose, 2008). Indeed, even if one took the view that consumers' choices are primarily hedonic-oriented, the near perfect correlation observed by appropriateness ratings and *expected* liking (Fig. 6.2) should point to a strong relationship between appropriateness and choice. To this day, surprisingly little attention has been devoted to empirically verifying and quantifying this relationship.

These effects were recently studied by Giacalone and Jaeger (2017) in a series of 14 studies spanning a wide range of product categories and usage contexts. In all studies, participants evaluated a set of stimuli chosen to represent fixed levels of IBU appropriateness (low, medium, high) in a target context, and completed a choice task using a 7-point choice likelihood scale, or a discrete choice using the best-worst scaling method. Linear regression results from all studies consistently indicated that the IBU level significantly predicted choice response. The results were robust with respect to variation in product category and experimental protocol. Effect sizes varied substantially between studies, with appropriateness level explaining from a minimum of 3% to >60% of variance in consumers' choice responses, and this variation was linearly related to the range in appropriateness in the product sets. Overall, the results strongly support the notion that IBU appropriateness is a significant predictor of food choice, and hence a meaningful product performance measure.

6.3.1.2 Appropriateness and emotions

Some authors have emphasized the important effect that situational appropriateness may have for the measurement of food-related emotions in consumer research, a topic that has enjoyed increased attention in recent years (Köster & Mojet, 2015; Meiselman, 2015). Such effects were elegantly elucidated in a series of papers by Piqueras-Fiszman and Jaeger (2014a, 2014b, 2014c, 2015a, 2015b): using a variety of methods and products. These authors demonstrated that the usage context affected consumer emotional responses toward food products, and specifically that the frequency of use of positive (*vs.* negative) emotion terms strongly depended on whether a product was considered appropriate (*vs.* inappropriate) for the target usage context.

This research has obviously important implications for emotion measurement. Although sometimes one may want to separate the product effect from the context effect, neglecting the situational component of emotional responses hinders the ecological validity of the results. Therefore, a future research direction with both theoretical and practical interest would then be to study how individual aspects of the usage context (e.g., time of the day, presence of others, etc.) affect perceived appropriateness and emotional responses. The afore-mentioned papers (in particular Piqueras-Fiszman & Jaeger, 2015a) already showcase examples of how the impact of variations in usage context dimensions on appropriateness can be quantified.

6.3.2 Situation-based consumer segmentation

As observed earlier, appropriateness data tend to have high reproducibility and a lower level of variability than liking or other attitudinal data. However, inter-individual variation occurs, and there might be situations in which one may want to study it.

Extending the reasoning that appropriateness of food is socially determined, culture or ethnic groups can be expected to be a major source of variation. This is supported by extant data (e.g., Baird & Schutz, 1976; Jaeger, 2000; Nantachai et al., 1991). For example, Jaeger (2000) demonstrated differential patterns of situational appropriateness for apples between New Zealand and Samoan consumers related to the cultural orientation (individualism *vs.* collectivism) between these two countries.

Within a specific culture, Rucker and Schutz (1982) have presented examples of consumer segmentation on the basis of a factor analysis of appropriateness data, concluding that this data can be used to segment consumers in a meaningful way. It is not known where these segments correspond to any predefined data in terms of demographic, psychographic, or behavioral variables. In the already mentioned paper by Sosa and collaborators (Sosa et al., 2005) focusing on the relationships between degree of appropriateness and (self-reported) frequency of consumption, the authors found appropriateness to be highly correlated to consumption when considering averaged data, but also a significant spread in individual responses. However, segmenting consumers by demographics did not improve prediction, leading the authors to conclude that the effect of appropriateness on consumption frequencies must be mediated

by non-demographic factors. Evidently, understanding the specific composition of an appropriateness-based consumer segment would be interesting in the context of both product development and marketing, as well in other areas in which understanding product-usage patterns for specific consumer groups is of interest (e.g., in a public health context). Nevertheless, this aspect has not yet been explicitly addressed by extant research.

In general, constructs relating to general attitudes toward food (e.g., food neophobia), and to previous experience with a product category (e.g., product knowledge, involvement, etc.) are more likely to affect appropriateness evaluation. For example, Giacalone and Jaeger (2016) suggest that consumers who are knowledgeable about a product may process product elements more confidently, and therefore infer appropriate usage contexts more easily compared with less experienced consumers (e.g., a wine connoisseur may quickly envision food pairing of a wine based on grape variety or vintage, whereas a novice would not). The next section reviews recent advances in this area that may explain why this is the case.

6.3.2.1 *Familiarity as a moderator of appropriateness*

Previous studies have suggested that consumers' familiarity with a product may be an important factor in shaping appropriateness of use evaluations (e.g., Jaeger et al., 2005; Tuorila et al., 1994).

The influence of product familiarity in explaining consumers' evaluations of appropriateness of use has been recently investigated in two papers by Giacalone, Jaeger, and collaborators (Giacalone et al., 2015; Giacalone & Jaeger, 2016), which demonstrated that product familiarity is positively related to product versatility, defined as the total number of situations a product is deemed appropriate for. In other words, as product familiarity increases, products are perceived as appropriate for a larger number of uses. Conversely, consumers experienced difficulty in identifying appropriate uses for unfamiliar products. This finding may explain the relationship between appropriateness and frequency of purchase (Sosa et al., 2005; Shepherd et al., 1993). Generally, consumers have little incentive to sample unfamiliar foods and beverages, and therefore familiar products may be routinely chosen because they are more readily available, and more likely to enter the consideration set because of, for example, previous satisfactory consumption experiences. Thus, it makes sense that product familiarity is a strong predictor of repeat purchase, at least when consumers make generic food provisioning decisions ("stocking up" the household pantry with no particular end usage in mind) (Giacalone et al., 2015).

These results also suggest that consumers may find it challenging to envisage how unfamiliar food products can be incorporated into their existing dietary practices, resulting in food-related usage patterns that are resistant to change. This is, of course, a major barrier to consumer adoption of new products.

In the future, it would be interesting to elucidate the mechanisms by which unfamiliar products are perceived as less appropriate. One possibility is that familiarity breeds comfort, making existing usage patterns and routines more desirable.

Additionally, consumers might worry that an unfamiliar product might not taste as expected, and be a potential waste of money. The cognitive effort associated with evaluating unfamiliar products or product features may itself be a source of negative bias against new products. For food and beverage specifically, risk perception may be a factor. For example, consumers may think that familiar products are healthier and safer. Additionally, the potential social embarrassment coming from poor choices (e.g., serving a product that is contaminated, or simply does not taste as expected) might also lead consumers to rate unfamiliar products as generally less appropriate. Uncovering these motives should be a focus in future appropriateness research, and it would be especially needed in understanding how this negative bias toward novel products can be overridden. Giacalone and Jaeger (2016) also suggest devoting greater attention to retail-level strategies that can suggest product usage, such as free trials (useful if the low appropriateness is rooted in risk aversion), and a goal-based shelf display (i.e., where products are organized around intended uses or benefits), which may be a good strategy to improve appropriateness in a usage context not currently associated with that product (e.g., reduced alcohol wine in the health food section).

6.3.3 Immersive technologies in appropriateness research

Immersive technologies, where consumers experience and interact with a computer generated virtual environment, are currently a major trend in sensory and consumer science. The possibility of conducting product evaluations in a fully immersive virtual environment is clearly a hotbed for context research, and is seen as a major opportunity to enhance the external validity of CLT tests (Jaeger & Porcherot, 2017). The application of VR is dealt with extensively in this book, and the reader is referred to Chapters 16 and 23 for in-depth discussions.

A relevant question to ask here is what does the increasing availability of immersive technologies mean for appropriateness evaluations? This is difficult to answer sensibly, because at the time of writing, the VR technology is still in its infancy, and we probably need more experience to fully understand its value in sensory and consumer science. In general, VR technologies should offer interesting possibilities for addressing two of the perceived drawbacks of appropriateness research related to the elicitation of usage contexts in response to written descriptions (see Section 6.2.2.3): the low level of immersiveness, and the potential misalignment of these mental images between consumers (as well as between consumers and researchers). In principle, the use of VR technology can remedy both.

That aside, in this author's opinion, IBU and VR do rather different things, and are likely to have different uses. The VR approach seems most useful when the goal is go in depth into one (or a few) target contexts that are of interest to the researcher or product developer, or when that context is actually the only one relevant to that product category, for example, a car cabin for evaluation of a car air freshener. Here, the VR approach can be very useful to provide results (e.g., liking, willingness to pay) with a higher external validity than a context-neutral CLT-test. From this perspective, the VR approach is more in line with the tradition of conducting situational research in

natural or physically manipulated environments (e.g., Bell et al., 1994; Di Monaco et al., 2014; Petit & Sieffermann, 2007; Sester et al., 2013).

While the IBU approach cannot replace real contexts to the same degree as (in principle) VR can, it affords the possibility of evaluating a larger number of products and usage contexts in the same session. This makes it especially useful for evaluating the performance of different products across different usage contexts (do consumers understand how to use the products? Which benefits do each product deliver, which products may be replaced or repositioned, etc.), and their perceived versatility. This information is, of course, most useful for products that have a variety of applicable contexts (typically the case for food and beverages), and when one is interested in understanding the cognitive structures pertaining to product usage within a specific product category.

From a practical perspective, the IBU method overall seems much better suited for the context of current CLT tests, which, for the most part, involve actual samples, because when it comes to eating while immersed in a VR environment, reaching a high enough level of realism and usability still presents a significant challenge.

6.4 Conclusions

This chapter has discussed situational appropriateness in the context of sensory and consumer research, with particular focus on the IBU method proposed by Schutz (1988, 1994) as an adjunct to CLT evaluations. Appropriateness, intended as perceived fit with a specific usage situation, is a basic product-related construct distinct from other affective and attitudinal measures. The end-goal of CLT tests should be to ensure that the test products not only score high on absolute performance indicators (acceptability, willingness to pay, etc.), but also high appropriateness for the usage situation envisioned by the product developer. As recently emphasized by Jaeger and Porcherot (2017), product understanding cannot be complete without appropriateness for use evaluations. While the IBU approach cannot fully replace real contexts in terms of external validity, it does provide an easy and flexible way to gather contextual information in a CLT setting. Like any method, the IBU approach has its own pros and cons, which should be weighted based on the specific goals against alternative approaches (including CLT, non-CLT, and immersive technologies) to investigate situational influences in product-related research.

Recent research has indicated that IBU appropriateness ratings (1) may be a strong predictor of consumer food choices, willingness to use, and reported intake, and (2) underpin other product-related variables, such as emotional responses and perceived product uniqueness. Taken collectively, these studies indicate that appropriateness should be considered as an important product performance criterion in CLT testing. Given the complexity of consumer behavior, the ability to predict product success must ultimately encompass a combination of product performance indicators. Situational appropriateness is, however, an important part of this mix that should not be overlooked.

References

Almli, V. L., Næs, T., Enderli, G., Sulmont-Rossé, C., Issanchou, S., & Hersleth, M. (2011). Consumers' acceptance of innovations in traditional cheese. A comparative study in France and Norway. *Appetite*, *57*, 110–120.

Bach, V., Kidmose, U., Thybo, A. K., & Edelenbos, M. (2013). Sensory quality and appropriateness of raw and boiled Jerusalem artichoke tubers (*Helianthus tuberosus* L.). *Journal of the Science of Food and Agriculture*, *93*, 1211–1218.

Bach, V., Mikkelsen, L., Kidmose, U., & Edelenbos, M. (2015). Culinary preparation of beetroot (*Beta vulgaris* L.): The impact on sensory quality and appropriateness. *Journal of the Science of Food and Agriculture*, *95*, 1852–1859.

Baird, P. C., & Schutz, H. G. (1976). The marketing concept applied to "selling" good nutrition. *Journal of Nutrition Education*, *8*, 13–17.

Belk, R. W. (1975). Situational variables and consumer behavior. *Journal of Consumer Research*, *2*, 157–164.

Bell, R., Meiselman, H. L., Pierson, B. J., & Reeve, W. G. (1994). Effects of adding an Italian theme to a restaurant on the perceived ethnicity, acceptability, and selection of foods. *Appetite*, *22*, 11–24.

Bisogni, C. A., Winter Falk, L., Madore, E., Blake, C. E., Jastran, M., Sobal, J., et al. (2007). Dimensions of everyday eating and drinking episodes. *Appetite*, *48*, 218–231.

Blake, C. E., Bisogni, C. A., Sobal, J., Devine, C. M., & Jastran, M. (2007). Classifying foods in contexts: How adults categorize foods for different eating settings. *Appetite*, *49*, 500–510.

Bruhn, C. M., & Schutz, H. G. (1986). Consumer perceptions of dairy and related-use foods. *Food Technology*, *40*, 79–85.

Cardello, A. V., & Meiselman, H. L. (2018). Contextual influences on consumer responses to food products. In: G. Ares & P. Varela (Eds.), *vol. 2 Methods in consumer research* (pp. 3–54): Elsevier.

Cardello, A. V., Pineau, B., Paisley, A. G., Roigard, C. M., Chheang, S. L., & Guo, L. F. (2016). Cognitive and emotional differentiators for beer: An exploratory study focusing on "uniqueness" *Food Quality and Preference*, *54*, 23–38.

Cardello, A. V., Schutz, H., Snow, C., & Lesher, L. (2000). Predictors of food acceptance, consumption and satisfaction in specific eating situations. *Food Quality and Preference*, *11*, 201–216.

Cardello, A. V., & Schutz, H. G. (1996). Food appropriateness measures as an adjunct to consumer preference/acceptability evaluation. *Food Quality and Preference*, *7*, 239–249.

de Andrade, J. C., de Aguiar Sobral, L., Ares, G., & Deliza, R. (2016). Understanding consumers' perception of lamb meat using free word association. *Meat Science*, *117*, 68–74.

Di Monaco, R., Giacalone, D., Pepe, O., Masi, P., & Cavella, S. (2014). Effect of social interaction and meal accompaniments on acceptability of sourdough prepared croissants: An exploratory study. *Food Research International*, *66*, 325–331.

Elzerman, J. E., Hoek, A. C., van Boekel, M. A., & Luning, P. A. (2011). Consumer acceptance and appropriateness of meat substitutes in a meal context. *Food Quality and Preference*, *22*, 233–240.

Elzerman, J. E., Hoek, A. C., van Boekel, M. J., & Luning, P. A. (2015). Appropriateness, acceptance and sensory preferences based on visual information: A web-based survey on meat substitutes in a meal context. *Food Quality and Preference*, *42*, 56–65.

Elzerman, J. E., van Boekel, M. A., & Luning, P. A. (2013). Exploring meat substitutes: Consumer experiences and contextual factors. *British Food Journal*, *115*, 700–710.

Gains, N. (1994). The repertory grid approach. In H. J. H. MacFie & D. M. H. Thomson (Eds.), *Measurement of food preferences* (pp. 51–76). Boston, MA: Springer.

Geertsen, J., Allesen-Holm, B. H., & Giacalone, D. (2016). Consumer-led development of novel sea-buckthorn based beverages. *Journal of Sensory Studies*, *31*, 245–255.

Giacalone, D. (2018). Product performance optimization. In: G. Ares & P. Varela (Eds.), *vol. 1 Methods in consumer research* (pp. 159–185): Elsevier.

Giacalone, D., Frøst, M. B., Bredie, W. L., Pineau, B., Hunter, D. C., Paisley, A. G., et al. (2015). Situational appropriateness of beer is influenced by product familiarity. *Food Quality and Preference*, *39*, 16–27.

Giacalone, D., & Jaeger, S. R. (2016). Better the devil you know? How product familiarity affects usage versatility of foods and beverages. *Journal of Economic Psychology*, *55*, 120–138.

Giacalone, D., & Jaeger, S. R. (2017). Situational appropriateness (item-by-use): an overlooked predictor of food choice. *12th Pangborn Sensory Science Symposium, 20–24 August 2017, Providence, RI, USA*.

Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*, *40*, 33–51.

Hartwell, H. J., Shepherd, P. A., Edwards, J. S., & Johns, N. (2016). What do patients value in the hospital meal experience? *Appetite*, *96*, 293–298.

Hein, K. A., Hamid, N., Jaeger, S. R., & Delahunty, C. M. (2010). Application of a written scenario to evoke a consumption context in a laboratory setting: Effects on hedonic ratings. *Food Quality and Preference*, *21*, 410–416.

Hein, K. A., Hamid, N., Jaeger, S. R., & Delahunty, C. M. (2012). Effects of evoked consumption contexts on hedonic ratings: A case study with two fruit beverages. *Food Quality and Preference*, *26*, 35–44.

Hersleth, M., Berggren, R., Westad, F., & Martens, M. (2005). Perception of bread: A comparison of consumers and trained assessors. *Journal of Food Science*, *70*, 95–101.

Hersleth, M., Lengard, V., Verbeke, W., Guerrero, L., & Næs, T. (2011). Consumers' acceptance of innovations in dry-cured ham: Impact of reduced salt content, prolonged aging time and new origin. *Food Quality and Preference*, *22*, 31–41.

Hough, G., Wakeling, I., Mucci, A., Chambers, E., IV, Gallardo, I. M., & Alves, L. R. (2006). Number of consumers necessary for sensory acceptability tests. *Food Quality and Preference*, *17*, 522–526.

Jack, F. R., & Piggott, J. R. (1991). Free choice profiling in consumer research. *Food Quality and Preference*, *3*, 129–134.

Jack, F. R., Piggott, J. R., & Patterson, A. (1994). Use and appropriateness in cheese choice, and an evaluation of attributes influencing appropriateness. *Food Quality and Preference*, *5*, 281–290.

Jaeger, S. R. (2000). Uncovering cultural differences in choice behaviour between Samoan and New Zealand consumers: A case study with apples. *Food Quality and Preference*, *11*, 405–417.

Jaeger, S. R., Cardello, A. V., Chheang, S. L., Beresford, M. K., Hedderley, D. I., & Pineau, B. (2017a). Holistic and consumer-centric assessment of beer: A multi-measurement approach. *Food Research International*, *99*, 287–297.

Jaeger, S. R., Cardello, A. V., Jin, D., Hunter, D. C., Roigard, C. M., & Hedderley, D. I. (2017b). Product uniqueness: Further exploration and application of a consumer-based methodology. *Food Quality and Preference*, *60*, 59–71.

Jaeger, S. R., Giacalone, D., Roigard, C. M., Pineau, B., Vidal, L., Giménez, A., et al. (2013). Investigation of bias of hedonic scores when co-eliciting product attribute information using CATA questions. *Food Quality and Preference*, *30*, 242–249.

Jaeger, S. R., Hedderley, D., & MacFie, H. J. (2001). Methodological issues in conjoint analysis: A case study. *European Journal of Marketing*, *35*, 1217–1239.
Jaeger, S. R., Hort, J., Porcherot, C., Ares, G., Pecore, S., & MacFie, H. J. H. (2017). Future directions in sensory and consumer science: Four perspectives and audience voting. *Food Quality and Preference*, *56*, 301–309.
Jaeger, S. R., & Porcherot, C. (2017). Consumption context in consumer research: Methodological perspectives. *Current Opinion in Food Science*, *15*, 30–37.
Jaeger, S. R., Rossiter, K. L., & Lau, K. (2005). Consumer perception of novel fruit and familiar fruit: A repertory grid application. *Journal of the Science of Food and Agriculture*, *85*, 480–488.
Jager, S. R., & Rose, J. M. (2008). Stated choice experimentation, contextual influences and food choice: A case study. *Food Quality and Preference*, *19*, 539–564.
Kelly, G. A. (1955). *The psychology of personal construct*. New York, NY: WW & Norton Co.
Köster, E. P. (2003). The psychology of food choice: Some often encountered fallacies. *Food Quality and Preference*, *14*, 359–373.
Köster, E. P., & Mojet, J. (2015). From mood to food and from food to mood: A psychologycal perspective on the measurement of food related emotions in consumer research. *Food Research International*, *76*, 180–191.
Lähteenmäki, L., & Tuorila, H. (1995). Consistency of liking and appropriateness ratings and their relation to consumption in a product test of ice cream. *Appetite*, *25*, 189–198.
Lähteenmäki, L., & Tuorila, H. (1997). Item-by-use appropriateness of drinks varying in sweetener and fat content. *Food Quality and Preference*, *8*, 85–90.
Lai, A. (1991). Consumption situation and product knowledge in the adoption of a new product. *European Journal of Marketing*, *25*, 55–67.
Lau, D., Hanada, L., Kaminskyj, O., & Krondl, M. (1979). Predicting food use by measuring attitudes and preference. *Food Product Development*, *13*, 66–72.
Marshall, D. (1993). Appropriate meal occasions: Understanding conventions and exploring situational influences on food choice. *International Review of Retail, Distribution and Consumer Research*, *3*, 279–301.
Martens, M., Schutz, H. G., Risvik, E., Rodbotten, M., & Raundhardt, O. (1987). Consumer perception of nutritional value related to other quality attributes and to chemical components. In J. Solms, D. R. Booth, & R. M. Pangborn (Eds.), *Food acceptance and nutrition* (pp. 321–355). London, UK: Academic Press.
McEwan, J. A., & Thomson, D. M. H. (1988). An investigation of the factors influencing consumer acceptance of chocolate confectionery using the repertory grid method. In D. M. H. Thomson (Ed.), *Food acceptability* (pp. 347–362). London, UK: Elsevier.
Meiselman, H. L. (2006). The role of context in food choice, food acceptance, and food consumption. R. Shepherd & M. Raats (Eds.), *The Psychology of Food Choice: vol. 3. Frontiers in nutritional science* (pp. 179–200). Wallingford, UK: CABI.
Meiselman, H. L. (2008). Experiencing food products within a phisycal and social context. In H. N. J. Schifferstein & P. Hekkert (Eds.), *Product Experience*. Amsterdam, NL: Elsevier.
Meiselman, H. L. (2013). The future in sensory/consumer research: Evolving to a better science. *Food Quality and Preference*, *27*, 208–214.
Meiselman, H. L. (2015). A review of the current state of emotion research in product development. *Food Research International*, *76*, 192–199.
Meiselman, H. L., Johnson, J. L., Reeve, W., & Crouch, J. E. (2000). Demonstration of the influence of the eating environment on food acceptance. *Appetite*, *35*, 231–237.
Mejlholm, O., & Martens, M. (2006). Beer identity in Denmark. *Food Quality and Preference*, *17*, 108–115.

Monteleone, E., Raats, M. M., & Mela, D. J. (1997). Perceptions of starchy food dishes: Application of the repertory grid method. *Appetite*, *28*, 255–265.
Nantachai, K., Petty, M. F., & Scriven, F. M. (1991). An application of contextual evaluation to allow simultaneous food product development for domestic and export markets. *Food Quality and Preference*, *3*, 13–22.
Petit, C., & Sieffermann, J. M. (2007). Testing consumer preferences for iced-coffee: Does the drinking environment have any influence? *Food Quality and Preference*, *18*, 161–172.
Piggott, J. R., Sheen, M. R., & Apostolidou, S. G. (1990). Consumer perceptions of whiskies and other alcoholic products. *Food Quality and Preference*, *2*, 177–185.
Piqueras-Fiszman, B., & Jaeger, S. R. (2014a). The impact of evoked consumption contexts and appropriateness on emotion responses. *Food Quality and Preference*, *32*, 277–288.
Piqueras-Fiszman, B., & Jaeger, S. R. (2014b). Emotion responses under evoked consumption contexts: A focus on the consumers' frequency of product consumption and the stability of responses. *Food Quality and Preference*, *35*, 24–31.
Piqueras-Fiszman, B., & Jaeger, S. R. (2014c). The impact of the means of context evocation on consumers' emotional associations towards eating occasions. *Food Quality and Preference*, *37*, 61–70.
Piqueras-Fiszman, B., & Jaeger, S. R. (2015a). The effect of product–context appropriateness on emotion associations in evoked eating occasions. *Food Quality and Preference*, *40*, 49–60.
Piqueras-Fiszman, B., & Jaeger, S. R. (2015b). Emotions associated with mealtimes: Memorable meals and typical evening meals. *Food Research International*, *76*, 243–252.
Raats, M. M., & Shepherd, R. (1991/1992). An evaluation of the use and perceived appropriateness of milk using the repertory grid method and the 'item-by-use' appropriateness method. *Food Quality and Preference*, *3*, 89–100.
Ratneshwar, S., & Shocker, A. D. (1991). Substitution in use and the role of usage context in product category structures. *Journal of Marketing Research*, *28*, 281–295.
Resurreccion, A. V. (1986). Consumer use patterns for fresh and processed vegetable products. *International Journal of Consumer Studies*, *10*, 317–332.
Rosas-Nexticapa, M., Angulo, O., & O'Mahony, M. (2005). How well does the 9-point hedonic scale predict purchase frequency? *Journal of Sensory Studies*, *20*, 313–331.
Rucker, M. H., & Schutz, H. G. (1982). Development of consumer typologies from appropriateness ratings. *Proceedings of the 1982 Academy of Marketing Science (AMS) Annual Conference*, Cham: Springer587.
Schutz, H. G., Fridgen, J. D., & Damrell, J. D. (1975). Consumer perceptions of rice and related products. *Journal of Food Science*, *40*, 277–281.
Schutz, H. G. (1988). Beyond preference: Appropriateness as a measure of contextual acceptance of food. In D. M. H. Thomson (Ed.), *Food acceptability* (pp. 115–134). Amsterdam, NL: Elsevier.
Schutz, H. G. (1994). Appropriateness as a measure of the cognitive-contextual aspects of food acceptance. In D. M. H. Thomson (Ed.), *Measurement of food preferences* (pp. 25–50). London: Blackie Academic.
Schutz, H. G. (1995). Eating situations, food appropriateness, and consumption. In B. M. Marriott (Ed.), *Not eating enough: Overcoming underconsumption of military operational rations* (pp. 341–359). Washington, DC: National Academies Press.
Schutz, H. G. (1999). Consumer data – Sense and nonsense. *Food Quality and Preference*, *10*, 245–251.
Schutz, H. G., Cardello, A. V., & Winterhalter, C. (2005). Perceptions of fiber and fabric uses and the factors contributing to military clothing comfort and satisfaction. *Textile Research Journal*, *75*, 223–232.

Schutz, H. G., & Jaeger, S. R. (2010). Innovation in sensory practices and education: An interview with Howard Schutz. In S. R. Jaeger & H. MacFie (Eds.), *Consumer-driven innovation in food and personal care products* (pp. 102–131). Oxford, UK: Woodhead Publishing.

Schutz, H. G., & Martens, M. (2001). Appropriateness as a cognitive-contextual measure of food attitudes. In L. Frewer, E. Risvik, & H. Schifferstein (Eds.), *Food, people and society* (pp. 247–266). Berlin: Springer.

Schutz, H. G., & Ortega, J. H. (1974). Consumer attidues towards wine. *American Journal of Enology and Viticulture*, *25*, 33–38.

Schutz, H. G., & Phillips, B. A. (1976). Consumer perceptions of textiles. *Family and Consumer Sciences Research Journal*, *5*, 2–14.

Schutz, H. G., & Rucker, M. H. (1975). A comparison of variable configurations across scale lengths: An empirical study. *Educational and Psychological Assessment*, *35*, 319–324.

Schutz, H. G., Rucker, M. H., & Russell, G. F. (1975). Food and food-use classification systems. *Food Technology*, *29*, 50–64.

Scriven, F. M., Gains, N., Green, S. R., & Thomson, D. M. H. (1989). A contextual evaluation of alcoholic beverages using the repertory grid method. *International Journal of Food Science & Technology*, *24*, 173–182.

Sester, C., Deroy, O., Sutan, A., Galia, F., Desmarchelier, J.-F., Valentin, D., et al. (2013). "Having a drink in a bar": An immersive approach to explore the effects of context on food choice. *Food Quality and Preference*, *28*, 23–31.

Shepherd, R., Schutz, H. G., & Sparks, P. (1993). Prediction of frequency of food consumption from use attitude and demographic variables. *Food Quality and Preference*, *4*, 89.

Sosa, M., Martínez, C., Arruiz, F., Hough, G., & Mucci, A. (2005). Degree of appropriateness and frequency of consumption of mayonnaise, ketchup, mustard and similar sauces in Argentina. *Food Quality and Preference*, *16*, 667–674.

Stefflre, V. (1971). Some eliciting and computational procedures for descriptive semantics. In P. Kay, P. Reich, & M. McLaran (Eds.), *Explorations in mathematical anthropology*. Cambridge, MA: MIT Press.

Stolzenbach, S., Bredie, W. L., Christensen, R. H., & Byrne, D. V. (2016). Understanding liking in relation to sensory characteristics, consumer concept associations, arousal potential and "appropriateness for use" using apple juice as an application. *Journal of Sensory Studies*, *31*, 135–142.

Thomson, D. M. H., & McEwan, J. M. (1988). An application of the repertory grid method to investigate consumer perceptions of food. *Appetite*, *10*, 181–193.

Tuorila, H., Hyvönen, L., & Vainio, L. (1994). Pleasantness of cookies, juice and their combinations rated in brief taste tests and following ad libitum consumption. *Journal of Sensory Studies*, *9*, 205–216.

Tversky, A., Sattah, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*, 371–384.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, *39*, 1179–1189.

Wansink, B. (2004). Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annual Review of Nutrition*, *24*, 455–479.

Warlop, L., & Ratneshwar, S. (1993). The role of usage context in consumer choice: A problem solving perspective. *Advances in Consumer Research*, *20*, 377–382.

White, F. D., Resurreccion, A. V. A., & Lillard, D. A. (1988). Consumer characteristics, food related attitudes and consumption of meat and meat products. *Nutrition Research*, *8*, 1333–1344.