# The pursuit of ecological validity through contextual methodologies

Adriana Galinanes Plaza, Julien J. Delarue, Laure Saulais

# The pursuit of ecological validity through contextual methodologies

**AUTHOR NAMES**

Galiñanes Plaza, A.[a,b], Delarue, J.[a,*], Saulais, L.[b,c]

**AFFILIATIONS**

[a]UMR Ingénerie Procédes Aliments, AgroParisTech, INRA, Université Paris-Saclay, 91300 Massy, France

[b]Center for Food and Hospitality Research, Institut Paul Bocuse, Chateau du Vivier, BP 25, 69131 Ecully Cedex, France

[c] UMR GAEL, INRA, F-38000 Grenoble, France

[*]Corresponding Author:

*E-mail address:* julien.delarue@agroparistech.fr; tel. +33 (0) 1 69 93 50 10; Fax +33 (0) 1 69 93 51 74. UMR Ingénerie Procédés Aliments, AgroParisTech, INRA, Université Paris-Saclay, 91300 Massy, France.

Contents

## 1. Introduction and background

It is vastly recognized that context impacts consumers' liking and choice of food, with direct implications regarding the validity of measures of the latter obtained in a given context. This issue is key for the food industry, whose strategic choices require reliable models of consumers' liking and behavior in order to predict the commercial success of a product. Yet, the everyday practice of consumer tests appears very heterogeneous regarding the inclusion of context variables, which may contribute to the low reliability of hedonic data used in the industry.

Since Meiselman in 1992 proposed to study real foods in real contexts (Meiselman, 1992), several studies have been conducted in natural consumption settings in an effort to improve the ecological validity of consumer data used in sensory science (Bell & Pliner, 2003; de Castro, 1994; Hetherington, Anderson, Norton, & Newson, 2006; Marshall & Bell, 2003). However, the gain in realism of studies in natural contexts is obtained to the detriment of control over context variables, questioning the reproducibility and transferability of the results.

In the past decades, several approaches have been developed in order to fill the gap between laboratory and natural contexts. They encompass evoked context studies, immersive technologies or the use of virtual reality. These approaches are intended to provide richer contextual realism to standard laboratory approaches by playing on contextual variables such as the physical or social contexts, or by using advanced technology in the case of the virtual reality. However, there are no standardized criteria to determine the type of variable that should or should not be added, and how and when they should be. Therefore, the question of validity and transferability of the data obtained in such conditions remains.

Based on a narrative review, this article discusses the added value of contextual approaches to increase the validity of consumer and sensory data. We argue that the addition of contextual cues in experimental approaches should be based on sufficient experimental evidence gathered within a clear theoretical framework. This review examines the notion of validity and ecological validity through the prism of different experimental disciplines (and particularly consumer psychology and behavioral economics) and draws some implications for sensory and consumer science. We review the recent research on context studies and the effect of context on consumers' liking, choice and intake. We also discuss the use of contextual variables in laboratory settings and the emerging use of new methodologies.

This article sets out to (1) define an analytical framework for assessing the relevance of moving towards more ecological validity; (2) assess evidence on how contextual effects should be taken into account in sensory and consumer science studies; and (3) identify the conditions and potential critical points for the design of experiments that take into account context to ensure ecological validity.

## 2. The concept of validity in sensory and consumer studies

### 2.1.Evaluating the validity of an experiment: internal, external and ecological validity

The experimental approach is used in various scientific fields concerned with individual behaviors. In particular, consumer psychology and economics use experiments to investigate consumer behaviors and preferences. In these fields, the role of theory in the experimental approach is significant, although not systematic (for a discussion on the role of theory in experimental economics, the reader is directed to Card, DellaVigna & Malmendier, 2011). For instance, experiments in economics aim to either (i) test theoretical assumptions, (ii) generate data on a little known phenomenon or (iii) evaluate the potential impact of policy scenarios or private sector innovations (Saulais, Muller & Lesgards, 2017). In consumer psychology, experiments use conceptual models and psychology theories (Kempen et al., 2017, Köster, 2009,) such as the Theory of Planned Behavior or the Expectancy-value theory (Ajzen, 1991). While studies in sensory science share this overall goal of better understanding consumer behavior, they often focus on operational objectives, such as to support product development through consumer tests.

In the various scientific fields relating to consumer science, experiments range from controlled, standardized laboratory experiments (standard approach) to natural experiments (experiments run in natural contexts), including different types of field experiments or field data (for more detailed information on field experiments, the reader is directed to the seminal papers by Carpenter et al., 2004 and Harrisson & List, 2004; and to the Fréchette & Schotter, 2015 (Part IV: The Lab and the Field) for a more recent view).

In general, the validity of experimental data is assessed from two complementary perspectives: internal and external validity. While internal validity refers to the ability of experimental data to provide understanding and to explain the causal relations within an experiment, external validity refers to the ability of the results of a given experiment to be generalized to other situations (Guala, 2012; Roe & Just, 2009). Therefore, moving from controlled to natural experiments implies a tradeoff between these two perspectives.

Ecological validity refers to the representation of the studied stimuli in an environment. This concept was introduced by Egon Brunswik in the area of the psychology of perception (Brunswik, 1943; Brunswik, 1955). “Representative design” addresses the ecological validity issue by considering a stimuli representative of the organism-environment relation. Brunswik therefore proposes to move from the study of people to the study of situations, replacing proper sampling of participants with representative sampling of a situation or task; and moving from “artificial” to “natural” contexts (Diehl, Wahl, & Freund, 2017). On the other hand, Brofenbrenner (1977) also includes the role of the researcher in the definition of ecological validity. The degree of ecological validity may be determined

by the researchers who should ensure that the environment experienced by the subjects has similar properties to the context of interest.

The ecological validity of a study thus depends on whether the task performed in an experimental context is relevant in the context of interest. If a researcher runs an experiment in the context of interest without modifying the ecology of that particular context, the internal validity as well as the ecological validity of that context can be ensured. However, if the researcher runs an experiment in a context that highly differs from the context of interest or has to modify it to establish internal validity, the inferences for ecological validity may not be guaranteed.

This definition of ecological validity generates an ambiguity between the notions of external and ecological validity making it difficult to understand the real purpose of adding ecological value to consumer and sensory studies. Nevertheless, we can assume that a greater ecological validity leads to a greater external validity of the results.

Some of the main features of laboratory experiments is the required control over the studied stimuli and the control of the environment in which the experimental study takes place. Laboratory or central location experiments may also allow better control of participants' characteristics (e.g. weigh, hunger state, fasting period…). These factors ensure the ability to explain causal relationships between the stimuli and response. Natural experiments may lack of control over those factors, however they ensure greater ecological validity as natural relationships between the participants and the stimuli occur without restrictions or control of the environment. As an intermediate approach, field experiments attempt to reinforce both internal validity, obtained through strict control over the experimental task, and external validity through the use of a natural physical context, following the rationale that if causality is determined by internal validity, the probability that this relationship (stimuli-response) will be relevant in another ecologically valid setting may increase (Roe & Just, 2009).

These concepts, defined below (Table 1) highlight the importance of three features of an experiment when considering whether it is ecologically valid: the nature of the environment, the nature of the stimuli (in this paper we will refer to the nature of the product) and, the nature of the task. Following the works of experimental economists, we propose to consider an additional criterion: the participants – and more precisely, the nature of the pool of participants and the experience they can bring to the task (Carpenter et al., 2004).

Table 1 Definitions and quotes

| The different types of experiments | Notions commonly used to evaluate experimental data |
|---|---|
| **Laboratory or controlled experiments:** "allows underlying causal relations to become manifest at the level of empirical regularities. In | **Validity:** "the best available approximation to the truth of a given proposition, inference, or |

| | |
|---|---|
| a competently performed experiment, single causal connections can be "read off" directly from statistical associations." (Guala, 2012, p.613)<br>**Field experiments:** "define what might be better called an ideal experiment, in the sense that one is able to observe a subject in a controlled setting but where the subject does not perceive any of the controls as being unnatural and there is no deception being practiced." (Harrisson & List, 2004, p.1010)<br>**Natural experiments:** "researcher cannot manipulate the stimulus or influence the data generation process. Rather, the researcher takes advantage of a change in context or setting that occurs for some subjects due to natural causes or social changes beyond the researcher's and subjects' influence" (Roe & Just, 2009, p.1267) | conclusion." (Trochim, 2006)<br>**Robustness:** "measure of the method's capability to remain unaffected by small, but deliberate variations in method parameters (environment, protocol, laboratory, equipment, staff, ...)." (Boutrolle, Arranz, Rogeaux, & Delarue, 2005, p.707)<br>**Reliability:** "the degree to which the result of a measurement, calculation, or specification can be depended on to be accurate." (Oxford Online Dictionary, « Reliability », https://en.oxforddictionaries.com/ viewed online July 2nd, 2018)<br>**Replicability:** "the ability of a scientific experiment or trial to be repeated to obtain a consistent result." (Oxford Online Dictionary, « Replicability », https://en.oxforddictionaries.com/ viewed online July 2nd, 2018) |

## 2.2.Critical points in sensory and consumer studies

In sensory and consumer sciences, laboratories and central location test (CLT) have long been considered the "gold" standard for the study of consumers' liking and behavior. Those scenarios have offered great reliability and robustness of results due to the control of experimental variables through the application of standards (e.g. the AFNOR V09-500 in France) which establishes a methodological framework to explain causal relations. However, in the last decades, the high rate of market failures of new food products that had been selected on the sole basis of CLT, has prompted researchers and industrials to question the ability of these methodological approaches to provide reliable data (Garber, Hyatt, & Starr, 2003; Jaeger et al., 2017b; Köster & Mojet, 2012).

Using the perspective of the four criteria listed above, we try to identify the main critical points that should be considered when assessing the validity of experimental data in sensory and consumer science.

### 2.2.1.Experimental environment

Context was defined by Meiselman, (2006) as the specific physical, social and situational conditions in which food and beverages are consumed. Several studies have shown that the context in which food is

182 evaluated impacts consumers' liking scores and food choices (Edwards, Meiselman, Edwards, &
183 Lesher, 2003; King, Weber, Meiselman, & Lv, 2004; Meiselman, Johnson, Reeve, & Crouch, 2000;
184 Stroebele & De Castro, 2004). These effects can be seen as a result of the role of context as a whole,
185 or more specifically as a result of the presence or absence of some specific contextual variables in a
186 given setting. These aspects will be more specifically addressed in the next section of the article
187 (section 3)
188
189 2.2.2.Nature of the product
190 In this review, focus is placed on studies related to food products, although most considerations would
191 also apply to other product categories. In laboratory settings, food products are usually evaluated as
192 single items (bite or dish) and not as part of a meal; even the portion's size is usually smaller than in
193 more natural settings. However, several studies have shown that products evaluated as part of a meal
194 are higher appreciated than individual items (King, Meiselman, Hottenstein, Work, & Cronk, 2007;
195 King et al., 2004). Rozin & Tuorila, (1993) have described the concept of "eating reference unit" as
196 the size of the tested food (bite, dish, meal, diet pattern) over time. Each reference unit has a different
197 level of complexity, temporal and spatial importance, and research application. For example, a bite is a
198 unit of reference eaten in a short period of time, in a single space and it is used by sensory and product
199 developers; however, a meal is a unit of reference more complex that includes smaller reference unit
200 as bites and that would be used by food service and institutional researchers (Meiselman, 2006).
201 However, in studies taking meals into account rather than isolated products, the definition of "meal" is
202 not standardized, as it depends on the researchers' culture and orientations (Meiselman, 2006; Pliner,
203 Bell, Road, Bell, & Meiselman, 2004).
204 Another critical aspect regarding the ecological validity of the product is its method of preparation.
205 Sensory tests usually employ optimized, standardized cooking methods and minimize variations
206 between batches of products. However, the method of food preparation is involved in the formulation
207 of the hedonic judgement, therefore questioning the ecological validity of the standardized approach
208 (Delarue & Boutrolle, 2010). Several studies have reported a direct effect of preparation methods on
209 liking and discrimination when consumers have the freedom to taste products according to their own
210 habits as they do in natural conditions (Matuszewska, Baryłko-Pikielna, Szczecinska, &
211 Radzanowska, 1997; Posri, Macfie, & Henson, 2001). Variations in preparation methods occur in real
212 life situations, where optimized conditions are rarely met. Yet the standardized tests rarely account for
213 the possible impacts of these variations in the data obtained.
214

215 2.2.3.Selection of participants
216 The mindset of participants when performing a study is a key element in the pursuit of ecological
217 validity. Initial beliefs, attitudes, intentions, knowledge and exposure can all have a significant impact
218 on perceptions and decisions, yet they are rarely taken into account in the interpretation of sensory
219 tests (Bernard & Liu, 2017; Boutrolle, Delarue, Köster, Aranz, & Danzart, 2009; Cardello, Bell, &
220 Kramer, 1996; Edwards & Hartwell, 2009; Kempen et al., 2017; Mahon et al., 2006; Tuorila et al.,
221 2015).
222 In addition to this, the way participants are involved in the test seems to impact consumers' evaluation.
223 Recent studies have pointed out the motivation and involvement of participants as a critical factor
224 when analyzing and comparing different type of experiments (Bangcuyo et al., 2015; Hathaway &
225 Simons, 2017).
226 The way the participants are selected and recruited may also constitute an issue. One of the main
227 criticisms made to inferences drawn in sensory and consumer studies has been the use of non-
228 representative populations. This concern is primarily directed to studies conducted for academic
229 purposes, which frequently use student populations. However, this factor only needs to be considered
230 if the mechanisms or tasks involved in a particular behavior depend on the population type. Depending
231 on the research question, specific populations may be required and in this case, the recruitment of the
232 wrong population may compromise the generalization of the results to a more diverse population
233 (Harrisson & List, 2004).
234
235 2.2.4.Evaluation task
236 The features of the experimental task (experimental procedure or instrumental measure) may also have
237 a significant impact on the respondents' behavior – and therefore on the validity of data. The
238 importance of the nature of the evaluation task performed, as well as the psychological processes
239 involved in the task, have been the focus of several studies in the fields of experimental economics and
240 experimental psychology (Harrison & List, 2004). In sensory and consumers' studies, participants
241 generally answer a questionnaire after tasting a product. The framing of a task, the number and the
242 way of asking the questions have been found to have an impact on consumers' responses (Cardello,
243 2017; Kwak, Ahn, Lee, Kreger, & Lee, 2013; Kwak & Lee, 2016; Lim, 2011; Prescott, Lee, & Kim,
244 2011). Furthermore, some factors such as attention or time perception are known to play a significant
245 role in judgement and decision-making and may directly affect the outcome of a hedonic test or a
246 choice experiment (Dijksterhuis, Smith, van Baaren, & Wigboldus, 2005; Köster, 2003).
247 Another critical point related to the task is the incentive to reply. The presence of incentives directly
248 associated to an experimental task has been shown to have an impact on the way participants report
249 their willingness to pay for a product. In the absence of an incentive (and even in the presence of a

remuneration for their participation), responses tend to exhibit a hypothetical bias, which often manifests in the form of an over-evaluation of the product compared with a consequential task (Carson & Groves, 2007; Shogren, 2005). However, despite its possible implications for new product development, this question has not, to our knowledge, been investigated in the field of hedonic evaluation yet.

## 3. **Increasing ecological validity: what do context studies say?**

### 3.1.From laboratory to natural settings

As a way of addressing the concerns identified in the previous section regarding the validity of such data, it has been suggested to move from controlled settings towards more natural environments – that is to say, to use more contextualized approaches.

Indeed, as an alternative to the laboratory, consumers can be studied in non-standardized, natural consumption environments. The advantage of this field approach is that it reinforces the ecological validity of the experimental setting (environment), allowing researchers to study the interactions between the multiple contextual variables and the consumer's behavior. Regarding the product, while a food product in a laboratory is tested alone and punctually (such as a food product tested as a single dish and presented in a small quantity), the same stimulus in a natural environment (such as a restaurant) may occur in a different, more ecological manner (such as a food product consumed within a meal, in a large quantity). Regarding the task, participants can be unaware of the existence and of the purpose of the study (pure observation of choices or food intake) or be made aware only of some aspects, at the end of the consumption (questionnaires that can be delivered once participants have finished eating or have selected their food) (Lin & Mattila, 2010).

While adding contextual elements may reinforce ecological validity by nature, we are still not sure about the transferability of the data obtained in natural environments in other contexts - not only because of the environment, but also because the stimulus or product itself and, the features of the task performed are different. In the following subsections, we examine more closely the question of ecological validity of context studies.

### 3.2.Do context parameters play a role in the validity of data?

The way to see ecological validity and its potential effects on consumer judgment has direct methodological implications. In the field of sensory and consumer science, studies looking at the validity of contextualized experiments fall into two categories: those that approach the issue of ecological validity as a whole (the experimental context consist of a combination of the environment and the task performed and, attempts to keep most of them as close to natural as possible) and those

that focus on specific factors that are found to have an impact on the measures and, try to make these more ecologically valid.

The studies following a global approach compare scores on food liking and choices in different natural environments (restaurants, canteens, prisons) with those obtained on laboratory or central location settings showing differences on hedonic scores (Edwards, Meiselman, Edwards, & Lesher, 2003; King, Weber, Meiselman, & Lv, 2004; Meiselman, Johnson, Reeve, & Crouch, 2000). Those differences are usually related to the degree of discrimination among products – consumers being more discriminant in natural settings than in laboratory settings – or to the higher scores on natural settings versus laboratory settings. The studies focusing on context variables compare how the addition of contextual variables in controlled experiments affect food liking and choice (King et al., 2004; Stroebele & De Castro, 2004; Weber, King, & Meiselman, 2004). We may first notice that several classifications of contextual variables have been proposed: Rozin & Tuorila (1993) divide contextual variables into either product and non-product variables and subdivide them in simultaneous and temporal contextual factors; Meiselman (1996), proposes to distinguish between three categories of variables (the situation, the individual and the product); whereas Stroebele & De Castro (2004), divide the contextual variables into social context variables, physical surroundings, time related characteristics and distraction and/or television viewing. From these studies, it is difficult to fully disentangle the various factors and isolate a specific context effect. The relevance of those contextual variables thus remains unclear. To date, the lack of knowledge of the combined effects of these contextual variables on consumers' responses compromises the ability to identify causal relationships through experimental approaches. In practice, a consequence of this is that participants to a test may not perceive the study context the way the researcher assumes they would. This questions the ecological validity as defined by Brofenbrenner.

The issue seen as a whole would naturally lead to global changes in the test design, while dividing context into separate variables would bring targeted improvements of the experimental setup, keeping the rest of the task and environment potentially non ecological.

### 3.3.Key determinants of ecological validity: a literature review

#### 3.3.1. Methodology

For this literature review, a search on Google Scholar and Science Direct was conducted using the following keywords: 'context'; 'consumption context'; 'social facilitation'; 'food liking'; 'food choice'; 'food intake'. These keywords were used in combination to identify studies on the effect of the contextual factors (context, consumption context, social facilitation) on consumers' evaluation and behaviors (food liking, food choice, food intake). The reference lists and citations of eligible publications were also reviewed to identify pertinent literature.

A criterion for inclusion in the review was that the study had an experimental design in which either food liking, choice or intake was manipulated by a contextual variable (physical, social or food related). Table 2 shows a complete list of all the studies related to context effects following a a) global, b) separated variable and/or c) global and separated variable approach. We analyzed how those studies try to answer to the question of ecological validity by considering the four factors (participant, stimuli as food product, environment and task) previously presented. Twenty articles were identified that met these selection criteria. Of these, the majority (13) measured food acceptability as the dependent variable of interest, whereas nine articles investigated consumers' choice and intake as regards of meal duration and social facilitation.

On the other hand, in the interpretation of the table we also discuss studies that did not meet our inclusion criteria, but which provided additional insight as regards the use of context and ecological validity.

333 Table 2. Summary of 20 context studies (using a) global approach, b) separated variable approach and c) global and separated variable approach).

**a) Global approach**

| Study | Studied response | Studied factor | Selection of participants | Nature of the product | Experimental environment | Evaluation task | Results | Comments |
|---|---|---|---|---|---|---|---|---|
| de Castro, (1994) | Food intake | **Social facilitation** | 515 participants<br>Participants were pre-recruited and remunerated<br>Between-group design | Regular meals | Natural consumption contexts | Food diary for 7 consecutive days: what was eaten or drunk, time, amount, preparation method, type and gender of people eating with. Hunger, degree of elation and anxiety ratings | **Meals eaten with other people were larger and longer** compared to meals eaten alone. Meals eaten with spouse and family were larger and eaten faster, while meals eaten with friends were larger and of longer duration | Ecological validity and external validity may have been ensured because no changes were done in the contexts and the regular task (eat) was not affected<br>No food type comparison |
| Meiselman et al., (2000) | Food acceptance | **Eating location** | Cross-cultural study:<br>74 and 125 participants (UK data)<br>Participants were recruited for all locations except sensory laboratory<br>Between-group design | • Menu based on canned food<br>• Menu main dish Chicken fettuccine Alfredo | • Training restaurant vs Student cafeteria (UK);<br>• Training rest. vs Food lab vs cafeteria (USA) | Food acceptance on a 9-point hedonic scale (UK)<br>Food attributes (flavor, texture, color, overall rating) on a 7-point hedonic scale (USA) | **Hedonic scores were 1 point higher in the restaurant** > cafeteria. Regarding hedonic attributes (texture, flavor, color), ratings were higher in the restaurant > cafeteria ones | The tested canned food may be unfamiliar to the UK tested population.<br>The nature of the task (questionnaire distribution and number of questions) differed between and within contexts.<br>Participants in the lab condition were in a very specific context and mood state (students in attendance to take a final exam) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Bell & Pliner, (2003) | Meal duration | **Number of people at the table** | 1124 regular clients<br>Participants were unaware of the study and not remunerated<br>Between-group design | Regular meals | • Worksite cafeteria<br>• Moderately priced restaurant<br>• Fast-food restaurant | Purely observational<br>No questionnaire | **Significant effect of group size on meal duration** in the three settings, however the effect was smaller in the fast-food setting | Ecological validity and external validity is ensured because no changes were done in the contexts |
| Edwards et al., (2003) | Food acceptability | **Eating location** | Participants:<br>• Army training camp: 44<br>• University staff refectory/ 38<br>• Private boarding school: 88<br>• Freshman's buffet: 83<br>• Private party: 78<br>• Residential home (elderly): 43<br>• Student refectory: 33<br>• Day care center (elderly): 33<br>• University 4-star restaurant: 19<br>• Hotel 4-star restaurant: 32<br>Participants were unaware of the current study and not remunerated<br>Between-group design | Chicken à la King and Rice | Ten locations, representing different types of food service situations | Demographic questionnaire + appearance, texture, taste and overall acceptability ratings on a 9-point scale + satiety ratings on a 6-point scale | **Contexts affected acceptability ratings**: different scores were obtained as regards product sensory attributes (appearance, taste and texture as well as satiety) | Food preparation is context dependent; therefore, acceptability may differ from one context to another due to sensory properties modification.<br>Contexts also differed on service style, dinning, choice, etc.<br>=> Ecological validity is ensured however results may not be comparable across contexts |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Kozlowska et al., (2003) | Predictive value of hedonic test | **Eating location** | 35 elderly people<br>33 young people<br>Participants were recruited<br>Between and within-group design | 5 apple juices with different sugar concentrations | • Central Location Test (CLT)<br>• Home Use Test (HUT) | Overall liking on a 9-point hedonic scales | **Context do not have a significant effect on hedonic scores** even if higher scores were obtained at home than in laboratory.<br>**Poor prediction of laboratory scores of juice consumption** | The nature of the product differs among contexts (50ml v 150ml) as regards the eating reference unit.<br>Scope for choice differs across contexts, although it may be key to ecological validity<br>The evaluation task differs among contexts: participants answer questionnaires at different times (HUT at the end of the day as a recall) what could affect the attention participants put on the product and therefore, on the final hedonic score |
| Boutrolle, Delarue, Arranz, Rogeaux, & Köster (2007) | Hedonic scores | **Eating location**<br>**Product type** | Participants: regular users of the tested products<br>• Study 1: 240/context<br>• Study 2: 240/context<br>• Study 3: 160/context<br>Participants were recruited in-situ (CLT) or pre-recruited by phone (HUT)<br>Between-group design | 2 variants of each product:<br>• Milk beverage<br>• Salted crackers<br>• Sparkling water<br>Products were sequentially presented | • CLT<br>• HUT | Overall liking scores on a 10-point hedonic scale.<br>CLT: 2 products evaluated during 1 session<br>HUT: After one week of testing participants got the second product and repeated the same task | Products got **higher scores at the HUT**.<br>The influence of the method used depends on the type of product (how products are usually eaten)<br>Pure monadic were slightly higher than monadic sequential scores | The evaluation task differed between contexts (time for evaluation);<br>Ecological validity of HUT may depend on both the environment and the task itself (natural product consumption) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Morizet, Depezay, Combris, Picard, & Giboreau, (2012) | Food choice | **Labelling** | Participants:<br>• Non label: 125<br>• Basic label: 116<br>• Moderated Label: 131<br>Participants were recruited<br>Between-group design | Carrot and broccoli dishes | Three school canteens:<br>• School 1: 140<br>• School 2: 113<br>• School 3: 111 | Chef give or not information about the vegetable options<br>Food choice at lunch time | Children chose significantly more often the familiar version of the dish when no information was given<br>The addition of a descriptive **label led to an increased frequency of choice for the new vegetable dish for carrots** only, and not for broccoli | Ecological validity is ensured as no contextual variable is highly modified. Only the information is manipulated |

**b) Separated variable approach**

| Study | Studied response | Studied factor | Selection of participants | Nature of the product | Experimental environment | Evaluation task | Results | Comments |
|---|---|---|---|---|---|---|---|---|
| Hersleth et al., (2003) | Liking | **Eating location**<br>**Food accompaniment** | 55 participants: likers of wine<br>Participants were pre-recruited<br>Within-group design | Eight different wines + dummy wine<br>Food accompaniment : crackers with cheese, carrots and broccoli with dip and tortilla chips with mild salsa | 4 contexts:<br>• Laboratory: with/without food<br>• Reception room: with /without food | Participants taste the 4 wines at each time and were asked to rate their liking on a 9-point hedonic scale | Sensory differences among wines and contexts significantly influenced liking scores (same size effect)<br>**Food accompaniment had a positive effect** on liking scores | Sensory stimuli differ when the wine is tested together with foods. Using a reception room allowed social interaction and food accompaniment in a natural way |
| King et al., (2004) | Food acceptability | **Social interaction**<br>**Eating location** | Participants: regular consumer of the tested products | Side salad with dressing<br>Small pizza<br>Iced tea | 6 contexts:<br>• T1: Laboratory + individual | Overall liking for the entire meal + overall liking for each meal | **Meal situation** had a strongest **positive effect on tea and salad**; **social** | The number of participants varied a lot from on context to another (from 35 to 106) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | **Choice** | • Test 1: 104<br>• Test 2: 93<br>• Test 3: 106<br>• Test 4: 106<br>• Test 5: 101<br>• Test 6: 35<br><br>Participants were recruited (CLT) and/or contacted by phone or advertisements at the local offices (Test 3-5)<br><br>Between-within group design | Each component has two flavor variants | items<br>• T2: Laboratory + meal<br>• T3: Laboratory + meal + social interaction<br>• T4: Mock restaurant + meal + social interaction<br>• T5: Mock restaurant + meal + social interaction + choice<br>• T6: Restaurant | component on a 9-point structured hedonic scale<br>Demographic information | **interaction negatively affected pizza** hedonic scores; **eating location had a weak** but positive **effect on pizza and tea and a negative effect on salad; and choice** had a **positive effect on salad** | The nature of the product differs from one context to another (meal versus individual meal components) as well as the preparation method because of equipment differences (CLT vs Restaurant). This limits the comparison between hedonic scores obtained in different contexts.<br>The evaluation task implied by one context (questionnaire distribution, choice option, etc.) contributes positively to ecological validity in real and recreated restaurant, but is not applied to the standard CLT |
| Weber et al., (2004) | Food consumption | **Social interaction**<br>**Eating location**<br>**Cutlery**<br>**Choice** | Participants: regular consumer of the tested products<br>• Test 1: 93<br>• Test 2: 106<br>• Test 3: 106<br>• Test 4: 101<br><br>Participants were pre-recruited<br><br>Between-within group design | Pizza + salad + tea: 2 variants of each product | 4 contexts:<br>• T1: Laboratory + meal<br>• T2: Mock restaurant + meal + social interaction<br>• T3: Mock restaurant + meal + social interaction + silver cutlery | Portion size estimation on a portion size scale | **Salad consumption was higher when there was a choice of dressings** in an enhanced environment. **Pizza and tea consumption were higher in an enhanced restaurant-like environment**. **Social interaction alone has no impact on food consumption** | Participants have time to eat and then answer the questionnaire in all context. Social facilitation and choice may increase ecological validity |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | • T4: Mock restaurant + meal + social interaction + silver cutlery + choice | | | |
| Hersleth, Ueland, Allain, & Næs, (2005) | Food acceptance | **Eating location Social facilitation Meal accompaniments** | 87 participants: regular cheese consumers<br>Participants were pre-recruited<br>Within-group design | Hard and semi-hard Norwegian cheeses | 3 contexts:<br>• Laboratory<br>• Club house of a soccer team<br>• HUT | Overall liking on a 9-point hedonic scale + Questionnaire in HUT about social interaction, food eaten together with and beverages | **No differences among contexts** and **social facilitation** | Product experience differs between contexts due to the possibility to eat the cheese together with other products at the HUT (bread, crisp bread, biscuits) so the evaluations may not be comparable.<br>The evaluation task differed between contexts (time for evaluation);<br>Ecological validity of HUT may depend on both the environment and the task itself (natural product consumption) |
| Hetherington et al., (2006) | Food intake | **Social facilitation Relation among participants** | 37 participants<br>Participants were pre-recruited and remunerated<br>Within-group design | Different type of products: bread rolls, potato crisps, fresh green salad, etc. | 3 contexts:<br>• T1: Laboratory<br>• T2: Laboratory + TV<br>• T3: Laboratory + negative social facilitation | Food dairy before the test + Appetite and mood ratings on VAS + eat product + recall how much they had eaten using photographs of 6 possible portion size | **Energy intake was significantly enhanced by presence of familiar others and watching TV** | The experimental environment at T2 and T4 may have influence the evaluation task increasing the ecological validity of the results as regards consumers' intake. Consumers may have experienced a natural consumption situation |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | • T4: Laboratory + positive social facilitation | | | |
| Stroebele & de Castro, (2006) | Food intake and meal duration | **Music** | 78 participants<br>Participants were pre-recruited<br>Within-group design | Food and drink intake of 7 consecutive days | Natural consumption contexts | Food diary: amount and type of eaten food, where, when with whom, for how long, presence of music, music speed and volume on a 7-point scale | **Music increased food and drink intake and longer meal duration**. No significant differences were found in music speed or volume | Social facilitation and meal occasions are confounding variables which can limit the comparison between contexts |
| King, Meiselman, Hottenstein, Work, & Cronk, (2007) | Food acceptability | **Meal situation Social interaction Eating location Choice** | Participants: regular consumers of the tested product<br>• Test 1: 74<br>• Test 2: 83<br>• Test 3: 386<br>Participants were pre-recruited on test 1 and 2<br>Between-group design | • Iced tea<br>• tossed salad with Italian dressing<br>• garlic bread sticks<br>• cannelloni with meat filling<br>• meat lasagna | 3 contexts:<br>• T1: CLT<br>• T2: national Italian chain restaurant<br>• T3: nation-wide in-store satisfaction survey in the same chain restaurant | Overall liking on a 9-point structured hedonic scale (just on CLT) + overall rating on a 6-point structured hedonic scale | **Location and food choice** had the strongest **positive effects on acceptance** ratings, while **social facilitation and enhanced environment** had **no significant effec**t on the acceptability scores | The number of participants highly differs among contexts.<br>The nature of the product differs from one context to another (portion size and preparation). The evaluation task differs from T1 and T2 compares to T3 (whereas in T1 and T2 questionnaires were presented at the beginning of the meal in T3 each product was accompanied by the specific questionnaires). In this case higher number of questions were presented on T1 and T2 which may could affect the evaluation task |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Zeinstra et al., (2010) | Liking | **Preparation method** | 116 participants:<br>• 46: 4-6ys<br>• 25: 7-8ys<br>• 23: 11-12ys<br>• 22: 18-25ys<br>Participants were pre-recruited<br>Between-group design | Carrots and French beans:<br>• mashed<br>• steamed<br>• boiled<br>• stir-fried<br>• grilled<br>• deep-fried | Restaurant | Familiarity questionnaire + ranking test + testing and rating with 3 smiley faces (like, neutral, dislike) + attributes rating + preference ranking | **Vegetable liking was related to a uniform surface and the typical vegetable taste**. Brown coloring and a granular texture were negatively related | Small number of participants.<br>In spite of effort to conduct the test in an experimental restaurant, children were seating alone which may not have been representative of a regular lunch at school |
| Piqueras-Fiszman, Alcaide, Roura, & Spence, (2012) | Food perception | **Product presentation: shape and color of the plate** | Participants:<br>• Study 1: 53<br>• Study 2: 51<br>Participants were recruited<br>Within-group design | Strawberry mouse | Laboratory:<br>• Study 1: white vs black dishes<br>• Study 2: triangular vs squared vs rounded | Taste one spoonful of the sample and rate perceived sweetness, flavor intensity, and quality of the strawberry mousse on an unstructured 10-cm-long scales + liking on a 9-point hedonic scale | Mousse was perceived more **intense and sweeter in a white plate than in a black plate. The shape did not affect** | The nature of the product/task differs between a dessert spoonful and a dessert in terms of reference unit |
| Di Monaco, Giacalone, Pepe, Masi, & Cavella, (2014) | Food acceptability | **Social interaction Drink accompaniments** | Participants: regular consumers of the tested product<br>• Test 1: 32<br>• Test 2: 33<br>• Test 3: 30<br>Participants were pre-recruited<br>Between-group design | 5 frozen croissants | 3 contexts:<br>• T1: Laboratory<br>• T2: Laboratory + social facilitation<br>• T3: Laboratory + social facilitation + drink options | Overall liking + attribute liking (appearance, odor, flavor, taste and texture.) on a 9-point hedonic scales<br>Freshness on a 9-point scale | **Social interaction negatively affected** all the **liking scores** when compared to the control group<br>**Social + drink accompaniments seemed to increase hedonic scores** | Ecological validity could be questioned since participants did not know each other.<br>However, evaluating croissants with a drink is closer to usual eating habits |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| García-Segovia et al., (2015) | Food acceptance and intake | **Eating location**<br>**Table setting** | Participants:<br>• Test 1: 94<br>• Test 2: 90<br>Participants were pre-recruited | Ready-to-eat (RTE) herb-roasted chicken | 2 contexts:<br>• T1: Laboratory<br>• T2: More realistic context (Room, experimental home-style dining room, experimental restaurant)<br>3 table settings:<br>• plastic tray<br>• home-style table<br>• gourmet | Hunger status on a 9-point Likert scale before and after eating<br>• Before eating: Impression of the appearance of the table setting on a 9-point hedonic scale + the willingness to eat on a 9-point Likert scale<br>• After eating: overall flavor and overall impression of the served food on two 9-point hedonic scales + portion size impression on a 9-point Likert scale | **Appearance was higher rated** on the gourmet setting as well as **the willingness to eat** in realistic contexts whereas in the laboratory the table setting did not affect the scores<br>**Intake differs** depending on the context (laboratory < realistic context) and table setting (gourmet < home and plastic tray) | Even in contexts designed to be realistic, the task differed from natural situation, and participants were not allowed to talk and they did not have the possibility to select their meal which may affect the ecological validity of the data |

### c) Global and separated variable approach

| Study | Studied response | Studied factor | Selection of participants | Nature of the product | Experimental environment | Evaluation task | Results | Comments |
|---|---|---|---|---|---|---|---|---|
| De Graaf et al., (2005) | Liking ratings | **Eating location**<br>**Choice** | Participants and soldiers:<br>• T1: 199<br>• T2: 36<br>• T3: 36 | • Main dishes (unfamiliar and familiar)<br>• 4 groupings | 3contexts:<br>• Field study (military camp)<br>• Laboratory | Liking ratings on a 9-point hedonic scale<br>Field environment: before breakfast, | **High correlation** between field and laboratory scores **for snacks** but not for main dishes. | Nature of the product differs between contexts (entire meal vs small bites); this compromises the ecological validity of |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Participants were pre-recruited<br>Between-group design | of snacks (sweet commercial, savory commercial, sweet military, savory military),<br>• 3 entire meal menus | • Laboratory + choice | soldiers chose three meals for the that included some of the tested products and evaluated them<br>Participants chose three menus in a short period of time.<br>Laboratory: participants tested small bites of the stimuli<br>Choice simulation: participants chose three products from a range of stimuli + tasted and evaluated in the laboratory | **Correlations improved** when laboratory subjects were offered a **choice of foods** | lab data as small bites cannot be compared to regular meal consumption situations.<br>The nature of the task (questionnaire filling) differs within the field and between the lab experiment. There is no control over the conditions under which foods were tested and evaluated (by soldiers) which may compromise the comparison of the data.<br>Scope for choice may increase ecological validity as the task is closer to natural consumption situations |
| Petit & Siefferman, (2007) | Liking and consumption | **Product preparation Eating location** | Participants:<br>• T1: 96<br>• T2: 96<br>• T3: 52<br>• T4: 55<br>Participants were recruited for the laboratory study<br>Between-group design | Iced coffee: water and milk based | 4 contexts:<br>• T1: Laboratory<br>• T2: Laboratory + contextual elements (curtains, pictures, etc.)<br>• T3: meeting room<br>• T4: cafeteria | • Overall liking and ranging on a 21-point scale + short questionnaire (demographic + consumption habits) | **No significant differences between T1 and T2**<br>**Significant product differences** in T3<br>Differences between laboratory and natural settings | The nature of the product differs in each context as the preparation method differs.<br>The experimental environment in the T2 may has not been perceived as representative of real life because of external elements that were incongruent (ex: curtains, candles, etc…).<br>Differences regarding the |

room temperature in the different studies (22°C vs 28°C) may have also impacted the evaluation of the iced coffee (a product typically consumed in summer)
The evaluation task also differs among contexts: sequential presentation versus simultaneous

334

335 3.3.2. Main results

336 As can be seen from Table 2, some studies show that context effects on food acceptability differ
337 depending on product categories. Social facilitation shows to increase meal duration as well as food
338 intake whereas food choice seems to increase food acceptability.
339 Similarities and differences are found when comparing the results from studies following a global
340 approach with those following a separated variable approach. The studies following a global approach
341 study consumer behavior through observation (there is no control over the contexts, products,
342 participants and task) or task modulation. When only observation is used, ecological validity is
343 ensured as consumers behave in their regular basis. In this type of studies, food choice, meal duration
344 and intake can be analyzed, however food perception or liking cannot. When the task is modulated
345 (questionnaire filling, food diary), social facilitation increases food intake and differences in hedonic
346 scores are observed across studies. However, these results are controversial as in some situations no
347 differences were observed (Kozlowska et al., 2003). These studies ensured ecological validity as the
348 contexts, products and participants are not altered, however the transferability of the results into
349 another context should be questioned.
350 The studies following a separated variable approach modify not just one contextual factor but several
351 factors at a time (for example the nature of the product or the evaluation task) decreasing the internal
352 validity of the results as well as the ecological validity. In this type of studies, the effect of context on
353 product category should be highlighted as differences between snacks and meals ratings are observed,
354 as well as the effect of the use of congruent elements on consumers' liking. This type of studies has
355 also shown controversial results, being significant in some cases and irrelevant in others (Hersleth et
356 al., 2005; Petit & Sieffermann, 2007).
357 The following parts discuss the outcomes of the literature review regarding the four factors from the
358 previously proposed framework to analyze ecological validity.
359

360 3.3.3. Experimental environment

361 Context has shown to have a certain impact on consumers' liking (Boutrolle et al., 2007; De Graaf et
362 al., 2005; Edwards et al., 2003; García-Segovia et al., 2015). The experimental environment is the
363 most studied factor in the literature on context. However, the comparison of completely different
364 contexts or the addition of contextual variables have led to controversial results as we have previously
365 indicated. The ecological validity of the results can be compromised due to the use of different
366 participant pool in the case of the global approach (different age, social status, etc.) or to the use of
367 incongruent elements in the case of the separated variable approach (García-Segovia et al., 2015; Petit
368 & Sieffermann, 2007). As shown in Table 2, participants and contexts are confounding elements (i.e.

we cannot dissociate both variables) because comparative studies are usually conducted according to a between-group design.

Besides, consumers' expectations and beliefs towards specific food consumption contexts play a key role on consumers' judgement (Bernard & Liu, 2017; Köster, 2003). Hence, it must be stressed that comparing laboratory settings to natural consumption contexts may lead to results as different as comparing hedonic scores from two natural contexts (e.g. school canteen and restaurant). Not only the products may differ in both situations, but also consumers' expectations. Unfortunately, participants' expectations are never really taken into account in studies on context even if they could help to explain differences in consumer behavior and hedonic scores.

### 3.3.4. Nature of the product

Concerning the nature of the product, when the served food sample in a laboratory setting is not representative of the regular amount, preparation and presentation of the same food in a natural setting, it may be hazardous to compare studies because the product/meal combination may not be representative of participants' previous experiences and may convey dissonance and related biases (Rozin & Tuorila, 1993). In fact, we can observe how some products like snacks are able to "pass across contexts" without significant differences on the hedonic scores whereas meals do not (De Graaf et al., 2005; Edwards et al., 2003). This aspect linked to the product category is important to ensure the ecological validity of the results in both global and separated variable approach.

As regards the effect of food combination and sequence of food items during a meal, it is interesting to notice that most of the research on human eating behavior has been focused on food items instead of food combinations. In the last decades, researchers have shown that suitable food combinations result in more pleasant recipes and this is translated in higher overall hedonic scores (Di Monaco et al., 2014; Elzerman, Hoek, Boekel, & Luning, 2011; Hersleth et al., 2003; Pagliarini, Gabbiadini, & Ratti, 2005). While others have also studied how much of each meal component contributes to that (Jimenez et al., 2015; Meiselman, 2006). In addition to this, the sequence and appropriateness of mealtimes when evaluating products has produced different results (Boutrolle et al., 2007; Cardello, Schutz, Snow, & Lesher, 2000; King et al., 2004; Meiselman, 2006). Therefore, the study of products as food items instead of part of a meal may contribute to misleading results that cannot be generalized from one context to another.

Another important aspect that has been already mentioned is that consumers and locations are most often confounded variables. They cannot be studied independently as they are intimately related to consumers' expectations and mindset. Even if a food is exactly the same in two different contexts, consumers may not bring to those contexts the same experience, beliefs and/or expectations. As a consequence, even if they like a given food in one context, consumers may prefer another one that fits

404 better another context. Besides, when comparing consumption settings and particularly meals, the
405 preparation method is a key element in the variability of the sensory properties of the product and may
406 be the source of beliefs, that could, in turn, affect hedonic responses (De Graaf et al., 2005; Edwards &
407 Hartwell, 2009).
408
409 3.3.5.Selection of participants
410 In Table 2, we highlighted the following participant-related aspects found in the literature on context:
411 remuneration, group size and relations and, consumers' familiarity. Most of the participants in
412 laboratory settings are recruited on-purpose and compensated whereas participants to natural context
413 studies are not. This aspect can have a strong impact on consumers' implication and therefore, on
414 obtained data. However, remuneration of participants has not been really explored in the literature on
415 context. On the other hand, some of the studies have compared hedonic scores among different
416 contexts were the studied population was too small to generalize their findings (Edwards et al., 2003;
417 Zeinstra et al., 2010). Moreover, the degree of relation between participants have shown to have
418 different impact on consumers' behavior. When participants know each other they behave in their
419 regular basis whereas when it is not the case, negative correlation with the hedonic scores is obtained
420 (Di Monaco et al., 2014). As regards consumers' familiarity toward the tested products, it must be
421 noted that most of the studies have recruited regular consumers of the tested product. This is an
422 important factor when comparing contexts because some studies have shown that products familiarity
423 may reduce contexts' effects whereas unfamiliar products may be more context-dependent (Giacalone
424 et al., 2015; Hersleth et al., 2005; Kim, Jombart, Valentin, & Kim, 2015). However, we should be very
425 cautious with this notion because in the case of main dishes, familiarity may also be related to
426 particular consumption contexts.
427
428 3.3.6. Evaluation task
429 Table 2 reveals that different tasks have been applied across studies: comparison of overall impression
430 of served food by 9-point hedonic scale, comparison of overall liking by visual analogue scale (VAS),
431 comparison of food attributes, comparison of consumers' willingness to pay, etc. (De Graaf et al.,
432 2005; García-Segovia, Harrington, & Seo, 2015; Kozlowska et al., 2003; Meiselman et al., 2000).
433 Besides, we may observe that, even when the task is the same, hedonic scales and questionnaires
434 frequently differ from one experiment to another as well as from one study to another. We also notice
435 important differences in experimental procedures when comparing one context to another. For
436 example, questionnaires are distributed or displayed differently in different contexts (e.g. paper and
437 pencil vs. digital screen). The same goes with the way to ask participants to test the products, etc. It
438 should also be added that tests in laboratory or in central location do not usually account for the fact in

natural situations consumers may have the possibility to choose the food they want to eat. This may have important consequences in consumers' mindset, not to mention the product experience itself. All in all, the lack of standardization of protocols in the reviewed literature may (at least partly) explain the lack of consistent results as regards the effects of context on consumers' evaluation and behavior.

We argue that ecological validity cannot be seen as independent of internal validity but complementary, and that the focus should be shifted from a search for realism to the definition of clear criteria for transferability from one context to another. Moreover, the focus should be placed on how to isolate the causal effect rather than on the realism from one context to another in order to explain differences among contexts. The pursuit of ecological validity may be seen as a good opportunity to implement the methodologies currently used in the laboratory and try to find a satisfying compromise between the laboratory results and natural setting data.

## 4. New methodological approaches: towards increased transferability?

Rolls & Shide (1992) already anticipated the need to bring together the best features of laboratories and natural consumption contexts in order to study the interactions between contextual variables, but in a controlled way. We identify five approaches designed to address the question of ecological validity. The first one, the classical approach, is the use of natural context that we already described in section 3. The four other approaches are more recent: Living Labs, evoked contexts, immersive contexts and virtual reality. Some of these methods have been described in previous reviews, in particular by Jaeger & Porcherot, (2017c).

We will first define each type of approach, and then characterize the different studies according to this typology.

(a) *Living labs* - Even if a no clear definition for Living labs is found in the literature, the authors have decided to use the definition given by Dell'Era & Landoni (2014) (p.139) where Living Lab is defined as "a design research methodology aimed at co-creating innovation through the involvement of aware users in a real-life setting". In Living labs, the researcher can control and record a selected number of contextual variables and the interaction between them, within a natural consumption situation. Living lab experiments can be seen as an attempt to compromise with the limitations and advantages of laboratory and field experiments, as the control of contextual variables increases the internal validity of the study, while the situation is kept as ecological as possible. Examples of Living labs dedicated to food studies are "The Restaurant of the Future" Wageningen, Netherlands (Hinton et al., 2013; Zeinstra et al., 2010), "The Grill Room" in Bournemouth, United Kingdom (Bell, Meiselman, Pierson, & Reeve, 1994; Meiselman et al., 2000) and "The Living Lab" at the Research Centre of the Institute

Paul Bocuse in Ecully, France (Allirot et al., 2014; Iborra-Bernad, Saulais, Petit, & Giboreau, 2018).

(b) *Evoked contexts* - In the evoked contexts approach, the researcher places the consumer in a typical laboratory evaluation task, but uses either text, audio recordings, and/or pictures that evoke what would be a natural consumption situation of the product (Jaeger & Porcherot, 2017c). In this case, consumers have to imagine themselves in a particular situation and evaluate a product or a set of products. This approach is well established in other disciplines such as marketing studies (Bitner, 1990; Daunt & Greer, 2015; Esmark, Noble, & Breazeale, 2017).

(c) *Immersive contexts* - To define immersive contexts, we should first define what immersion means. Immersion is defined by Witmer & Singer (1998) as "a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences". The main difference between immersive and evoked context approaches is that consumers do not have to imagine themselves in a particular consumption situation, but they experience it instead. Three main features describe the characteristics of immersive contexts: lack of awareness of time, loss of awareness of the real world, involvement and a sense of being in the task environment (Jennett et al., 2008). These approaches usually imply a wealth of means (videos displayed on large screens, multisensory stimulation, including temperature, background sounds, odours, etc.). Within this category, we can also include the recreated environments. Recreated environments are a form of immersive approach where the setting reproduces the physical natural environment where the food consumption would be done, and consumers actually experience a similar situation as in a natural context.

(d) *Virtual reality* - Finally, the virtual reality approach is defined by the "use of virtual environments to present digitally recreated real world activities to participants via immersive (head-mounted displays) and non-immersive (2D computer screens) mediums" (Parson, 2015). We argue that including non-immersive mediums such as the 2D computer screen in the virtual reality definition may create certain confusion with the evoked and immersive contexts categories, where such tools can be also used. For this reason, we have considered only virtual reality studies where head-mounted displays are used.

Following these definitions, Table 3 provides an analysis of these four new methodological approaches through the prism of the four criteria of experimental validity that were previously discussed.

509 Table 3 Summary of new methodological approach studies (a) Living Labs, b) Evoked contexts, c) Immersive contexts, d) Virtual reality)

**a) Living Labs**

| Study | Studied response | Studied Factor | Selection of participants | Nature of the product | Experimental environment | Evaluation task | Results | Comments |
|---|---|---|---|---|---|---|---|---|
| Bell et al., (1994) | Food acceptability and selection | **Decoration** | Participants: regular consumers<br>• Test 1: 63<br>• Test 2: 75<br>Participants were neither recruited nor remunerate<br>Between-group design | A full restaurant menu | 2 contexts:<br>• T1: regular decoration and British food names<br>• T2: Italian decoration and Italian food names | Perceived ethnicity rating and hedonic scores of each component of the meal on a 9-point hedonic scale | **Food acceptability was not affected by decoration** but food selection | Consumers on Italian decoration will may perceived the context as a particular day that may nudges their choices.<br>Scope for choice increase ecological validity of both studies |
| Allirot et al., (2014) | Food intake | **Food choice** | 17 participants<br>Participants were recruited and remunerated<br>Within-group design | A breakfast consumed in one eating episode (F1) and another one consumed in 4 eating episodes (F4). Buffet meal versus standardized meal | 2 contexts:<br>• Experimental restaurant<br>• Laboratory | Appetite rating in a VAS* + blood sampling | In F4, **participants consumed less food in grams and less energy from low energy dense foods at the buffet**, but total energy intakes were not different between conditions | Combination of laboratory and field experiments to ensure the ecological validity of the experiment.<br>No comparison with natural context<br>Scope for choice increase ecological validity of both studies |

**b) Evoked contexts**

| Study | Studied response | Studied factor | Selection of participants | Nature of the product | Experimental environment | Evaluation task | Results | Comments |
|---|---|---|---|---|---|---|---|---|
| Hein, Hamid, Jaeger, & Delahunty, (2010) | Hedonic ratings | **Eating location Product** | Participants: regular apple juice consumers<br>• Test 1: 72<br>• Test 2: 70<br>Participants were recruited and remunerated<br>Between-group design | 4 apple juices: different concentrations of citric acid and strawberry flavor | 2 contexts:<br>• T1: laboratory<br>• T2: evoked refreshing drink occasion | • T1: overall liking on a 9-point hedonic scale + 2 questions about task difficulty and accuracy on a 9-point scale + open ended question about purpose of the study<br>• Evoked context: Participants described their own context + same questionnaires as T1 + 2 more questions about projection task | **Differences in hedonic ratings of the samples were observed between the two contexts**.<br>Tendency of **greater discrimination in evoked context**. Task was considered easier in evoked context and provided information more accurate | The experimental environment might be different for each participant making difficult context comparison.<br>The evaluation task may differ from one participant to another: time needed to project themselves, the accuracy of the scenario, etc. |
| Hein, Hamid, Jaeger, & Delahunty, (2012) | Hedonic ratings | **Eating location Product** | Participants: regular consumers of the tested product<br>• Test 1: 64<br>• Test 2: 62<br>• Test 3: 63<br>• Test 4: 68 | Appel and blackcurrant juice: different concentrations of citric acid and strawberry flavor/ sucrose and orange flavor | 4 contexts:<br>• T1: laboratory<br>• T2: evoked refreshing drink occasion<br>• T3: evoked breakfast | Similar task to the previous study.<br>• T1: after evaluate apple juices participants evaluate the blackcurrant ones | **Higher effect of evoked consumption contexts on hedonic response** was observed for the blackcurrant juice compared to the apple juice. Lower hedonic ratings for | The blackcurrant juice may not be representative for the evoked contexts<br>The experimental environment might be different for each participant making difficult context comparison |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Participants were recruited<br>Between-group design | | situation<br>• T4: evoked movie situation | • T2, T3, T4: one more question about context appropriateness | blackcurrant juice on evoked than laboratory context | The evaluation task may differ from one participant to another: time needed to project themselves, the accuracy of the scenario, etc. |
| Giacalone et al., (2015) | Situational appropriateness | **Familiarity degree**<br>**Eating location** | Participants:<br>• Study 1: 76<br>• Study 2: 97<br>• Study 3: 93<br>• Study 4: 145<br>Participants were recruited and remunerated<br>Within-group design | 9 images of commercially available beers: different familiarity degree | Laboratory + verbal or pictorial contexts (sports, home, alone, etc.) | Participants rated all the usages they perceived appropriate for beer consumption through a checklist task | **Context affected differently familiar and unfamiliar products** | The use of pictorial contexts may help consumers to better project themselves in a particular context of consumption |
| Hersleth, Monteleone, Segtnan, & Næs, (2015) | Intrinsic and extrinsic product cues | **Eating location** | 120 participants: regular consumers of the tested product<br>Participants were recruited<br>Within-group design | 6 types of dry-cured ham | 2 contexts:<br>• Evoked traditional meal<br>• Evoked novel meal | Evaluation of intrinsic characteristics on a 9-point hedonic scale.<br>+<br>extrinsic characteristics on a 9-point scale + question about most common eating situation | **Evoked meal contexts affected both the intrinsic and the extrinsic ratings**, with the strongest effect for the extrinsic ratings.<br>Consumers were somewhat more discriminating when evoking a traditional meal than a novel meal | The use of pictorial contexts may help consumers to better project themselves in a particular context of consumption<br>Role of product/recipe familiarity and expectations is highlighted in this study as a key element when studying consumers' evaluation in food contexts |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Lusk, Hamid, Delahunty, & Jaeger, (2015) | Hedonic responses | **Eating location Evaluation task** | Participants: regular consumers of the tested product:<br>• Study 1: 65<br>• Study 2: 48<br>Participants were recruited and remunerated<br>Between-group design | 4 apple juices:<br>• 2 common<br>• 2 premium) | Evoked refreshing drink occasion | • S1: Overall liking on a best-worst scaling + questions about task complexity and response accuracy<br>• S2: 9-point hedonic scale + questions about task complexity and response accuracy | **Higher product discrimination** was obtained with Best–worst scaling. **Best–worst scaling was perceived as more difficult than the 9-pt scale.** No difference between the two methods on the perceived accuracy of the given information | The evaluation task differs between studies: S1 participants taste 3 times same products during which may lead to better product discrimination |
| de Andrade et al., (2016) | Purchase intention | **Eating location Product presentation** | Participants: regular lamb consumers:<br>• Study 1: 157<br>• Study 2: 171<br>Participants were recruited<br>Between-group design | Lamb meat | 2 contexts:<br>• A celebratory lunch with family over the weekend (weekend lunch context)<br>• A dinner at home after a day's work (weekday dinner context) | Purchase intention on a 9-point structured scale (rating based-conjoint analysis) | **Purchase intention scores were significantly affected by the evoked context**. Consumers were more willing to purchase lamb meat when the celebratory weekend lunch context was considered. In both contexts, price was the most important variable | The use of pictures may help consumers to better project themselves in a particular context of purchase occasion however The evaluation task may differ from one participant to another: time needed to project themselves, the accuracy of the scenario, etc. |
| Kim, Lee, & Kim, (2016) | Food acceptability | **Eating location** | 200 participants: regular consumers of the tested product<br>Participants were recruited | 2 types of coffee | 4 contexts:<br>• T1: Laboratory<br>• T2: Laboratory + evocation<br>• T3: | • T1: overall liking on a 9-point hedonic scale.<br>• T2, T3, T4: vividness of evocation on 9- | **Vividness of evocation lasted longer in the simulated café setting, implying physical cues reinforcing cognitive** | The experimental environment as well as the task may have not been representative for the consumers. The evaluation task may differ from one participant to |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Between-group design | | Laboratory + physical elements<br>• T4: Laboratory + physical elements + evocation | point category scale + liking scores on a 9-point hedonic scale + involvement questionnaire | **evocation**<br>No significant effect of evocation and context but evocation*context*product | another: time needed to project themselves, the accuracy of the scenario, etc. There is no a clear causal relation between the adding contextual elements on the simulated coffee context |
| Jaeger et al., (2017a) | Hedonic product discrimination and sensory characterization | **Eating location** | Cross cultural study: 1727 participants<br>Participants were recruited<br>Between-subject design | Food and drinks familiar for the studied population | 2 contexts:<br>• T1: Laboratory<br>• T2: Evoked drinking a particular product<br>• T3: Evoked breakfast on a weekend morning | • Overall liking on a 9-point hedonic scale + CATA questions + questions about task difficulty and engagement | **No consistent trend in the results regarding the influence of evoked contexts**<br>The evoked context did not systematically influence the results<br>Higher product discrimination on controlled conditions | The experimental environment as well as the task may have not been representative for the consumers. The evaluation task may differ from one participant to another: time needed to project themselves, the accuracy of the scenario, etc. |

### c) Immersive contexts

| Study | Studied response | Studied factor | Selection of participants | Nature of the product | Experimental environment | Evaluation task | Results | Comments |
|---|---|---|---|---|---|---|---|---|
| Sester et al., (2013) | Drink choice | **Drinking location** | Participants:<br>• Study 1A: 93<br>• Study 1B: 83<br>• Study 2A: 60<br>• Study 2B: 60<br>Participants were recruited for the study 2 | Drinks | 4 Contexts<br>• S1,2A: immersive bar with warmth furniture<br>• S1,2B: immersive bar with cold | • S1A, B: look at the clips + order a drink from a list + bar warmth rating on a 10-point scale<br>• Study 2A, B: triangle test + questionnaire | **S1A, B: immersive contexts influenced drink choice** (association between drinks and video clips).<br>S2A, B: drink choices were different according to the | The nature of the task in the S2 is not representative of the natural environment (triangle test).<br>Results should be replicate in another immersive or natural environments |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Within- and between-group design | | furniture<br>• S1A, B: + 5 different clips | (warmth scale and bar appropriateness) | ambience | Task in S1 is different to the S2 so the robustness of the data should be questioned |
| Bangcuyo et al., (2015) | Liking and preference | **Drinking location** | 46 participants: regular coffee consumers<br>Participants were recruited and remunerated<br>Within-subject design | 5 coffees | 2 contexts:<br>• Laboratory<br>• Immersive café | • Laboratory: demographic questionnaire + acceptability on a 9-point hedonic scale + ranking + Engagement Questionnaire<br>• Café: same evaluation task after a month | **Significant differences in preference order and liking were found between contexts Participants were more discriminants in the immersive coffee** and results more reliable predictor of future coffee liking (replication)<br>Immersive coffee increased engagement | The nature of the immersive context may be representative for a coffee drinking situation<br>Contrary to the evaluation task that may have not be representative of a natural situation (5 testing coffees). However, results from replication seems to confirm products liking |
| Hathaway & Simons, (2017) | Data quality and panelist engagement | **Eating location** | 59 participants: regular consumers of the tested products<br>Participants were recruited and remunerated<br>Within-subject design | 4 commercially available chocolate chip cookies | 4 contexts:<br>• Laboratory<br>• Mixed immersive domestic kitchen<br>• Full immersion domestic kitchen | Demographic questionnaire + acceptability on a 9-point hedonic scale + Engagement questionnaire | **Participants were more discriminants in the full immersion context**.<br>Similar levels of engagement in the two immersive conditions<br>Different results as regards the first and second replication with no liking differences among contexts | The nature of the product, a familiar product that may be eaten in different type of contexts may have caused this effect on liking; certain product categories are more context-dependent than others |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Holthuysen et al., (2017) | Overall-liking and just-about-ratings | **Eating location Product** | Participants:<br>• Study 1, 2: 242<br>• Study 3: 222<br>Participants were recruited for the controlled and recreated condition<br>Between-group design | 2 airplane meals: 2 variants of each | 3 contexts:<br>• S1: Laboratory<br>• S2: Recreated airplane<br>• S3: Real airplane | • S1: selection of the product + overall liking on a Visual Analogue Scale.<br>• S2, S3: evaluation of the previous tested sample but as a full meal + overall liking on a VAS+ sensory attribute rating on a JAR scale | **Recreated and actual airplane were more discriminant than the controlled condition**. No significant difference on ratings between recreated and natural context were found | First study that compares new approaches to natural conditions |
| Liu, Hooker, Parasidis, & Simons, (2017) | Food quality, nutritional content, liking | **Labelling** | 120 participants: regular peanut butter consumers divided in 3 labelling conditions<br>Participants were recruited and remunerated<br>Between-group design | Peanut butter | **Immersive context**: virtual grocery store + 3 conditions:<br>• Blind<br>• Labelled<br>• Labelled + verbal call-out | Acceptability on a 9-point hedonic scale + WTP + Food quality questionnaire + demographic information | **Labelling improved product quality and nutritional content perception but not liking and WTP**. Verbal call out improved food quality, nutritional content and WTP | |
| Sinesio et al., (2018) | Food perception and liking | **Eating location** | 48 participants: regular consumers of the tested product<br>Participants were recruited | Salad tomato and wild rocket salad at different storage time | 2 contexts:<br>• Study 1: Laboratory<br>• Study 2: Immersive holiday farm dining room | • S1: overall liking on a 9-point hedonic scale + perceived freshness on a 9-point scale.<br>• S2: same evaluation task + | **Liking scores were higher in the immersive environment setting than in the lab**. However, higher discrimination as regards storage time | The nature of the task at the lab could increase consumers' attention leading to a higher discrimination |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Within-subject design | | | Engagement Questionnaire | was found in the lab | |

**d) Virtual reality**

| Study | Studied response | Studied factor | Selection of participants | Nature of the product | Experimental environment | Evaluation task | Results | Comments |
|---|---|---|---|---|---|---|---|---|
| van Herpen, van den Broek, van Trijp, & Yu, (2016) | Shopping behavior | **Location** | 100 participants: regular consumers of the tested products and buyers of the tested real supermarket<br>Participants were recruited<br>Between-group design | Milk, fruit and vegetables and biscuits | 3 contexts:<br>• Real supermarket<br>• 3D VR supermarket<br>• Laboratory + supermarket pictures | Same task in the 3 contexts:<br>• buy a list of products fill a questionnaire | **Similar results were obtained in the VR condition and the real context** for milk and biscuits. However, participants bought more products and spent more money (for biscuits and fruit & vegetables), in VR and picture condition | The nature of the environment and evaluation task in the picture condition and VR may not be representative of the actual behavior. We should consider the effect of the VR devices on the evaluation task |
| Andersen et al., (2018) | Desires and liking | **Location** | 60 participants<br>Participants were recruited<br>Between and within-group design | Beverages and skin care lotion odor | 3 contexts:<br>• Laboratory<br>• 3D VR beach<br>• Laboratory + beach picture | Thirst, hunger and emptiness/fullness on VAS + familiarity with VR + desires for particular beverages on a 10-point scale + odor skin liking on a 10-point scale + engagement on a 7-point Likert scale + level of excitement + choice of a snack | **Desire for cold vs hot beverages was significantly higher in the beach exposures** and particularly for VR. After exposure, beach and laboratory contexts did not differ in beverages desires. Beach scenarios did not affect odor liking and no retention effect on choice behavior was observed | The nature of the stimuli, VR environment and task may have been representative of natural environments. However, a natural context comparison would be needed to prove the predictability and ecological validity of the applied methodology as electronic devices could impact consumers' evaluation ("wow" effect) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | at the end of the test | Beach scenarios increased participant' engagement especially VR | |
| Ouellet, Boller, Corriveau-Lecavalier, Cloutier, & Belleville, (2018) | Feasibility and construct validity of a new methodology | **Memory** | Participants:<br>• Study 1 :49<br>• Study 2:35<br>Participants were recruited<br>Between and within-group design | List of products | Virtual supermarket | • S1: memorize a list of products + buy those products<br>• Study 2: Multifactorial Memory Questionnaire (MMQ) + same task S1 | **The virtual store showed to have an appropriate level of difficulty, supporting the feasibility and construct validity of the task** according to everyday memory tasks' results | The nature of the task (memorize) may have been representative of natural context. However a natural context comparison would be needed to prove those results |
| Schnack, Wright, & Holdershaw, (2018) | Telepresence and usability | **Location** | Participants:<br>• Study 1: 62<br>• Study 2: 49<br>Participants were recruited<br>Between-group design | Food products | 2 contexts:<br>• S1: 3D VR supermarket<br>• S2: Laboratory + screen with supermarket images | Same task in both contexts: products purchasing + Presence Questionnaire (7-point Likert scale) + usability and open ended questions | **Immersive Virtual Reality improves participants' telepresence and usability**. A significant age group and gaming experience was shown | The nature of the task (purchase) may have been representative of natural context. However, a natural context comparison would be needed to prove those results as electronic devices could impact consumers' evaluation ("wow" effect) |
| Siegrist et al., (2018) | Food selection | **Location Task** | Participants:<br>• Study 1A: 37<br>• Study 1B: 31<br>• Study 2: 50<br>Participants were recruited and remunerated | Cereals | 2 contexts:<br>• S1A: Real life supermarket<br>• S1B: 3D VR supermarket<br>• S2: 3D VR supermarket | • S1A, B: similar task. Select one cereal package for kid's camp + one for a specific type of diet<br>• S2: select a healthy or tasty | S1A, B: **no significant differences between contexts** were found. S2: **significant differences between tasks were found**. Participants spend more time for the healthy package | The use of Eye tracking devices in the real life context may have reduce the ecological validity of the experiment providing similar results between S1A and S1B |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Between and within- group design | | | cereal package depending on the test condition (healthy or tasty) | selection than the tasty one | |
| Ung, Menozzi, Hartmann, & Siegrist, (2018) | Energy content | **Location** | 34 participants<br>Participants were recruited and remunerated<br>Within-subject design | 3 types of foods (Fake food buffet) | 2 contexts:<br>• Laboratory buffet<br>• 3D VR buffet | Same task in both contexts: serve themselves a meal similar to what they would normally have for lunch | **No significant differences as regards the energy content** between settings were found | The nature of the environment in the laboratory may not have been representative of the natural consumption setting even as a fake buffet.<br>As regards the task (serving) this may be representative for the participants. However, a natural context comparison would be needed to prove those results |

Before analyzing each approach, as a general comment, we would like to highlight that the results obtained from each approach may differ depending on the nature of the product (product category) and the familiarity with the product. Certain products may be more affected by situation-specific cues than others. Therefore, special attention should be given to these aspects when analyzing and comparing products evaluations from one context to another.

As it can be seen in the Living Lab studies (Table 3, section a) the characteristics of the participants, the nature of the product and the environment are kept as realistic as possible, whereas the evaluation task through the inclusion of questionnaires may compromise the external validity of the results in a certain way. Consumers experience a natural consumption situation, therefore the transferability of the data to another setting that follows similar patterns can be achieved. However, the use of this type of settings may be costlier and require additional logistics compared to the use of other contextual methodologies.

Concerning the evoked context studies, this approach is easy to apply and inexpensive because not physical elements are added. However, the degree to which participants project themselves to the evoked context is not controlled, despite attempts to measure vividness of evocations, making generalization of results to other contexts difficult (Köster, 2003). Therefore, the gain in ecological validity due to evocation of a consumption situation is difficult to assess, and may very well be outweighed by the loss due to artificiality of the projective task implied by such a procedure.

Immersive approaches have been hypothesized to improve consumers' involvement as well as product discrimination as participants may experience similar psychological processes that in natural contexts (Andersen, Kraus, Ritz, & Bredie, 2018). As it was previously discussed, consumers' experiences and prior beliefs about particular contexts are key elements when conducting sensory evaluations in contexts studies (Köster, 2003). The fact that consumers experience a natural consumption situation even if it is under controlled conditions may ensure the ecological validity of the results and improve the external validity. However, as it can be seen in Table 3 - section c, there is a lack of standardization of the contextual variables in the immersive studies that have been conducted so far, - different degrees of immersion can be shown - therefore there is limited knowledge about the relevance of each contextual variable and their contribution to the outcome of experimental studies. Moreover, the higher costs that these methodologies involve have been highlighted as main drawbacks in their use.

To our knowledge, so far only one published study has attempted to compare immersive and natural settings methodologies. In a study of the impact of context on food evaluation of airplane meals, Holthuysen, Vrijhof, de Wijk, & Kremer, (2017) compared overall liking and just-about-right ratings in laboratory, recreated airplane and an actual plane. Recreated and actual plane settings showed similar results, contrary to laboratory settings. However, in this case it should be highlighted that the

actual immersive context was a recreated environment. A flight was recreated through the use of a physical environment (cabin creation), use of boarding passes and hand luggage, flight instructions, regular time of flight, etc. Unlike most immersive tests, recreated environments do not place participants in a location where screens, sounds or smells are combined. Further work is therefore needed in the definition and categorization of immersive experiments and on the comparison of external validity between this approach and natural settings.

Finally, an increasingly popular methodological approach to improve the ecological validity is the use of virtual reality. Until now, most of these studies have focused on consumers' purchasing behavior in food stores. This methodological approach has offered controversial results as regards product discrimination and consumer behavior (Dreyfuss, Porcherot, Sinesio, Henneberg, Depoortere, & McEwan, 2018). Whereas in some studies similar results have been obtained in virtual and natural environments, in other situations an over effect has been reported. The virtual reality allows participants to place themselves in particular contexts (telepresence) and improve products usability increasing the engagement in the task. However, in some situations, depending on the type of used technology, the use of electronic devices may compromise the "natural" experience and biases the obtained results even if consumers are used to this type of technology. Moreover, the nature of the environment remains non-ecological when 2D computer screens are used as well as the product evaluation task, especially when the research question is related to product acceptability. For further discussion about virtual reality studies, the reader is directed to Stelick & Dando, (2018).

## 5. **Contribution**

### 5.1.Research

Our analysis of context studies in sensory and consumer science considers four critical points when evaluating the need for a given contextual parameter: the experimental environment, the nature of the product, the selection of participants, and the evaluation task. This review adds evidence to the lack of standardized methodologies and analytical framework highlighted by several previous reviews, as well as the problems of robustness and reliability of the results that it induces. We suggest that the use of contextual variables needs to be assessed according to their contribution to ecological, but also internal validity.

There has been a lot of research on the effects of context on consumers' hedonic response, food choice or intake, however the overall inconsistency of findings renders difficult their integration into clear guidelines to improve the ecological validity of a study. In particular, to date, the results are too context-specific, product category specific or task specific to enunciate more general principles that could be used to develop such a framework. This has led to the emergence of new methodological

579 approaches, with limited effort to determine how, and whether, each of these methods may
580 complement or outrank the other.
581 This review also highlights that contexts and consumers are confounded variables that make the
582 generalization of the results even more hazardous, as hedonic responses are deeply related to
583 consumers' expectations, which are themselves related to each context.
584 Therefore, we propose to pursue the ecological validity in sensory and consumer studies from a
585 complementary perspective, in which laboratory and new methodological approaches work together in
586 complementarity. When and how we should consider ecological validity as a goal in research on
587 context should be the most important question. Living labs and immersive studies may be able to
588 reinforce ecological validity when looking at consumers' choice or purchase intention. However, no
589 study has yet examined the external validity of data acquired in such conditions. It is advisable for
590 researchers to plan studies to compare similar methodological approaches (internal comparisons of
591 living labs and immersive studies) across different contexts and different product categories in order to
592 gain better knowledge and understanding of the reliability of the applied methodologies.
593 In line with the theories of behavioral economics, in particular Prospect Theory (Tversky &
594 Kahneman, 1991), we also propose to give more attention to context-induced reference points when
595 evaluating products. Beliefs (prior experiences) associated to a particular context may indeed play a
596 role by predisposing consumers to a different framework of evaluation. Reference points have been
597 shown to greatly modulate judgement and decision making. Even if very few studies have focused on
598 the effects of context and beliefs on food evaluation, the reference framework of evaluation is likely to
599 be an important factor explaining context effects (Bernard & Liu, 2017; Kempen et al., 2017).
600
601 5.2.Practical implications
602 This review has several practical implications. Firstly, we observe that, even though contextual
603 variables have been found to modulate consumer behavior, we cannot establish clear operational
604 recommendations because of the heterogeneity of results found in the literature.
605 However, this review provides a framework and criteria to assess ecological validity, which could
606 contribute to increase methodological thoroughness in the fields of sensory and consumers' studies,
607 providing workable outcomes to the private sector, notably for product development.
608 Among all the attempts to improve context, based on our review of the (limited number of) works
609 using recent methodologies, it seems that consumers are more engaged in the task and able to
610 experience a natural context in living lab and immersive approaches. A possible explanation could be
611 that unlike in evoked settings and virtual reality, participants do not have to put too much effort in
612 imagining a consumption situation or use electronic devices which could make the task more

ecologically valid. However, this type of experiments can be expensive and difficult to logistically handle, and this conclusion needs to be strengthened by more comparative data.

As regards the nature of the food, it is important to consider, especially in the context of new product development, the type of product that the test aims to evaluate, at which stage of development process the data is needed, and in which settings the final product will be consumed. It has been shown that the impact of context depends on the product category and units of evaluation (e.g. product vs dish). Moreover, familiarity towards the tested product seems to modulate the contexts effects: while a product familiar to consumers can be eaten in several contexts, unfamiliar products can be related to particular occasions and consumption contexts. In the early stages of product development, when specific sensory product characteristics should be defined, laboratory settings should be considered as the best solution. However, when it comes to the choice or purchase intention, more naturalistic environments may be needed to ensure product success.

Although survey institutes and stakeholders in the industry are well aware of the necessity to recruit consumer samples that are representative of a target population, other participant-related factors (the way the participants are recruited and the incentives they receive to take part in the studies) are less considered and yet may also be relevant concerning the validity of hedonic results. Besides, some studies have shown that is important to consider participants' prior experiences, expectations and beliefs when testing a food, as those factors can tell us more about the consumer and the way he/she will behave in a specific context. These aspects are particularly important when evaluating full dishes. In particular, when comparing natural contexts (institutional meals, restaurants, etc.), food preparation has been shown to have a direct impact on the sensory properties of a product and to indirectly influence consumers' evaluation due to the associations made between context and served food.

Finally, as regards the evaluation task, we should consider several aspects. First of all, when comparing contexts, we should ensure that the task and the experimental procedure are the same in order to be able to compare the results. Moreover, it is important to keep in mind, consumers will be more focused on the task performance, therefore on the product itself, in laboratory settings than in the natural consumption settings where the hedonic score can include other aspects such as the actual experience, environment, etc. Therefore, further research is needed to improve the understanding of the effect of experimental procedures and instrumental measures used when comparing settings on the participants' evaluation processes.

## 6. Limitations

The lack of homogeneity in the definition and the lack of consistency and standardization in the use of contextual variables and associated tools to measure consumers' behavior may have limited the conclusions that could be drawn from this review.

Another important point is that, although the literature has shown different ways of classifying contextual variables, the relative weight and significance of those variables on consumer behavior need further assessment, especially through replicated studies. Moreover, as it has been shown, several experimental procedures are used through the different studies, thus making it difficult to compare their findings. We suggest that further research should dedicate more attention to the understanding of the nature of the task.

## 7. Conclusion & Perspectives

Increasing the number of consumer studies in natural settings was pointed as one of the most important challenges for research during the 11th Pangborn Sensory Science Symposium (Jaeger et al., 2017b). In the past decades, sensory and consumer scientists have tried to move from laboratory experiments to natural experiments and different alternative approaches, such as evoked or immersive contexts or virtual reality, have emerged with the purpose of ensuring better ecological validity. Ecological validity is achieved if participants perceive the experimental environment, the food they taste and the task they perform to be representative of a natural consumption situation.

On the other hand, as Guala (2012) proposes, internal validity should be firstly addressed to tackle the problem of external validity. By knowing under which circumstances the results can be extrapolated may allow us to find the specific reasons to explain why results may not be generalized. The problem of external validity might be related to the lack of important factors or the presence of artificial conditions in the experimental design that are far from the natural situations. However, is it important to determine the extent to which those factors can be transferred and reproduced in the laboratory, whether this is always possible, and what is the degree of ecological validity and realism that the researcher should assume and seek depending on the purpose and finality of the study.

## 8. Acknowledgements

The authors acknowledge Herb Meiselman for his kind comments and suggestions on this paper first draft. This review is part of a PhD project funded by the Food Hygiene Scientific Society (SSHA).

## 9. References

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2), 179-211. https://doi.org/10.1016/0749-5978(91)90020-T

Allirot, X., Seyssel, K., Saulais, L., Roth, H., Charrié, A., Drai, J., … Laville, M. (2014). Effects of a breakfast spread out over time on the food intake at lunch and the hormonal responses in obese men. *Physiology & Behavior*, *127*, 37–44. https://doi.org/10.1016/j.physbeh.2014.01.004

Andersen, I. N. S. K., Kraus, A. A., Ritz, C., & Bredie, W. L. P. (2018). Desires for beverages and liking of skin care product odors in imaginative and immersive virtual reality beach contexts. *Food Research International*. https://doi.org/10.1016/J.FOODRES.2018.01.027

685 Bangcuyo, R. G., Smith, K. J., Zumach, J. L., Pierce, A. M., Guttman, G. A., & Simons, C. T. (2015).
686 The use of immersive technologies to improve consumer testing: The role of ecological validity,
687 context and engagement in evaluating coffee. *Food Quality and Preference*, *41*, 84–95.
688 https://doi.org/10.1016/j.foodqual.2014.11.017

689 Bell, R., Meiselman, H. L., Pierson, B. J., & Reeve, W. G. (1994). Effects of adding an Italian theme
690 to a restaurant on perceived ethnicity, acceptability, and selection of foods. *Appetite*, *22*.
691 https://doi.org/10.1006/appe.1994.1002

692 Bell, R., & Pliner, P. L. (2003). Time to eat: The relationship between the number of people eating and
693 meal duration in three lunch settings. *Appetite*, *41*(2), 215–218. https://doi.org/10.1016/S0195-
694 6663(03)00109-0

695 Bernard, J. C., & Liu, Y. (2017). Are beliefs stronger than taste? A field experiment on organic and
696 local apples. *Food Quality and Preference*, *61*(October 2016), 55–62.
697 https://doi.org/10.1016/j.foodqual.2017.05.005

698 Bitner, M. J. (1990). Evaluating Service Encounters: The Effects of Physical Surroundings and
699 Employee Responses. Journal of Marketing, 54(2), 69. https://doi.org/10.2307/1251871

700 Boutrolle, I., Arranz, D., Rogeaux, M., & Delarue, J. (2005). Comparing central location test and
701 home use test results: Application of a new criterion. *Food Quality and Preference*, *16*(8), 704–
702 713. https://doi.org/10.1016/j.foodqual.2005.03.015

703 Boutrolle, I., Delarue, J., Arranz, D., Rogeaux, M., & Köster, E. P. (2007). Central location test vs.
704 home use test: Contrasting results depending on product type. *Food Quality and Preference*,
705 *18*(3), 490–499. https://doi.org/10.1016/j.foodqual.2006.06.003

706 Boutrolle, I., Delarue, J., Köster, E., Aranz, D., & Danzart, M. (2009). Use of a test of perceived
707 authenticity to trigger affective responses when testing food. *Food Quality and Preference*,
708 *20*(6), 418–426. https://doi.org/10.1016/j.foodqual.2009.03.004

709 Bressoud, E. (2013). Testing FMCG innovations: experimental real store versus virtual. *Journal of*
710 *Product & Brand Management*, *22*(4), 286–292. https://doi.org/10.1108/JPBM-05-2012-0141

711 Brofenbrenner, U. (1977). Toward an Experimental Ecology of Human Development. *American*
712 *Psychologist*, *32*(7), 513–531. https://doi.org/10.1037/0003-066X.32.7.513

713 Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*,
714 *50*(3), 255–272. https://doi.org/10.1037/h0060889

715 Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology.
716 *Psychological Review*, *62*(3), 193–217. https://doi.org/10.1037/h0047470

717 Card, D., DellaVigna, S., & Malmendier, U. (2011). The Role of Theory in Field Experiments. Journal
718 of Economic Perspectives, 25(3), 39–62. https://doi.org/10.1257/jep.25.3.39

719 Cardello, A. V., Bell, R., & Kramer, M. (1996). Attitudes of consumers toward military and other
720 institutional foods. *Food Quality and Preference*, *7*(1), 7–20.

721 Cardello, A. V. (2017). Hedonic scaling: assumptions, contexts and frames of reference. *Current*
722 *Opinion in Food Science*, *15*, 14–21. https://doi.org/10.1016/j.cofs.2017.05.002

723 Cardello, A. V, Schutz, H., Snow, C., & Lesher, L. (2000). Predictors of food acceptance,
724 consumption and satisfaction in specific eating situations. *Food Quality and Preference*, *11*(3),
725 201–216. https://doi.org/10.1016/S0950-3293(99)00055-5

726 Carpenter, J. P., Harrison, G. W., & List, J. A. (2004). Field Experiments in Economics: an
727 Introduction. *Research in Experimental Economics*, *10*(04), 1–15. https://doi.org/10.1016/S0193-
728 2306(04)10001-X

729 Carson, R. T., & Groves, T. (2007). Incentive and informational properties of preference questions.

730 *Environmental and Resource Economics*, *37*(1), 181–210. https://doi.org/10.1007/s10640-007-
731 9124-5

732 Daunt, K. L., & Greer, D. A. (2015). Unpacking the perceived opportunity to misbehave: The
733 influence of spatio-temporal and social dimensions on consumer misbehavior. *European Journal*
734 *of Marketing*, *49*(9/10), pp. 1505-1526. https://doi.org/10.1108/EJM-01-2014-0061

735 de Andrade, J. C., Nalério, É. S., Giongo, C., de Barcellos, M. D., Ares, G., & Deliza, R. (2016).
736 Influence of evoked contexts on rating-based conjoint analysis: Case study with lamb meat. *Food*
737 *Quality and Preference*, *53*, 168–175. https://doi.org/10.1016/j.foodqual.2016.06.013

738 de Castro, J. M. (1994). Family and friends produce greater social facilitation of food intake than other
739 companions. *Physiology & Behavior*, *56*(3), 445–455. https://doi.org/10.1016/0031-
740 9384(94)90286-0

741 De Graaf, C., Cardello, A. V., Matthew Kramer, F., Lesher, L. L., Meiselman, H. L., & Schutz, H. G.
742 (2005). A comparison between liking ratings obtained under laboratory and field conditions: The
743 role of choice. *Appetite*, *44*(1), 15–22. https://doi.org/10.1016/j.appet.2003.06.002

744 Delarue, & Boutrolle. (2010). The effects of context on liking : implications for hedonic measurements
745 in new product development. *Consumer-Driven Innovation in Food and Personal Car Products*.
746 Retrieved from http://prodinra.inra.fr/?locale=en#!ConsultNotice:257347

747 Dell'Era, C., & Landoni, P. (2014). Living Lab: A Methodology between User- Centred Design and
748 Participatory Design. Creativity and Innovation Management, 23(2), 137–154.
749 https://doi.org/10.1111/caim.12061

750 Di Monaco, R., Giacalone, D., Pepe, O., Masi, P., & Cavella, S. (2014). Effect of social interaction
751 and meal accompaniments on acceptability of sourdough prepared croissants: An exploratory
752 study. *Food Research International*, *66*, 325–331. https://doi.org/10.1016/j.foodres.2014.10.001

753 Diehl, M., Wahl, H.-W., & Freund, A. (2017). Ecological Validity as a Key Feature of External
754 Validity in Research on Human Development. *Research in Human Development*, *14*(3), 177–
755 181. https://doi.org/10.1080/15427609.2017.1340053

756 Dijksterhuis, A., Smith, P. K., van Baaren, R. B., & Wigboldus, D. H. J. (2005). The Unconscious
757 Consumer: Effects of Environment on Consumer Behavior. *Journal of Consumer Psychology*,
758 *15*(3), 193–202. https://doi.org/10.1207/s15327663jcp1503_3

759 Dreyfuss, L., Porcherot, C., Sinesio, F, Henneberg, S., Depoortere, L. & McEwan, J. A. (2018). In-
760 context consumer research: Benefits and opportunities for immersive techniques. Workshop
761 presented at *the 8th Eurosense – European Conference on Sensory and Consumer Research.* 2-5
762 september 2018,Verona, Italy.

763 Edwards, J. A., & Hartwell, H. J. (2009). Institutional meals. In H. L. Meiselman (Ed.), *Meals in*
764 *Science and Practice. Insterdisciplinary research and business applications* (pp. 102–127).
765 Cambridge: Woodhead Publishing Limited.

766 Edwards, J. S. A., Meiselman, H. L., Edwards, A., & Lesher, L. (2003). The influence of eating
767 location on the acceptability of identically prepared foods. *Food Quality and Preference*, *14*(8),
768 647–652. https://doi.org/10.1016/S0950-3293(02)00189-1

769 Elzerman, J. E., Hoek, A. C., Boekel, M. A. J. S. Van, & Luning, P. A. (2011). Consumer acceptance
770 and appropriateness of meat substitutes in a meal context. *Food Quality and Preference*, *22*(3),
771 233–240. https://doi.org/10.1016/j.foodqual.2010.10.006

772 Esmark, C. L., Noble, S. M., & Breazeale, M. J. (2017). I'll Be Watching You: Shoppers' Reactions to
773 Perceptions of Being Watched by Employees. Journal of Retailing, 93(3), 336–349.
774 https://doi.org/10.1016/j.jretai.2017.04.005

Garber, L. L., Hyatt, E. M., & Starr, R. G. (2003). Measuring consumer response to food products. *Food Quality and Preference*, *14*(1), 3–15. https://doi.org/10.1016/S0950-3293(02)00030-7

García-Segovia, P., Harrington, R. J., & Seo, H.-S. (2015). Influences of table setting and eating location on food acceptance and intake. *Food Quality and Preference*, *39*, 1–7. https://doi.org/10.1016/j.foodqual.2014.06.004

Giacalone, D., Frøst, M. B., Bredie, W. L. P., Pineau, B., Hunter, D. C., Paisley, A. G., … Jaeger, S. R. (2015). Situational appropriateness of beer is influenced by product familiarity. *Food Quality and Preference*, *39*, 16–27. https://doi.org/10.1016/J.FOODQUAL.2014.06.012

Fréchette, G. R., & Schotter, A. (2015). Handbook of Experimental Economic Methodology. Oxford: Oxford University Press.

Guala, F. (2012). Experimentation in Economics. *Philosophy of Economics*, (June 2005), 597–640. https://doi.org/10.1016/B978-0-444-51676-3.50021-X

Harrisson, G. W., & List, J. a. (2004). Field Experiments. *Journal of Economic Literature*, *XLII*(December), 1009–1055. https://doi.org/10.1257/0022051043004577

Hathaway, D., & Simons, C. T. (2017). The impact of multiple immersion levels on data quality and panelist engagement for the evaluation of cookies under a preparation-based scenario. *Food Quality and Preference*, *57*, 114–125. https://doi.org/10.1016/J.FOODQUAL.2016.12.009

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*(4), 410–416. https://doi.org/10.1016/j.foodqual.2009.10.003

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*(1), 35–44. https://doi.org/10.1016/j.foodqual.2012.02.014

Hersleth, M., Mevik, B.-H., Næs, T., & Guinard, J.-X. (2003). Effect of contextual factors on liking for wine—use of robust design methodology. *Food Quality and Preference*, *14*(7), 615–622. https://doi.org/10.1016/S0950-3293(02)00190-8

Hersleth, M., Monteleone, E., Segtnan, A., & Næs, T. (2015). Effects of evoked meal contexts on consumers' responses to intrinsic and extrinsic product attributes in dry-cured ham. *Food Quality and Preference*, *40*, 191–198. https://doi.org/10.1016/j.foodqual.2014.10.002

Hersleth, M., Ueland, Ø., Allain, H., & Næs, T. (2005). Consumer acceptance of cheese, influence of different testing conditions. *Food Quality and Preference*, *16*(2), 103–110. https://doi.org/10.1016/j.foodqual.2004.02.009

Hetherington, M. M., Anderson, A. S., Norton, G. N. M., & Newson, L. (2006). Situational effects on meal intake: A comparison of eating alone and eating with others. *Physiology & Behavior*, *88*(4), 498–505. https://doi.org/10.1016/j.physbeh.2006.04.025

Hinton, E. C., Brunstrom, J. M., Fay, S. H., Wilkinson, L. L., Ferriday, D., Rogers, P. J., & de Wijk, R. (2013). Using photography in 'The Restaur ant of the Future'. A useful way to assess portion selection and plate cleaning? *Appetite*, *63*, 31–35. https://doi.org/10.1016/j.appet.2012.12.008

Holthuysen, N. T. E., Vrijhof, M. N., de Wijk, R. A., & Kremer, S. (2017). "Welcome on board": Overall liking and just-about-right ratings of airplane meals in three different consumption contexts-laboratory, re-created airplane, and actual airplanethuysen. *Journal of Sensory Studies*, (December 2016), e12254. https://doi.org/10.1111/joss.12254

Iborra-Bernad, C., Saulais, L., Petit, E., & Giboreau, A. (2018). Sensory analysis and observational study in an experimental restaurant: Pilot study. *International Journal of Gastronomy and Food Science*, *13*, 47–51. https://doi.org/10.1016/j.ijgfs.2018.05.004

820 Jaeger, S. R., Fiszman, S., Reis, F., Chheang, S. L., Kam, K., Pineau, B., … Ares, G. (2017a).
821 Influence of evoked contexts on hedonic product discrimination and sensory characterizations
822 using CATA questions. *Food Quality and Preference*, *56*, 138–148.
823 https://doi.org/10.1016/j.foodqual.2016.10.003

824 Jaeger, S. R., Hort, J., Porcherot, C., Ares, G., Pecore, S., & MacFie, H. J. H. (2017b). Future
825 directions in sensory and consumer science: Four perspectives and audience voting. *Food*
826 *Quality and Preference*, *56*, 301–309. https://doi.org/10.1016/j.foodqual.2016.03.006

827 Jaeger, S. R., & Porcherot, C. (2017c). Consumption context in consumer research: methodological
828 perspectives. *Current Opinion in Food Science*, *15*, 30–37.
829 https://doi.org/10.1016/j.cofs.2017.05.001

830 Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring
831 and defining the experience of immersion in games. International Journal of Human Computer
832 Studies, 66(9), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004

833 Jimenez, M., Rodriguez, D., Greene, N., Zellner, D. A., Cardello, A. V, & Nestrud, M. (2015). Seeing
834 a meal is not eating it : Hedonic context effects differ for visually presented and actually eaten
835 foods. *Food Quality and Preference*, *41*, 96–102. https://doi.org/10.1016/j.foodqual.2014.11.015

836 Kempen, E., Kasambala, J., Christie, L., Symington, E., Jooste, L., & Van Eeden, T. (2017).
837 Expectancy-value theory contributes to understanding consumer attitudes towards cow's milk
838 alternatives and variants. *International Journal of Consumer Studies*, *41*(3), 245–252.
839 https://doi.org/10.1111/ijcs.12331

840 Kim, S.-E., Lee, S. M., & Kim, K.-O. (2016). Consumer acceptability of coffee as affected by
841 situational conditions and involvement. *Food Quality and Preference*, *52*, 124–132.
842 https://doi.org/10.1016/j.foodqual.2016.04.008

843 Kim, Y.-K., Jombart, L., Valentin, D., & Kim, K.-O. (2015). Familiarity and liking playing a role on
844 the perception of trained panelists: A cross-cultural study on teas. *Food Research International*,
845 *71*, 155–164. https://doi.org/10.1016/j.foodres.2015.03.022

846 King, S. C., Meiselman, H. L., Hottenstein, A. W., Work, T. M., & Cronk, V. (2007). The effects of
847 contextual variables on food acceptability: A confirmatory study. *Food Quality and Preference*,
848 *18*(1), 58–65. https://doi.org/10.1016/j.foodqual.2005.07.014

849 King, S. C., Weber, A. J., Meiselman, H. L., & Lv, N. (2004). The effect of meal situation, social
850 interaction, physical environment and choice on food acceptability. *Food Quality and*
851 *Preference*, *15*(7–8), 645–653. https://doi.org/10.1016/j.foodqual.2004.04.010

852 Köster, E. P. (2003). The psychology of food choice: some often encountered fallacies. *Food Quality*
853 *and Preference*, *14*(5–6), 359–373. https://doi.org/10.1016/S0950-3293(03)00017-X

854 Köster, E. P. (2009). Diversity in the determinants of food choice: A psychological perspective. Food
855 Qual Prefer, 20. https://doi.org/10.1016/j.foodqual.2007.11.002

856 Köster, E. P., & Mojet, J. (2012). Flops analysis: A useful tool for future innovations (Part 2: The
857 reduction of future flop risks). *Agro Food Industry Hi-Tech*, *23*(2), 6–10.

858 Kozlowska, K., Jeruszka, M., Matuszewska, I., Roszkowski, W., Barylko-Pikielna, N., & Brzozowska,
859 A. (2003). Hedonic tests in different locations as predictors of apple juice consumption at home
860 in elderly and young subjects. *Food Quality and Preference*, *14*(8), 653–661.
861 https://doi.org/10.1016/S0950-3293(02)00207-0

862 Kwak, H. S., Ahn, B. H., Lee, Y., Kreger, J., & Lee, S.-Y. (2013). Correlation of liking and disliking
863 measurements in consumer acceptance tests. *Food Quality and Preference*, *30*(2), 86–92.
864 https://doi.org/10.1016/j.foodqual.2013.05.002

865 Kwak, H. S., & Lee, S.-Y. (2016). Presentation methods for unidirectional scales to measure

consumers' liking and disliking percepts. *Food Quality and Preference*, *51*, 20–26. https://doi.org/10.1016/j.foodqual.2016.02.016

Lim, J. (2011). Hedonic scaling: A review of methods and theory. *Food Quality and Preference*, *22*(8), 733–747. https://doi.org/10.1016/j.foodqual.2011.05.008

Lin, I. Y., & Mattila, A. S. (2010). Restaurant Servicescape, Service Encounter, and Perceived Congruency on Customers' Emotions and Satisfaction. *Journal of Hospitality Marketing & Management*, *19*(8), 819–841. https://doi.org/10.1080/19368623.2010.514547

Liu, R., Hooker, N. H., Parasidis, E., & Simons, C. T. (2017). A Natural Experiment: Using Immersive Technologies to Study the Impact of "All-Natural" Labeling on Perceived Food Quality, Nutritional Content, and Liking. *Journal of Food Science*, *82*(3), 825–833. https://doi.org/10.1111/1750-3841.13639

Lusk, K. A., Hamid, N., Delahunty, C. M., & Jaeger, S. R. (2015). *Effects of an evoked refreshing consumption context on hedonic responses to apple juice measured using best–worst scaling and the 9-pt hedonic category scale*. *Food Quality and Preference* (Vol. 43). https://doi.org/10.1016/j.foodqual.2015.01.007

Mahon, D., Cowan, C., & McCarthy, M. (2006). The role of attitudes, subjective norm, perceived control and habit in the consumption of ready meals and takeaways in Great Britain. *Food Quality and Preference*, *17*(6), 474–481. https://doi.org/10.1016/j.foodqual.2005.06.001

Marshall, D., & Bell, R. (2003). Meal construction: exploring the relationship between eating occasion and location. *Food Quality and Preference*, *14*(1), 53–64. https://doi.org/10.1016/S0950-3293(02)00015-0

Matuszewska, I., Baryłko-Pikielna, N., Szczecinska, A., & Radzanowska, J. (1997). Comparison of three procedures for consumer assessment of fat spreads : Short report. *Polish Journal of Food and Nutrition Sciences*, *6*(3), 139–142. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=2815534

Meiselman, H. L. (1992). Obstacles to studying real people eating real meals in real situations. *Appetite*, *19*, 84–86.

Meiselman, H. L. (1996). The contextual basis for food acceptance, food choice and food intake: the food, the situation and the individual. In *Food choice, acceptance and consumption* (pp. 239–263). https://doi.org/10.1007/978-1-4613-1221-5_6

Meiselman, H. L. (2006). The role of context in food choice, food acceptance and food consumption. In R. Shepherd & M. Raats (Eds.), *The Psychology of Food Choice* (pp. 179–199). London.

Meiselman, H. L., Johnson, J. L., Reeve, W., & Crouch, J. E. (2000). Demonstrations of the influence of the eating environment on food acceptance. *Appetite*, *35*(3), 231–237. https://doi.org/10.1006/appe.2000.0360

Morizet, D., Depezay, L., Combris, P., Picard, D., & Giboreau, A. (2012). Effect of labeling on new vegetable dish acceptance in preadolescent children. *Appetite*, *59*(2), 399–402. https://doi.org/10.1016/j.appet.2012.05.030

Ouellet, É., Boller, B., Corriveau-Lecavalier, N., Cloutier, S., & Belleville, S. (2018). The Virtual Shop: A new immersive virtual reality environment and scenario for the assessment of everyday memory. *Journal of Neuroscience Methods*, *303*, 126–135. https://doi.org/10.1016/J.JNEUMETH.2018.03.010

Oxford Dictionaries. English Oxford Living Dictionaries. https://en.oxforddictionaries.com/ (2nd July 2018)

Pagliarini, E., Gabbiadini, N., & Ratti, S. (2005). Consumer testing with children on food combinations for school lunch, *16*, 131–138. https://doi.org/10.1016/j.foodqual.2004.03.001

912 Parsons, T. D. (2015). Virtual Reality for Enhanced Ecological Validity and Experimental Control in
913 the Clinical, Affective and Social Neurosciences. Frontiers in Human Neuroscience,
914 9(December), 1–19. https://doi.org/10.3389/fnhum.2015.00660

915 Petit, C., & Sieffermann, J. M. (2007). Testing consumer preferences for iced-coffee: Does the
916 drinking environment have any influence? *Food Quality and Preference*, *18*(1), 161–172.
917 https://doi.org/10.1016/j.foodqual.2006.05.008

918 Piqueras-Fiszman, B., Alcaide, J., Roura, E., & Spence, C. (2012). Is it the plate or is it the food?
919 Assessing the influence of the color (black or white) and shape of the plate on the perception of
920 the food placed on it. *Food Qual Prefer*, *24*. https://doi.org/10.1016/j.foodqual.2011.08.011

921 Pliner, P. L., Bell, R., Road, M., Bell, R., & Meiselman, H. L. (2004). Workshop summary : What to
922 eat : A multi-discipline view of meals, *15*, 901–905.
923 https://doi.org/10.1016/j.foodqual.2004.05.003

924 Posri, W., Macfie, H., & Henson, S. (2001). Improving the predictability of consumer liking from
925 central location test in tea. In *The 4th Pangborn Sensory Science Symposium: A sense odyssey*.

926 Prescott, J., Lee, S. M., & Kim, K. O. (2011). Analytic approaches to evaluation modify hedonic
927 responses. *Food Quality and Preference*, *22*(4), 391–393.
928 https://doi.org/10.1016/j.foodqual.2011.01.007

929 Roe, B. E., & Just, D. R. (2009). Internal and external validity in economics research: Tradeoffs
930 between experiments, field experiments, natural experiments, and field data. *American Journal*
931 *of Agricultural Economics*, *91*(5), 1266–1271. https://doi.org/10.1111/j.1467-8276.2009.01295.x

932 Rolls, B. J., & Shide, D. J. (1992). Both naturalistic and laboratory-based studies contribute to the
933 understanding of human eating behavior. *Appetite*. https://doi.org/10.1016/0195-6663(92)90240-
934 7

935 Rozin, P., & Tuorila, H. (1993). Simultaneous and temporal contextual influences on food acceptance.
936 *Food Quality and Preference*, *4*(1), 11–20. https://doi.org/10.1016/0950-3293(93)90309-T

937 Saulais, L., Muller, L. & Lesgards, V. (2017). Whispering in the ears of… companies ? Experimental
938 Economics as a Tool for decision in the private sector [Murmurer à l'oreille... de l'industriel ?
939 L'économie expérimentale comme outil d'aide à la décision en entreprise]. *Revue économique*,
940 vol. 68,(5), 925-939. doi:10.3917/reco.pr2.0086.

941 Schnack, A., Wright, M. J., & Holdershaw, J. L. (2018). Immersive virtual reality technology in a
942 three-dimensional virtual simulated store: Investigating telepresence and usability. *Food*
943 *Research International*. https://doi.org/10.1016/J.FOODRES.2018.01.028

944 Sester, C., Deroy, O., Sutan, A., Galia, F., Desmarchelier, J., Valentin, D., & Dacremont, C. (2013). '"
945 Having a drink in a bar "': An immersive approach to explore the effects of context on drink
946 choice, *28*, 23–31. https://doi.org/10.1016/j.foodqual.2012.07.006

947 Shogren, J. F. (2005). Chapter 19 Experimental Methods and Valuation. *Handbook of Environmental*
948 *Economics*, *2*, 969–1027. https://doi.org/10.1016/S1574-0099(05)02019-X

949 Siegrist, M., Ung, C.-Y., Zank, M., Marinello, M., Kunz, A., Hartmann, C., & Menozzi, M. (2018).
950 Consumers' food selection behaviors in three-dimensional (3D) virtual reality. *Food Research*
951 *International*. https://doi.org/10.1016/J.FOODRES.2018.02.033

952 Sinesio, F., Saba, A., Peparaio, M., Saggia Civitelli, E., Paoletti, F., & Moneta, E. (2018). Capturing
953 consumer perception of vegetable freshness in a simulated real-life taste situation. *Food*
954 *Research International*, *105*, 764–771. https://doi.org/10.1016/J.FOODRES.2017.11.073

955 Stelick, A., & Dando, R. (2018). Thinking outside the booth — the eating environment , context and
956 ecological validity in sensory and consumer research. *Current Opinion in Food Science*, *21*, 26–

31. https://doi.org/10.1016/j.cofs.2018.05.005

Stroebele, N., & de Castro, J. M. (2006). Listening to music while eating is related to increases in people's food intake and meal duration. *Appetite*, *47*(3), 285–289. https://doi.org/10.1016/j.appet.2006.04.001

Stroebele, N., & De Castro, J. M. (2004). Effect of ambience on food intake and food choice. *Nutrition*, *20*(9), 821–838. https://doi.org/10.1016/j.nut.2004.05.012

Tuorila, H., Palmujoki, I., Kytö, E., Törnwall, O., & Vehkalahti, K. (2015). School meal acceptance depends on the dish, student, and context. *Food Quality and Preference*, *46*, 126–136. https://doi.org/10.1016/j.foodqual.2015.07.013

Trochim,W.M. (2006). Introduction to Validity. http://www.socialresearchmethods.net/kb/introval.php (24th September 2018).

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice : A Reference-Dependent Model Author ( s ): Amos Tversky and Daniel Kahneman Published by : Oxford University Press. *The Quarterly Journal of Economics*, *106*(4), 1039–1061. https://doi.org/10.2307/2937956

Ung, C.-Y., Menozzi, M., Hartmann, C., & Siegrist, M. (2018). Innovations in consumer research: The virtual food buffet. *Food Quality and Preference*, *63*, 12–17. https://doi.org/10.1016/J.FOODQUAL.2017.07.007

van Herpen, E., van den Broek, E., van Trijp, H. C. M., & Yu, T. (2016). Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. *Appetite*, *107*, 196–207. https://doi.org/10.1016/J.APPET.2016.07.033

Weber, A. J., King, S. C., & Meiselman, H. L. (2004). Effects of social interaction, physical environment and food choice freedom on consumption in a meal-testing environment. *Appetite*, *42*(1), 115–118. https://doi.org/10.1016/j.appet.2003.10.001

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments - A Presence Questionaire. Presence: Teleoperators and Virtual Environments, 3(???), 225–240. https://doi.org/https://doi.org/10.1162/105474698565686

Zeinstra, G. G., Koelen, M. A., Kok, F. J., & de Graaf, C. (2010). The influence of preparation method on children's liking for vegetables. *Food Quality and Preference*, *21*(8), 906–914. https://doi.org/10.1016/j.foodqual.2009.12.006