



HAL
open science

The pursuit of ecological validity through contextual methodologies

Adriana Galinanes Plaza, Julien J. Delarue, Laure Saulais

► **To cite this version:**

Adriana Galinanes Plaza, Julien J. Delarue, Laure Saulais. The pursuit of ecological validity through contextual methodologies. *Food Quality and Preference*, 2019, 73, pp.226-247. 10.1016/j.foodqual.2018.11.004 . hal-02625887

HAL Id: hal-02625887

<https://hal.inrae.fr/hal-02625887>

Submitted on 21 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

The pursuit of ecological validity through contextual methodologies

AUTHOR NAMES

Galiñanes Plaza, A.^{a,b}, Delarue, J.^{a,*}, Saulais, L.^{b,c}

AFFILIATIONS

^aUMR Ingénierie Procédés Aliments, AgroParisTech, INRA, Université Paris-Saclay, 91300 Massy, France

^bCenter 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 Ingénierie Procédés Aliments, AgroParisTech, INRA, Université Paris-Saclay, 91300 Massy, France.

35	Contents	
36	1. Introduction and background	3
37	2. The concept of validity in sensory and consumer studies	4
38	2.1. Evaluating the validity of an experiment: internal, external and ecological validity	4
39	2.2. Critical points in sensory and consumer studies	6
40	2.2.1. Experimental environment	6
41	2.2.2. Nature of the product	7
42	2.2.3. Selection of participants	8
43	2.2.4. Evaluation task	8
44	3. Increasing ecological validity: what do context studies say?	9
45	3.1. From laboratory to natural settings	9
46	3.2. Do context parameters play a role in the validity of data?	9
47	3.3. Key determinants of ecological validity: a literature review	10
48	3.3.1. Methodology	10
49	3.3.2. Main results	23
50	3.3.3. Experimental environment	23
51	3.3.4. Nature of the product	24
52	3.3.5. Selection of participants	25
53	3.3.6. Evaluation task	25
54	4. New methodological approaches: towards increased transferability?	26
55	5. Contribution	39
56	5.1. Research	39
57	5.2. Practical implications	40
58	6. Limitations	41
59	7. Conclusion & Perspectives	42
60	8. Acknowledgements	42
61	9. References	42

62
63
64
65

66 1. Introduction and background

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

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

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

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

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

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

101 2.1. Evaluating the validity of an experiment: internal, external and ecological validity

102 The experimental approach is used in various scientific fields concerned with individual behaviors. In
103 particular, consumer psychology and economics use experiments to investigate consumer behaviors
104 and preferences. In these fields, the role of theory in the experimental approach is significant, although
105 not systematic (for a discussion on the role of theory in experimental economics, the reader is directed
106 to Card, DellaVigna & Malmendier, 2011). For instance, experiments in economics aim to either (i)
107 test theoretical assumptions, (ii) generate data on a little known phenomenon or (iii) evaluate the
108 potential impact of policy scenarios or private sector innovations (Saulais, Muller & Lesgards, 2017).
109 In consumer psychology, experiments use conceptual models and psychology theories (Kempen et al.,
110 2017, Köster, 2009,) such as the Theory of Planned Behavior or the Expectancy-value theory (Ajzen,
111 1991). While studies in sensory science share this overall goal of better understanding consumer
112 behavior, they often focus on operational objectives, such as to support product development through
113 consumer tests.

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

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

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

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

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

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

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

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

163

164 Table 1 Definitions and quotes

<u>The different types of experiments</u>	<u>Notions commonly used to evaluate experimental data</u>
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)

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.” (Harrison & List, 2004, p.1010)

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)

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)

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)

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)

165

166 2.2.Critical points in sensory and consumer studies

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

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

178

179 2.2.1.Experimental environment

180 Context was defined by Meiselman, (2006) as the specific physical, social and situational conditions in
181 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 Leshner, 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 (Bangcuayo 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 (Harrison & 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

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

255

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

257 3.1. From laboratory to natural settings

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

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

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

277

278 3.2. Do context parameters play a role in the validity of data?

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

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

286 The studies following a global approach compare scores on food liking and choices in different natural
287 environments (restaurants, canteens, prisons) with those obtained on laboratory or central location
288 settings showing differences on hedonic scores (Edwards, Meiselman, Edwards, & Leshner, 2003;
289 King, Weber, Meiselman, & Lv, 2004; Meiselman, Johnson, Reeve, & Crouch, 2000). Those
290 differences are usually related to the degree of discrimination among products – consumers being
291 more discriminant in natural settings than in laboratory settings – or to the higher scores on natural
292 settings versus laboratory settings. The studies focusing on context variables compare how the
293 addition of contextual variables in controlled experiments affect food liking and choice (King et al.,
294 2004; Stroebele & De Castro, 2004; Weber, King, & Meiselman, 2004). We may first notice that
295 several classifications of contextual variables have been proposed: Rozin & Tuorila (1993) divide
296 contextual variables into either product and non-product variables and subdivide them in simultaneous
297 and temporal contextual factors; Meiselman (1996), proposes to distinguish between three categories
298 of variables (the situation, the individual and the product); whereas Stroebele & De Castro (2004),
299 divide the contextual variables into social context variables, physical surroundings, time related
300 characteristics and distraction and/or television viewing. From these studies, it is difficult to fully
301 disentangle the various factors and isolate a specific context effect. The relevance of those contextual
302 variables thus remains unclear. To date, the lack of knowledge of the combined effects of these
303 contextual variables on consumers' responses compromises the ability to identify causal relationships
304 through experimental approaches. In practice, a consequence of this is that participants to a test may
305 not perceive the study context the way the researcher assumes they would. This questions the
306 ecological validity as defined by Brofenbrenner.

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

310

311 3.3.Key determinants of ecological validity: a literature review

312 3.3.1. Methodology

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

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

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

331

332

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 Participants were pre-recruited and remunerated 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 No food type comparison
Meiselman et al., (2000)	Food acceptance	Eating location	Cross-cultural study: 74 and 125 participants (UK data) Participants were recruited for all locations except sensory laboratory Between-group design	<ul style="list-style-type: none"> • Menu based on canned food • Menu main dish Chicken fettuccine Alfredo 	<ul style="list-style-type: none"> • Training restaurant vs Student cafeteria (UK); • Training rest. vs Food lab vs cafeteria (USA) 	Food acceptance on a 9-point hedonic scale (UK) 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. The nature of the task (questionnaire distribution and number of questions) differed between and within contexts. 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	<p>1124 regular clients</p> <p>Participants were unaware of the study and not remunerated</p> <p>Between-group design</p>	Regular meals	<ul style="list-style-type: none"> • Worksite cafeteria • Moderately priced restaurant • Fast-food restaurant 	<p>Purely observational</p> <p>No questionnaire</p>	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	<p>Participants:</p> <ul style="list-style-type: none"> • Army training camp: 44 • University staff refectory/ 38 • Private boarding school: 88 • Freshman's buffet: 83 • Private party: 78 • Residential home (elderly): 43 • Student refectory: 33 • Day care center (elderly): 33 • University 4-star restaurant: 19 • Hotel 4-star restaurant: 32 <p>Participants were unaware of the current study and not remunerated</p> <p>Between-group design</p>	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)	<p>Food preparation is context dependent; therefore, acceptability may differ from one context to another due to sensory properties modification.</p> <p>Contexts also differed on service style, dinning, choice, etc.</p> <p>=> Ecological validity is ensured however results may not be comparable across contexts</p>

Kozłowska et al., (2003)	Predictive value of hedonic test	Eating location	35 elderly people 33 young people Participants were recruited Between and within-group design	5 apple juices with different sugar concentrations	<ul style="list-style-type: none"> • Central Location Test (CLT) • 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. 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. Scope for choice differs across contexts, although it may be key to ecological validity 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 Product type	Participants: regular users of the tested products <ul style="list-style-type: none"> • Study 1: 240/context • Study 2: 240/context • Study 3: 160/context Participants were recruited in-situ (CLT) or pre-recruited by phone (HUT) Between-group design	2 variants of each product: <ul style="list-style-type: none"> • Milk beverage • Salted crackers • Sparkling water Products were sequentially presented	<ul style="list-style-type: none"> • CLT • HUT 	Overall liking scores on a 10-point hedonic scale. CLT: 2 products evaluated during 1 session HUT: After one week of testing participants got the second product and repeated the same task	Products got higher scores at the HUT . The influence of the method used depends on the type of product (how products are usually eaten) Pure monadic were slightly higher than monadic sequential scores	The evaluation task differed between contexts (time for evaluation); 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: <ul style="list-style-type: none"> • Non label: 125 • Basic label: 116 • Moderated Label: 131 Participants were recruited Between-group design	Carrot and broccoli dishes	Three school canteens: <ul style="list-style-type: none"> • School 1: 140 • School 2: 113 • School 3: 111 	Chef give or not information about the vegetable options Food choice at lunch time	Children chose significantly more often the familiar version of the dish when no information was given 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 Food accompaniment	55 participants: likers of wine Participants were pre-recruited Within-group design	Eight different wines + dummy wine Food accompaniment : crackers with cheese, carrots and broccoli with dip and tortilla chips with mild salsa	4 contexts: <ul style="list-style-type: none"> • Laboratory: with/without food • 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) 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 Eating location	Participants: regular consumer of the tested products	Side salad with dressing Small pizza Iced tea	6 contexts: <ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • Test 1: 104 • Test 2: 93 • Test 3: 106 • Test 4: 106 • Test 5: 101 • Test 6: 35 <p>Participants were recruited (CLT) and/or contacted by phone or advertisements at the local offices (Test 3-5)</p> <p>Between-within group design</p>	Each component has two flavor variants	<ul style="list-style-type: none"> • T2: Laboratory + meal • T3: Laboratory + meal + social interaction • T4: Mock restaurant + meal + social interaction • T5: Mock restaurant + meal + social interaction + choice • T6: Restaurant 	items component on a 9-point structured hedonic scale 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. 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 Eating location Cutlery Choice	<p>Participants: regular consumer of the tested products</p> <ul style="list-style-type: none"> • Test 1: 93 • Test 2: 106 • Test 3: 106 • Test 4: 101 <p>Participants were pre-recruited</p> <p>Between-within group design</p>	Pizza + salad + tea: 2 variants of each product	4 contexts: <ul style="list-style-type: none"> • T1: Laboratory + meal • T2: Mock restaurant + meal + social interaction • 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

					<ul style="list-style-type: none"> • 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 Participants were pre-recruited Within-group design	Hard and semi-hard Norwegian cheeses	3 contexts: <ul style="list-style-type: none"> • Laboratory • Club house of a soccer team • 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. The evaluation task differed between contexts (time for evaluation); 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 Participants were pre-recruited and remunerated Within-group design	Different type of products: bread rolls, potato crisps, fresh green salad, etc.	3 contexts: <ul style="list-style-type: none"> • T1: Laboratory • T2: Laboratory + TV • 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

					<ul style="list-style-type: none"> • T4: Laboratory + positive social facilitation 			
Stroebele & de Castro, (2006)	Food intake and meal duration	Music	78 participants Participants were pre-recruited 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 <ul style="list-style-type: none"> • Test 1: 74 • Test 2: 83 • Test 3: 386 Participants were pre-recruited on test 1 and 2 Between-group design	<ul style="list-style-type: none"> • Iced tea • tossed salad with Italian dressing • garlic bread sticks • cannelloni with meat filling • meat lasagna 	3 contexts: <ul style="list-style-type: none"> • T1: CLT • T2: national Italian chain restaurant • 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 effect on the acceptability scores	The number of participants highly differs among contexts. 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: <ul style="list-style-type: none"> • 46: 4-6ys • 25: 7-8ys • 23: 11-12ys • 22: 18-25ys Participants were pre-recruited Between-group design	Carrots and French beans: <ul style="list-style-type: none"> • mashed • steamed • boiled • stir-fried • grilled • 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. 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-Fizman, Alcaide, Roura, & Spence, (2012)	Food perception	Product presentation: shape and color of the plate	Participants: <ul style="list-style-type: none"> • Study 1: 53 • Study 2: 51 Participants were recruited Within-group design	Strawberry mouse	Laboratory: <ul style="list-style-type: none"> • Study 1: white vs black dishes • 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 <ul style="list-style-type: none"> • Test 1: 32 • Test 2: 33 • Test 3: 30 Participants were pre-recruited Between-group design	5 frozen croissants	3 contexts: <ul style="list-style-type: none"> • T1: Laboratory • T2: Laboratory + social facilitation • T3: Laboratory + social facilitation + drink options 	Overall liking + attribute liking (appearance, odor, flavor, taste and texture.) on a 9-point hedonic scales Freshness on a 9-point scale	Social interaction negatively affected all the liking scores when compared to the control group Social + drink accompaniments seemed to increase hedonic scores	Ecological validity could be questioned since participants did not know each other. However, evaluating croissants with a drink is closer to usual eating habits

García-Segovia et al., (2015)	Food acceptance and intake	Eating location Table setting	Participants: <ul style="list-style-type: none"> • Test 1: 94 • Test 2: 90 Participants were pre-recruited	Ready-to-eat (RTE) herb-roasted chicken	2 contexts: <ul style="list-style-type: none"> • T1: Laboratory • T2: More realistic context (Room, experimental home-style dining room, experimental restaurant) 3 table settings: <ul style="list-style-type: none"> • plastic tray • home-style table • gourmet 	Hunger status on a 9-point Likert scale before and after eating <ul style="list-style-type: none"> • 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 • 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 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 Choice	Participants and soldiers: <ul style="list-style-type: none"> • T1: 199 • T2: 36 • T3: 36 	<ul style="list-style-type: none"> • Main dishes (unfamiliar and familiar) • 4 groupings 	3 contexts: <ul style="list-style-type: none"> • Field study (military camp) • Laboratory 	Liking ratings on a 9-point hedonic scale 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 Between-group design	of snacks (sweet commercial, savory commercial, sweet military, savory military), • 3 entire meal menus	• Laboratory + choice	soldiers chose three meals for the that included some of the tested products and evaluated them Participants chose three menus in a short period of time. Laboratory: participants tested small bites of the stimuli 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. 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. 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: • T1: 96 • T2: 96 • T3: 52 • T4: 55 Participants were recruited for the laboratory study Between-group design	Iced coffee: water and milk based	4 contexts: • T1: Laboratory • T2: Laboratory + contextual elements (curtains, pictures, etc.) • T3: meeting room • T4: cafeteria	• Overall liking and ranging on a 21-point scale + short questionnaire (demographic + consumption habits)	No significant differences between T1 and T2 Significant product differences in T3 Differences between laboratory and natural settings	The nature of the product differs in each context as the preparation method differs. 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...). 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

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 (Kozłowska 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.

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

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

378

379 3.3.4. Nature of the product

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

388 As regards the effect of food combination and sequence of food items during a meal, it is interesting to
389 notice that most of the research on human eating behavior has been focused on food items instead of
390 food combinations. In the last decades, researchers have shown that suitable food combinations result
391 in more pleasant recipes and this is translated in higher overall hedonic scores (Di Monaco et al.,
392 2014; Elzerman, Hoek, Boekel, & Luning, 2011; Hersleth et al., 2003; Pagliarini, Gabbiadini, & Ratti,
393 2005). While others have also studied how much of each meal component contributes to that (Jimenez
394 et al., 2015; Meiselman, 2006). In addition to this, the sequence and appropriateness of mealtimes
395 when evaluating products has produced different results (Boutrolle et al., 2007; Cardello, Schutz,
396 Snow, & Leshner, 2000; King et al., 2004; Meiselman, 2006). Therefore, the study of products as food
397 items instead of part of a meal may contribute to misleading results that cannot be generalized from
398 one context to another.

399 Another important aspect that has been already mentioned is that consumers and locations are most
400 often confounded variables. They cannot be studied independently as they are intimately related to
401 consumers' expectations and mindset. Even if a food is exactly the same in two different contexts,
402 consumers may not bring to those contexts the same experience, beliefs and/or expectations. As a
403 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 where 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; Kozłowska 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

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

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

451

452 4. **New methodological approaches: towards increased transferability?**

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

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

462 (a) *Living labs* - Even if a no clear definition for Living labs is found in the literature, the authors
463 have decided to use the definition given by Dell'Era & Landoni (2014) (p.139) where Living
464 Lab is defined as "a design research methodology aimed at co-creating innovation through the
465 involvement of aware users in a real-life setting". In Living labs, the researcher can control
466 and record a selected number of contextual variables and the interaction between them, within
467 a natural consumption situation. Living lab experiments can be seen as an attempt to
468 compromise with the limitations and advantages of laboratory and field experiments, as the
469 control of contextual variables increases the internal validity of the study, while the situation is
470 kept as ecological as possible. Examples of Living labs dedicated to food studies are "The
471 Restaurant of the Future" Wageningen, Netherlands (Hinton et al., 2013; Zeinstra et al., 2010),
472 "The Grill Room" in Bournemouth, United Kingdom (Bell, Meiselman, Pierson, & Reeve,
473 1994; Meiselman et al., 2000) and "The Living Lab" at the Research Centre of the Institute

474 Paul Bocuse in Ecully, France (Allirot et al., 2014; Iborra-Bernad, Saulais, Petit, & Giboreau,
475 2018).

476 (b) *Evoked contexts* - In the evoked contexts approach, the researcher places the consumer in a
477 typical laboratory evaluation task, but uses either text, audio recordings, and/or pictures that
478 evoke what would be a natural consumption situation of the product (Jaeger & Porcherot,
479 2017c). In this case, consumers have to imagine themselves in a particular situation and
480 evaluate a product or a set of products. This approach is well established in other disciplines
481 such as marketing studies (Bitner, 1990; Daunt & Greer, 2015; Esmark, Noble, & Breazeale,
482 2017).

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

497 (d) *Virtual reality* - Finally, the virtual reality approach is defined by the “use of virtual
498 environments to present digitally recreated real world activities to participants via immersive
499 (head-mounted displays) and non-immersive (2D computer screens) mediums” (Parson,
500 2015). We argue that including non-immersive mediums such as the 2D computer screen in
501 the virtual reality definition may create certain confusion with the evoked and immersive
502 contexts categories, where such tools can be also used. For this reason, we have considered
503 only virtual reality studies where head-mounted displays are used.

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

507

508

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 <ul style="list-style-type: none"> • Test 1: 63 • Test 2: 75 Participants were neither recruited nor remunerate Between-group design	A full restaurant menu	2 contexts: <ul style="list-style-type: none"> • T1: regular decoration and British food names • 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. Scope for choice increase ecological validity of both studies
Allirot et al., (2014)	Food intake	Food choice	17 participants Participants were recruited and remunerated 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: <ul style="list-style-type: none"> • Experimental restaurant • 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. No comparison with natural context 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 <ul style="list-style-type: none"> • Test 1: 72 • Test 2: 70 Participants were recruited and remunerated Between-group design	4 apple juices: different concentrations of citric acid and strawberry flavor	2 contexts: <ul style="list-style-type: none"> • T1: laboratory • T2: evoked refreshing drink occasion 	<ul style="list-style-type: none"> • 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 • 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. 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. 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 <ul style="list-style-type: none"> • Test 1: 64 • Test 2: 62 • Test 3: 63 • Test 4: 68 	Appel and blackcurrant juice: different concentrations of citric acid and strawberry flavor/ sucrose and orange flavor	4 contexts: <ul style="list-style-type: none"> • T1: laboratory • T2: evoked refreshing drink occasion • T3: evoked breakfast 	Similar task to the previous study. <ul style="list-style-type: none"> • 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 The experimental environment might be different for each participant making difficult context comparison

			Participants were recruited		situation		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.
			Between-group design		<ul style="list-style-type: none"> T4: evoked movie situation 	<ul style="list-style-type: none"> T2, T3, T4: one more question about context appropriateness 		
Giacalone et al., (2015)	Situational appropriateness	Familiarity degree Eating location	<p>Participants:</p> <ul style="list-style-type: none"> Study 1: 76 Study 2: 97 Study 3: 93 Study 4: 145 <p>Participants were recruited and remunerated</p> <p>Within-group design</p>	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	<p>120 participants: regular consumers of the tested product</p> <p>Participants were recruited</p> <p>Within-group design</p>	6 types of dry-cured ham	<p>2 contexts:</p> <ul style="list-style-type: none"> Evoked traditional meal Evoked novel meal 	Evaluation of intrinsic characteristics on a 9-point hedonic scale. + 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. 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 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: <ul style="list-style-type: none"> • Study 1: 65 • Study 2: 48 Participants were recruited and remunerated Between-group design	4 apple juices: <ul style="list-style-type: none"> • 2 common • 2 premium) 	Evoked refreshing drink occasion	<ul style="list-style-type: none"> • S1: Overall liking on a best-worst scaling + questions about task complexity and response accuracy • 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: <ul style="list-style-type: none"> • Study 1: 157 • Study 2: 171 Participants were recruited Between-group design	Lamb meat	2 contexts: <ul style="list-style-type: none"> • A celebratory lunch with family over the weekend (weekend lunch context) • 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 Participants were recruited	2 types of coffee	4 contexts: <ul style="list-style-type: none"> • T1: Laboratory • T2: Laboratory + evocation • T3: 	<ul style="list-style-type: none"> • T1: overall liking on a 9-point hedonic scale. • 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	point category scale + liking scores on a 9-point hedonic scale + involvement questionnaire	evocation 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 Participants were recruited Between-subject design	Food and drinks familiar for the studied population	2 contexts: • T1: Laboratory • T2: Evoked drinking a particular product • 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 The evoked context did not systematically influence the results 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: • Study 1A: 93 • Study 1B: 83 • Study 2A: 60 • Study 2B: 60 Participants were recruited for the study 2	Drinks	4 Contexts • S1,2A: immersive bar with warmth furniture • 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 • Study 2A, B: triangle test + questionnaire	S1A, B: immersive contexts influenced drink choice (association between drinks and video clips). 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). Results should be replicate in another immersive or natural environments

			Within- and between-group design		furniture • 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 Participants were recruited and remunerated Within-subject design	5 coffees	2 contexts: • Laboratory • Immersive café	• Laboratory: demographic questionnaire + acceptability on a 9-point hedonic scale + ranking + Engagement Questionnaire • 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) Immersive coffee increased engagement	The nature of the immersive context may be representative for a coffee drinking situation 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 Participants were recruited and remunerated Within-subject design	4 commercially available chocolate chip cookies	4 contexts: • Laboratory • Mixed immersive domestic kitchen • Full immersion domestic kitchen	Demographic questionnaire + acceptability on a 9-point hedonic scale + Engagement questionnaire	Participants were more discriminants in the full immersion context. Similar levels of engagement in the two immersive conditions 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	<p>Participants:</p> <ul style="list-style-type: none"> • Study 1, 2: 242 • Study 3: 222 <p>Participants were recruited for the controlled and recreated condition</p> <p>Between-group design</p>	2 airplane meals: 2 variants of each	3 contexts: <ul style="list-style-type: none"> • S1: Laboratory • S2: Recreated airplane • S3: Real airplane 	<ul style="list-style-type: none"> • S1: selection of the product + overall liking on a Visual Analogue Scale. • 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 	<p>Recreated and actual airplane were more discriminant than the controlled condition.</p> <p>No significant difference on ratings between recreated and natural context were found</p>	First study that compares new approaches to natural conditions
Liu, Hooker, Parasidis, & Simons, (2017)	Food quality, nutritional content, liking	Labelling	<p>120 participants: regular peanut butter consumers divided in 3 labelling conditions</p> <p>Participants were recruited and remunerated</p> <p>Between-group design</p>	Peanut butter	Immersive context: virtual grocery store + 3 conditions: <ul style="list-style-type: none"> • Blind • Labelled • Labelled + verbal call-out 	Acceptability on a 9-point hedonic scale + WTP + Food quality questionnaire + demographic information	<p>Labelling improved product quality and nutritional content perception but not liking and WTP.</p> <p>Verbal call out improved food quality, nutritional content and WTP</p>	
Sinesio et al., (2018)	Food perception and liking	Eating location	<p>48 participants: regular consumers of the tested product</p> <p>Participants were recruited</p>	Salad tomato and wild rocket salad at different storage time	2 contexts: <ul style="list-style-type: none"> • Study 1: Laboratory • Study 2: Immersive holiday farm dining room 	<ul style="list-style-type: none"> • S1: overall liking on a 9-point hedonic scale + perceived freshness on a 9-point scale. • S2: same evaluation task + 	<p>Liking scores were higher in the immersive environment setting than in the lab.</p> <p>However, higher discrimination as regards storage time</p>	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 Participants were recruited Between-group design	Milk, fruit and vegetables and biscuits	3 contexts: <ul style="list-style-type: none"> • Real supermarket • 3D VR supermarket • Laboratory + supermarket pictures 	Same task in the 3 contexts: <ul style="list-style-type: none"> • 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 Participants were recruited Between and within-group design	Beverages and skin care lotion odor	3 contexts: <ul style="list-style-type: none"> • Laboratory • 3D VR beach • 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	<p>Participants:</p> <ul style="list-style-type: none"> • Study 1 :49 • Study 2:35 <p>Participants were recruited</p> <p>Between and within-group design</p>	List of products	Virtual supermarket	<ul style="list-style-type: none"> • S1: memorize a list of products + buy those products • 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	<p>Participants:</p> <ul style="list-style-type: none"> • Study 1: 62 • Study 2: 49 <p>Participants were recruited</p> <p>Between-group design</p>	Food products	<p>2 contexts:</p> <ul style="list-style-type: none"> • S1: 3D VR supermarket • 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	<p>Participants:</p> <ul style="list-style-type: none"> • Study 1A: 37 • Study 1B: 31 • Study 2: 50 <p>Participants were recruited and remunerated</p>	Cereals	<p>2 contexts:</p> <ul style="list-style-type: none"> • S1A: Real life supermarket • S1B: 3D VR supermarket • S2: 3D VR supermarket 	<ul style="list-style-type: none"> • S1A, B: similar task. Select one cereal package for kid's camp + one for a specific type of diet • 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 Participants were recruited and remunerated Within-subject design	3 types of foods (Fake food buffet)	2 contexts: • Laboratory buffet • 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. As regards the task (serving) this may be representative for the participants. However, a natural context comparison would be needed to prove those results

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

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

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

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

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

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

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

564

565 5. Contribution

566 5.1. Research

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

574 There has been a lot of research on the effects of context on consumers' hedonic response, food choice
575 or intake, however the overall inconsistency of findings renders difficult their integration into clear
576 guidelines to improve the ecological validity of a study. In particular, to date, the results are too
577 context-specific, product category specific or task specific to enunciate more general principles that
578 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

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

615 As regards the nature of the food, it is important to consider, especially in the context of new product
616 development, the type of product that the test aims to evaluate, at which stage of development process
617 the data is needed, and in which settings the final product will be consumed. It has been shown that the
618 impact of context depends on the product category and units of evaluation (e.g. product vs dish).

619 Moreover, familiarity towards the tested product seems to modulate the contexts effects: while a
620 product familiar to consumers can be eaten in several contexts, unfamiliar products can be related to
621 particular occasions and consumption contexts. In the early stages of product development, when
622 specific sensory product characteristics should be defined, laboratory settings should be considered as
623 the best solution. However, when it comes to the choice or purchase intention, more naturalistic
624 environments may be needed to ensure product success.

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

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

643

644 **6. Limitations**

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

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

654

655 7. Conclusion & Perspectives

656 Increasing the number of consumer studies in natural settings was pointed as one of the most
657 important challenges for research during the 11th Pangborn Sensory Science Symposium (Jaeger et al.,
658 2017b). In the past decades, sensory and consumer scientists have tried to move from laboratory
659 experiments to natural experiments and different alternative approaches, such as evoked or immersive
660 contexts or virtual reality, have emerged with the purpose of ensuring better ecological validity.

661 Ecological validity is achieved if participants perceive the experimental environment, the food they
662 taste and the task they perform to be representative of a natural consumption situation.

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

671

672 8. Acknowledgements

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

675

676 9. References

- 677 Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision*
678 *processes*, 50(2), 179-211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- 679 Alliot, X., Seyssel, K., Saulais, L., Roth, H., Charrié, A., Drai, J., ... Laville, M. (2014). Effects of a
680 breakfast spread out over time on the food intake at lunch and the hormonal responses in obese
681 men. *Physiology & Behavior*, 127, 37-44. <https://doi.org/10.1016/j.physbeh.2014.01.004>
- 682 Andersen, I. N. S. K., Kraus, A. A., Ritz, C., & Bredie, W. L. P. (2018). Desires for beverages and
683 liking of skin care product odors in imaginative and immersive virtual reality beach contexts.
684 *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-](https://doi.org/10.1016/S0195-6663(03)00109-0)
694 [6663\(03\)00109-0](https://doi.org/10.1016/S0195-6663(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., & Leshner, 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](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-](https://doi.org/10.1016/S0193-2306(04)10001-X)
728 [2306\(04\)10001-X](https://doi.org/10.1016/S0193-2306(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-](https://doi.org/10.1007/s10640-007-9124-5)
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-](https://doi.org/10.1016/0031-9384(94)90286-0)
740 9384(94)90286-0
- 741 De Graaf, C., Cardello, A. V., Matthew Kramer, F., Leshner, 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. Interdisciplinary research and business applications* (pp. 102–127).
765 Cambridge: Woodhead Publishing Limited.
- 766 Edwards, J. S. A., Meiselman, H. L., Edwards, A., & Leshner, 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](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>

- 775 Garber, L. L., Hyatt, E. M., & Starr, R. G. (2003). Measuring consumer response to food products.
776 *Food Quality and Preference*, 14(1), 3–15. [https://doi.org/10.1016/S0950-3293\(02\)00030-7](https://doi.org/10.1016/S0950-3293(02)00030-7)
- 777 García-Segovia, P., Harrington, R. J., & Seo, H.-S. (2015). Influences of table setting and eating
778 location on food acceptance and intake. *Food Quality and Preference*, 39, 1–7.
779 <https://doi.org/10.1016/j.foodqual.2014.06.004>
- 780 Giacalone, D., Frøst, M. B., Bredie, W. L. P., Pineau, B., Hunter, D. C., Paisley, A. G., ... Jaeger, S.
781 R. (2015). Situational appropriateness of beer is influenced by product familiarity. *Food Quality*
782 *and Preference*, 39, 16–27. <https://doi.org/10.1016/J.FOODQUAL.2014.06.012>
- 783 Fréchette, G. R., & Schotter, A. (2015). *Handbook of Experimental Economic Methodology*. Oxford:
784 Oxford University Press.
- 785 Guala, F. (2012). Experimentation in Economics. *Philosophy of Economics*, (June 2005), 597–640.
786 <https://doi.org/10.1016/B978-0-444-51676-3.50021-X>
- 787 Harrison, G. W., & List, J. a. (2004). Field Experiments. *Journal of Economic Literature*,
788 XLII(December), 1009–1055. <https://doi.org/10.1257/0022051043004577>
- 789 Hathaway, D., & Simons, C. T. (2017). The impact of multiple immersion levels on data quality and
790 panelist engagement for the evaluation of cookies under a preparation-based scenario. *Food*
791 *Quality and Preference*, 57, 114–125. <https://doi.org/10.1016/J.FOODQUAL.2016.12.009>
- 792 Hein, K. A., Hamid, N., Jaeger, S. R., & Delahunty, C. M. (2010). Application of a written scenario to
793 evoke a consumption context in a laboratory setting: Effects on hedonic ratings. *Food Quality*
794 *and Preference*, 21(4), 410–416. <https://doi.org/10.1016/j.foodqual.2009.10.003>
- 795 Hein, K. A., Hamid, N., Jaeger, S. R., & Delahunty, C. M. (2012). Effects of evoked consumption
796 contexts on hedonic ratings: A case study with two fruit beverages. *Food Quality and*
797 *Preference*, 26(1), 35–44. <https://doi.org/10.1016/j.foodqual.2012.02.014>
- 798 Hersleth, M., Mevik, B.-H., Næs, T., & Guinard, J.-X. (2003). Effect of contextual factors on liking
799 for wine—use of robust design methodology. *Food Quality and Preference*, 14(7), 615–622.
800 [https://doi.org/10.1016/S0950-3293\(02\)00190-8](https://doi.org/10.1016/S0950-3293(02)00190-8)
- 801 Hersleth, M., Monteleone, E., Segtnan, A., & Næs, T. (2015). Effects of evoked meal contexts on
802 consumers' responses to intrinsic and extrinsic product attributes in dry-cured ham. *Food Quality*
803 *and Preference*, 40, 191–198. <https://doi.org/10.1016/j.foodqual.2014.10.002>
- 804 Hersleth, M., Ueland, Ø., Allain, H., & Næs, T. (2005). Consumer acceptance of cheese, influence of
805 different testing conditions. *Food Quality and Preference*, 16(2), 103–110.
806 <https://doi.org/10.1016/j.foodqual.2004.02.009>
- 807 Hetherington, M. M., Anderson, A. S., Norton, G. N. M., & Newson, L. (2006). Situational effects on
808 meal intake: A comparison of eating alone and eating with others. *Physiology & Behavior*, 88(4),
809 498–505. <https://doi.org/10.1016/j.physbeh.2006.04.025>
- 810 Hinton, E. C., Brunstrom, J. M., Fay, S. H., Wilkinson, L. L., Ferriday, D., Rogers, P. J., & de Wijk,
811 R. (2013). Using photography in 'The Restaurant of the Future'. A useful way to assess portion
812 selection and plate cleaning? *Appetite*, 63, 31–35. <https://doi.org/10.1016/j.appet.2012.12.008>
- 813 Holthuysen, N. T. E., Vrijhof, M. N., de Wijk, R. A., & Kremer, S. (2017). "Welcome on board":
814 Overall liking and just-about-right ratings of airplane meals in three different consumption
815 contexts-laboratory, re-created airplane, and actual airplane. *Journal of Sensory Studies*,
816 (December 2016), e12254. <https://doi.org/10.1111/joss.12254>
- 817 Iborra-Bernad, C., Saulais, L., Petit, E., & Giboreau, A. (2018). Sensory analysis and observational
818 study in an experimental restaurant: Pilot study. *International Journal of Gastronomy and Food*
819 *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., Tijds, 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](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 Kozłowska, 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](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

- 866 consumers' liking and disliking percepts. *Food Quality and Preference*, 51, 20–26.
867 <https://doi.org/10.1016/j.foodqual.2016.02.016>
- 868 Lim, J. (2011). Hedonic scaling: A review of methods and theory. *Food Quality and Preference*,
869 22(8), 733–747. <https://doi.org/10.1016/j.foodqual.2011.05.008>
- 870 Lin, I. Y., & Mattila, A. S. (2010). Restaurant Servicescape, Service Encounter, and Perceived
871 Congruency on Customers' Emotions and Satisfaction. *Journal of Hospitality Marketing &*
872 *Management*, 19(8), 819–841. <https://doi.org/10.1080/19368623.2010.514547>
- 873 Liu, R., Hooker, N. H., Parasidis, E., & Simons, C. T. (2017). A Natural Experiment: Using
874 Immersive Technologies to Study the Impact of “All-Natural” Labeling on Perceived Food
875 Quality, Nutritional Content, and Liking. *Journal of Food Science*, 82(3), 825–833.
876 <https://doi.org/10.1111/1750-3841.13639>
- 877 Lusk, K. A., Hamid, N., Delahunty, C. M., & Jaeger, S. R. (2015). *Effects of an evoked refreshing*
878 *consumption context on hedonic responses to apple juice measured using best–worst scaling and*
879 *the 9-pt hedonic category scale. Food Quality and Preference* (Vol. 43).
880 <https://doi.org/10.1016/j.foodqual.2015.01.007>
- 881 Mahon, D., Cowan, C., & McCarthy, M. (2006). The role of attitudes, subjective norm, perceived
882 control and habit in the consumption of ready meals and takeaways in Great Britain. *Food*
883 *Quality and Preference*, 17(6), 474–481. <https://doi.org/10.1016/j.foodqual.2005.06.001>
- 884 Marshall, D., & Bell, R. (2003). Meal construction: exploring the relationship between eating occasion
885 and location. *Food Quality and Preference*, 14(1), 53–64. [https://doi.org/10.1016/S0950-](https://doi.org/10.1016/S0950-3293(02)00015-0)
886 [3293\(02\)00015-0](https://doi.org/10.1016/S0950-3293(02)00015-0)
- 887 Matuszewska, I., Baryłko-Pikielna, N., Szczecinska, A., & Radzanowska, J. (1997). Comparison of
888 three procedures for consumer assessment of fat spreads : Short report. *Polish Journal of Food*
889 *and Nutrition Sciences*, 6(3), 139–142. Retrieved from
890 <http://cat.inist.fr/?aModele=afficheN&cpsid=2815534>
- 891 Meiselman, H. L. (1992). Obstacles to studying real people eating real meals in real situations.
892 *Appetite*, 19, 84–86.
- 893 Meiselman, H. L. (1996). The contextual basis for food acceptance, food choice and food intake: the
894 food, the situation and the individual. In *Food choice, acceptance and consumption* (pp. 239–
895 263). https://doi.org/10.1007/978-1-4613-1221-5_6
- 896 Meiselman, H. L. (2006). The role of context in food choice, food acceptance and food consumption.
897 In R. Shepherd & M. Raats (Eds.), *The Psychology of Food Choice* (pp. 179–199). London.
- 898 Meiselman, H. L., Johnson, J. L., Reeve, W., & Crouch, J. E. (2000). Demonstrations of the influence
899 of the eating environment on food acceptance. *Appetite*, 35(3), 231–237.
900 <https://doi.org/10.1006/appe.2000.0360>
- 901 Morizet, D., Depezay, L., Combris, P., Picard, D., & Giboreau, A. (2012). Effect of labeling on new
902 vegetable dish acceptance in preadolescent children. *Appetite*, 59(2), 399–402.
903 <https://doi.org/10.1016/j.appet.2012.05.030>
- 904 Ouellet, É., Boller, B., Corriveau-Lecavalier, N., Cloutier, S., & Belleville, S. (2018). The Virtual
905 Shop: A new immersive virtual reality environment and scenario for the assessment of everyday
906 memory. *Journal of Neuroscience Methods*, 303, 126–135.
907 <https://doi.org/10.1016/J.JNEUMETH.2018.03.010>
- 908 Oxford Dictionaries. English Oxford Living Dictionaries. <https://en.oxforddictionaries.com/> (2nd July
909 2018)
- 910 Pagliarini, E., Gabbiadini, N., & Ratti, S. (2005). Consumer testing with children on food
911 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-](https://doi.org/10.1016/0195-6663(92)90240-7)
 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](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](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., Peparaiio, 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–

- 957 31. <https://doi.org/10.1016/j.cofs.2018.05.005>
- 958 Stroebele, N., & de Castro, J. M. (2006). Listening to music while eating is related to increases in
959 people's food intake and meal duration. *Appetite*, *47*(3), 285–289.
960 <https://doi.org/10.1016/j.appet.2006.04.001>
- 961 Stroebele, N., & De Castro, J. M. (2004). Effect of ambience on food intake and food choice.
962 *Nutrition*, *20*(9), 821–838. <https://doi.org/10.1016/j.nut.2004.05.012>
- 963 Tuorila, H., Palmujoki, I., Kytö, E., Törnwall, O., & Vehkalahti, K. (2015). School meal acceptance
964 depends on the dish, student, and context. *Food Quality and Preference*, *46*, 126–136.
965 <https://doi.org/10.1016/j.foodqual.2015.07.013>
- 966 Trochim, W.M. (2006). Introduction to Validity. <http://www.socialresearchmethods.net/kb/introval.php>
967 (24th September 2018).
- 968 Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice : A Reference-Dependent
969 Model Author (s): Amos Tversky and Daniel Kahneman Published by : Oxford University
970 Press. *The Quarterly Journal of Economics*, *106*(4), 1039–1061. <https://doi.org/10.2307/2937956>
- 971 Ung, C.-Y., Menozzi, M., Hartmann, C., & Siegrist, M. (2018). Innovations in consumer research: The
972 virtual food buffet. *Food Quality and Preference*, *63*, 12–17.
973 <https://doi.org/10.1016/J.FOODQUAL.2017.07.007>
- 974 van Herpen, E., van den Broek, E., van Trijp, H. C. M., & Yu, T. (2016). Can a virtual supermarket
975 bring realism into the lab? Comparing shopping behavior using virtual and pictorial store
976 representations to behavior in a physical store. *Appetite*, *107*, 196–207.
977 <https://doi.org/10.1016/J.APPET.2016.07.033>
- 978 Weber, A. J., King, S. C., & Meiselman, H. L. (2004). Effects of social interaction, physical
979 environment and food choice freedom on consumption in a meal-testing environment. *Appetite*,
980 *42*(1), 115–118. <https://doi.org/10.1016/j.appet.2003.10.001>
- 981 Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments - A Presence
982 Questionnaire. *Presence: Teleoperators and Virtual Environments*, *3*(??), 225–240.
983 <https://doi.org/https://doi.org/10.1162/105474698565686>
- 984 Zeinstra, G. G., Koelen, M. A., Kok, F. J., & de Graaf, C. (2010). The influence of preparation method
985 on children's liking for vegetables. *Food Quality and Preference*, *21*(8), 906–914.
986 <https://doi.org/10.1016/j.foodqual.2009.12.006>
- 987