



HAL
open science

Temporal Dominance of Sensations paired with dynamic wanting in an ad libitum setting: A new method of sensory evaluation with consumers for a better understanding of beer drinkability

Takahiro Wakihira, Seiko Miyashita, Minoru Kobayashi, Kazuhiko Uemura,
Pascal Schlich

► To cite this version:

Takahiro Wakihira, Seiko Miyashita, Minoru Kobayashi, Kazuhiko Uemura, Pascal Schlich. Temporal Dominance of Sensations paired with dynamic wanting in an ad libitum setting: A new method of sensory evaluation with consumers for a better understanding of beer drinkability. Food Quality and Preference, 2020, 86, pp.103992. 10.1016/j.foodqual.2020.103992 . hal-02955299

HAL Id: hal-02955299

<https://hal.inrae.fr/hal-02955299>

Submitted on 15 Jul 2022

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 **[Title]**

2 Temporal Dominance of Sensations paired with dynamic wanting in an *ad libitum* setting: a
3 new method of sensory evaluation with consumers for a better understanding of beer
4 drinkability

5

6 **[Authors]**

7 Takahiro Wakihira^a, Seiko Miyashita^a, Minoru Kobayashi^a, Kazuhiko Uemura^a, Pascal
8 Schlich^{b,*}

9

10 ^a *Research Laboratories for Alcohol Beverages, Asahi Breweries, Ltd., 1-21, Midori*
11 *1-Chome, Moriya, Ibaraki 302-0106, Japan*

12 ^b *Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, CNRS, INRAE, Univ.*
13 *Bourgogne Franche-Comté, F-21000 Dijon, France*

14

15 *Corresponding author at: Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon,
16 CNRS, INRAE, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France

17 *E-mail address:* pascal.schlich@inrae.fr (P. Schlich)

18

19 **[Abstract]**

20 While drinkability is of paramount importance when discussing beer, there is no
21 established methodology to assess it. The main objective of the present study was to
22 develop a new method of sensory evaluation with consumers to obtain a better
23 understanding of beer drinkability. With the aim of being practically and effectively used in a
24 wide range of consumer tests in an *ad libitum* consumption setting, a new method called the
25 “Multiple-sip Drinkability Test” was developed to evaluate beer drinkability defined as “the
26 will of drinking”. The method is based on Temporal Dominance of Sensations (TDS) paired
27 with dynamic wanting.

28 This paper presents the test designs and the results from two studies of four commercial
29 Japanese beers, one conducted with an expert sensory panel and the other with a naïve
30 panel of consumers. The results revealed the importance of monitoring dynamic wanting
31 over sips. In both studies, product differences in *wanting*, almost nonexistent at the
32 beginning, gradually became larger.

33 These studies also elicited the characteristics of beer with high drinkability. The product
34 that attained the highest *wanting* scores in both studies was perceived as having less of a
35 standout flavor, thereby producing fewer build-up effects on sensory perceptions, which
36 suggests that the greater the sensory load produced by a beer, the less one wants to drink it

37 continuously. This methodology should be used with different types of beers and consumers
38 to obtain a broader understanding of the sensory drivers of beer drinkability and consumer
39 satisfaction.

40

41 **[Keywords]**

42 Temporal Dominance of Sensations (TDS), wanting, *ad libitum*, drinkability, Multiple-sip
43 Drinkability Test, beer

44

45 **[1] Introduction**

46 *1.1. Motivation for the research*

47 It is an indubitable fact that a manufacturer's mission is to keep delivering benefits to
48 consumers, which can be accomplished by examining the usage experience of their
49 products. Sensory and consumer science is likely one of the best means of bridging the gap
50 between a manufacturer and consumers caused by misunderstanding the consumers'
51 expectations. The authors have struggled for a long time with an issue in consumer
52 research in the beer industry; namely, why products that are liked/disliked in consumer
53 research are not necessarily the ones that sell well/poorly in the market, and vice versa. The
54 authors believe that the answer to this question does not simply lie in marketing investments
55 in products, and thus began this research to search for methodologies to better understand
56 consumer responses in a way that is relevant to consumption in real life as well as efficiently
57 applicable to product development in the beer industry.

58 Liking has been used in the beer industry as a key measurement for a very long time.
59 Liking is an easy construct to work with because the construct itself is perfectly clear and the
60 relationship between the word label (like/liking) and the construct is unambiguous to both
61 consumers and researchers (Thomson & Bailey, 2006). However, two beers which are liked
62 equally at first are not always equally liked at the end of drinking. This aspect has long been
63 discussed in the beer industry as a construct of "drinkability." In past years, the industry used
64 consumer tests called the "Session Test" or the "Sessionability Test" in which participants
65 were asked to drink large quantities of beer in one session and then researchers measured
66 the total volume of consumption for each beer (Greenhoff & Buck, 2006; Dickie et al. 2006).
67 These methods, however, include challenging issues related to ethics, health, feasibility, test
68 controls, and so forth. Thus, the beer industry has long desired the development of a new
69 scientific method to measure beer drinkability in consumer research.

70

71 *1.2. Previous studies on beer drinkability*

72 Several physiological studies on beer drinkability have been conducted in the last two

73 decades:

- 74 – Nagao et al. (1998) showed a correlation between beer drinkability and gastric emptying,
75 and suggested that measurement of the relaxed cross-sectional area of the pylorus
76 antrum was useful to evaluate stomach fullness during beer drinking.
- 77 – They also studied the sensory perceptions of stale beer and beer with an added
78 unpleasant taste and flavor in correlation with the urination rate, and found that the
79 unpleasant taste and off-flavor of beer resulted in a lower urination rate while drinking
80 beer (Nagao et al. 1999).
- 81 – Kojima et al. (2009) used a noninvasive biometric system to examine the relationship
82 among throat sensation, beer flavor, and swallowing motion while drinking beer. They
83 concluded that the sensation in the throat was an important factor in beer drinkability.

84 These types of physiological studies have disentangled some of the aspects of beer
85 drinkability with their objective measurements. Presently, though, there is a lack of
86 consumer-relevant subjective approaches that would add more value in understanding the
87 construct of drinkability by linking the results with those from physiological studies. While
88 bearing in mind that the ultimate goal of studying beer drinkability is to maximize consumers'
89 enjoyment of drinking beer, it is of paramount importance that we understand beer
90 drinkability from the consumers' perspective. Thus, there is a strong need to develop a
91 consumer-relevant approach for assessing beer drinkability.

92 According to Mattos and Moretti (2005, 2006), drinkability can be defined as “the will of
93 drinking or the weakness of the sensation of satiation” and “to be appreciated until the last
94 drop.” Several researchers have studied the impact of ingredients (malts, hops) and brewing
95 technology (yeast, fermentation) on sensory perceptions linked with beer drinkability
96 (Davies, 2006; Kaltner & Mitter, 2006; V. Opstaele et al. 2006; Taldi, 2006). Despite their
97 efforts to clarify factors affecting beer drinkability, it should be noted that each of these
98 researchers needed to create their own measurement as a response variable of drinkability.
99 Unfortunately, the absence of commonly accepted methods for measuring drinkability has
100 hampered efforts to clarify the factors affecting it. This is the impetus behind the present
101 study aimed at developing a new method to measure “the will of drinking” with a focus on the
102 sensory effect for beer drinkability, which could be widely used in consumer tests.

103

104 *1.3. Consumer research in food and drink categories*

105 The most popular type of consumer test with foods involves product trials at a central
106 location. These kinds of tests offer reasonably good control conditions, and the staff can be
107 well trained in product preparation and handling (Lawless & Heymann, 1998). Although the
108 product is not tested under its normal conditions of use, such as at home or at parties,

109 restaurants, etc., central location tests (CLTs) have the advantage that respondents
110 evaluate the product under controlled conditions, which enables any misunderstandings to
111 be cleared up (Meilgaard et al. 2016). With these points in mind, a consumer-relevant
112 method for assessing beer drinkability that can be widely applicable in CLTs would be a
113 valuable solution toward not only obtaining a greater scientific understanding of beer
114 drinkability, but also for product development, particularly during the early stage when many
115 prototypes need to be tested.

116 Being aware of the issue regarding CLTs, several researchers have studied the influence
117 on product acceptance of how food and drink are consumed by respondents in CLTs.
118 Matuszewska et al. (1997) compared three procedures for consumer assessment of fat
119 spreads, and concluded that having consumers spread a margarine sample *ad libitum* by
120 themselves on a standardized piece of bread was the most preferable procedure due to its
121 superior discrimination power and labor-saving sample preparation. Zandstra et al. (1999)
122 showed in their study on yogurts that the optimal sucrose concentration as determined by
123 the taste-and-spit test was higher than that determined from the *ad libitum* consumption test.
124 Posri and MacFie (2008) compared overall liking scores of tea bag products obtained from
125 three CLTs in which the degree of freedom when making tea was different, and concluded
126 that the provision of greater “freedom” in CLTs could result in greater discrimination between
127 tested products. Although the influence of testing conditions on product acceptance could
128 differ depending on the type of product, these previous studies indicated that improving the
129 design of CLTs based on consumer insight could potentially result in a better understanding
130 of estimating consumer response in real life.

131

132 *1.4. Beer studies from the perspective of sensory and hedonic perceptions*

133 Since the ultimate goal of studying beer drinkability is to maximize consumers' enjoyment
134 of drinking beer, it is crucial that we investigate the factors affecting beer drinkability in terms
135 of consumers' sensory perceptions as well as affection. Beer drinking behavior involves
136 dynamic rather than static sensations, which implies the necessity of incorporating temporal
137 aspects into evaluations to grasp changes in sensory perceptions over time while drinking.
138 Pineau et al. (2003) proposed a new method of sensory evaluation, called Temporal
139 Dominance of Sensations (TDS), in order to record several sensory attributes
140 simultaneously over time. According to them, TDS makes it possible to collect temporal data
141 during one single evaluation for up to 10 attributes of complex food products, and to draw
142 curves of the dominance of each attribute over time for each product (Pineau et al. 2009).
143 The same group of researchers also proposed an extended method of TDS, called
144 multiple-intake (sip or bite) TDS, in order to avoid potentially misleading results from a single

145 intake because, as Koster (2003) pointed out, the evaluation on a single intake only would
146 not capture the changes in sensory perceptions that occur during actual consumption. This
147 method made it possible to obtain a series of consecutive TDS evaluations at every intake,
148 which was more relevant to consumption of a product in real life (Vandeputte et al. 2011;
149 Pineau et al. 2013; Schlich et al. 2013). As an example, Zorn et al. (2014) applied
150 multiple-sip TDS methodology to evaluate different sweeteners in orange juice, and showed
151 its effectiveness in identifying differences in the dynamics of their sensory characteristics,
152 which had not been identified using static measurements. Furthermore, Thomas et al.
153 (2015) suggested a new methodology for characterizing temporal drivers of liking (TDL)
154 based on the ability of consumers to record their changes in liking and to perform a TDS
155 task.

156 Several researchers have recently reported on their studies of beer using temporal
157 aspects. Vazquez-Araujo et al. (2013) compared three different techniques of temporal
158 sensory methods for evaluating beer flavors: time intensity, TDS, and drinking profile. They
159 demonstrated that all three techniques are useful for evaluating beer flavor, and indicated
160 the pros and cons for each method. Simioni et al. (2018) reported a study on special beers
161 (Bohemian Pilsner, Witbier, Belgian Strong Ale Dubbel, and Russian Imperial Stout) in which
162 consumers took six sips (20 mL each) in total to evaluate a sample of the four different beers
163 on four consecutive days. The tasters performed TDS, followed by a liking evaluation, for
164 every sip of each beer sample until they finished taking the 6th sip. The authors concluded
165 that multiple-sip TDS associated with an acceptance test promoted an enhanced
166 understanding of consumer behavior regarding temporal description and sensory
167 acceptability over repeated consumption. Ramsey et al. (2018) applied a combination of
168 temporal liking (TL) and temporal Check-All-That-Apply (TCATA), with the aim of
169 determining the influence of ethanol concentration on consumer evaluation for lager beer.
170 Participants in their study drank 30 mL of a beer sample and continued the evaluation up to
171 60 seconds for TL or for TCATA in a separate session, and they repeated the same
172 evaluation method for four different beers during the same session. The researchers
173 demonstrated the effectiveness of using these combined temporal methodologies for
174 clarifying the influence of ethanol on the perception of liking and sensory attributes by beer
175 consumers.

176 These previous studies expanded the range of approaches for sensory and consumer
177 science in evaluating beer flavors. However, these studies were conducted using controlled
178 settings, such as the amount of beer in each sip or the number of sips per panelist. With the
179 caveat presented by Koster (2003) in mind, these approaches with limited and controlled
180 consumption might not provide sufficient information to understand consumers' responses

181 when drinking an entire beer. To the best of our knowledge, except for one study reported by
182 Thomas et al. (2016), no other studies have investigated the dynamics of sensory and
183 hedonic perceptions throughout the consumption of a whole package of a food or a full
184 portion of a drink in an *ad libitum* consumption setting. Thomas et al. (2016) showed the
185 effectiveness of using a method consisting of collecting TDS and dynamic liking data in the
186 same session and during the consumption of the full portion of an oral nutritional
187 supplement. However, no study on beer has been conducted from this perspective.
188 Furthermore, when discussing topics such as beer drinkability or sense of satiation
189 regarding food, collecting data on dynamic wanting rather than liking, as a measurement of
190 evoked affection, would make more sense in a correlation with the volume of beer or the
191 amount of food actually consumed.

192 193 *1.5. Scope of the present study*

194 Finally, two other topics should be mentioned when talking about beer drinkability:
195 “context of consumption” and “emotions.” Recently, Jaeger and Porcherot (2017) pointed
196 out in their review paper that “contextual influences partially shape food-related consumer
197 behavior,” and Nijman et al. (2019) elucidated the effects of consumption context towards
198 consumer responses to beer. These studies indicate the importance of considering
199 consumption context when assessing beer drinkability. Moreover, several studies have
200 recently reported the effectiveness of measuring emotions evoked when drinking beer with
201 the aim at differentiating products beyond liking (Cardello et al. 2016; Gomez-Corona et al.
202 2017; Jaeger et al. 2017; Jaeger et al. 2018; Silva et al. 2019). These studies indicate the
203 potential benefit of measuring emotions to better understand the construct of beer
204 drinkability for consumers. These two topics, “context of consumption” and “emotions,” were
205 not directly covered in the scope of the present study since its focus was to establish a basic
206 method to evaluate beer drinkability from a sensory perspective.

207 208 *1.6. Aim of the research*

209 The main objective of the present study was to develop a new method of sensory
210 evaluation with consumers to obtain a better understanding of beer drinkability. The aim was
211 to establish a new method to measure “the will of drinking”, which can be practically and
212 effectively used in a wide range of consumer tests in *ad libitum* consumption settings. The
213 ultimate goal of this study was to clarify the factors that affect beer drinkability, which could
214 be achieved by investigating the effects of manipulating explanatory variables such as
215 ingredients, recipes and brewing technologies. In this way, this research could contribute to
216 the development and marketing of products for maximizing consumers’ enjoyment of

217 drinking beer.

218

219 **[2] Methods**

220 *2.1. Overview of the approach in this study*

221 The investigation was divided into two different parts. In Phase 1, a test was conducted
222 with an expert beer panel to confirm the capability and feasibility of our methodology. In
223 Phase 2, a test was conducted with a consumer panel to confirm the validity of our
224 methodology with beer consumers. This type of approach was used to acquire more
225 implications and more aspects for discussion rather than jumping immediately into
226 consumer tests.

227

228 *2.1.1. Test design of Phase 1*

229 In Phase 1, a test was conducted with an expert beer panel consisting of 11 employees of
230 the R&D center at Asahi Breweries, Ltd. (Moriya, Japan) who were trained in tasting and
231 evaluating beer. These panelists were recruited from approximately 40 panelists of the
232 Asahi expert beer sensory panel who have from 6 months to nearly 20 years of experience
233 in beer evaluations. These panelists evaluate beer samples regularly once or twice a day to
234 contribute to beer sensory evaluations conducted in the R&D center. An annual training
235 session and the daily participation in beer sensory evaluations help them maintain their
236 ability to evaluate beer characteristics, which, in Phase 1, were based on consensus using
237 the definition of the following sensory attributes: *malty, hoppy, estery, sweet, sour, bitter,*
238 *astringent, and stimulating feeling.*

239 In Phase 1, the expert beer panel evaluated 4 products, with 2 repetitions each, for a total
240 of 8 evaluations per panelist. Panelists in Phase 1 drank a maximum of 500 mL of each of
241 the beer samples in 45 minutes. *(See 2.1.3. Rationale behind the test designs on maximum*
242 *consumption volume.)*

243

244 *2.1.2. Test design of Phase 2*

245 In Phase 2, a test was conducted with a naïve panel consisting of 44 participants (30 men
246 and 14 women, age range: 20s to 50s) who were recruited from the Asahi Group R&D
247 Center (Moriya, Japan). These panelists were not trained for beer evaluation, but were used
248 to drinking beer at least once a week. A drinking frequency of at least once a week and a
249 male-to-female ratio of 2:1 among the study group are standard recruitment criteria when
250 conducting consumer research in the beer industry in Japan.

251 In Phase 2, the naïve panel evaluated 3 products, with no repetition, for a total of 3
252 evaluations per panelist. Panelists in Phase 2 drank a maximum of 350 mL of each sample

253 in 30 minutes. (See 2.1.3. Rationale behind the test designs on maximum consumption
254 volume.)

255

256 2.1.3. Rationale behind the test designs on maximum consumption volume

257 In Japan, around 70% of beer sales are attributed to beers consumed at home (Fuji Keizai
258 Co., Ltd. 2019). Most of these beers for consumption at home are sold in the regular 350 mL
259 can or the larger 500 mL can. Since it was expected that monitoring evaluations for a longer
260 time would result in more discrimination among products, the 500 mL can was chosen for
261 the full portion drinkability test in Phase 1. The volume of total consumption was restricted to
262 500 mL to be in line with the guideline on responsible drinking (around 20 g pure alcohol per
263 day) established by the Ministry of Health, Labour and Welfare in Japan (*Kenkou Nippon*
264 21).

265 According to a survey conducted by Asahi Breweries in 2018, among beer drinkers (2000
266 men and 1000 women, age range: 20s to 50s) with a drinking frequency of at least once a
267 week, 71% of men do not consume more than 700 mL of beer (27% stop after drinking one
268 can of 350 mL and 44% stop after drinking one can of 500 mL or two cans of 350 mL) and
269 92% of women do not consume more than 700 mL (56% stop after drinking one can of 350
270 mL and 36% stop after drinking one can of 500 mL or two cans of 350 mL) in one session at
271 home. Thus, the investigation with beer consumers in Phase 2 focused on the size of a full
272 portion of a regular 350 mL can, which is at least half of the total consumption volume for a
273 majority (71% for men and 92% for women) of beer consumers in the Japanese market.
274 Another reason for choosing the 350 mL can was its potentially high acceptability and
275 feasibility when considering extensive use of our proposed methodology in consumer tests.
276 If evaluations of a full portion of 350 mL lead to sufficient product discrimination, it would be
277 considerably useful for the beer industry in product development, particularly at the early
278 stage when many prototypes need to be tested.

279

280 2.2. Test flow

281 Figure 1 shows the flow of the “Multiple-sip Drinkability Test.” First, the panelists take a sip
282 and participate in TDS for 30 seconds, followed by a Liking evaluation. Second, they wait
283 until they feel like drinking another sip. Just before they want to have the 2nd sip, they
284 evaluate their *residual sensation*, which is the degree of flavor perceptions left in the mouth
285 and nose. Next, they take the 2nd sip and evaluate its *ease of drinking*. Again, they wait until
286 they feel like drinking another sip. Just before they want to have the 3rd sip, they evaluate
287 how much they “Want to drink continuously.” For the 4th, 5th, and 6th sips, they perform the
288 same evaluation methods as those for the 1st, 2nd, and 3rd sips, respectively. Likewise, they

289 repeat the same evaluations every 3 sips. (Note: *Residual sensation* was only asked in
290 Phase 1. See section 2.7 for further details.)

291 Most importantly, both of the tests in Phase 1 and Phase 2 were conducted in an *ad*
292 *libitum* setting in terms of the drinking amount and pace. Neither the amount of beer that
293 should be drunk in each sip nor the timing of the sips was set in the tests. Panelists could
294 also stop drinking whenever they wanted.

295

296 2.3. Test products

297 Four commercial beer samples marketed in Japan were ~~used~~ selected for this study.
298 Detailed product information is not mentioned in the present paper to avoid identifying
299 specific brands. ~~Instead, these samples were labeled as B1, B2, B3 and B4.~~ The main
300 analytical and sensory characteristics of the samples are shown in Table 1. Samples B1, B3
301 and B4 were selected ~~as test samples in this study~~ because they were major brands sold in
302 the market and because each possessed different sensory characteristics according to
303 previous sensory tests evaluated by the Asahi expert beer sensory panel. Sample B2,
304 however, was known to have quite similar characteristics to those of sample B4. Only a
305 slight difference was previously reported between samples B4 and B2; namely, sample B2
306 was perceived as more hoppy and less malty than B4 in a very subtle way. The reason why
307 we included Sample B2 was to see if our new methodology was able to detect any
308 difference between it and B4 in the *ad libitum* setting of drinking. However, no significant
309 difference was observed between samples B4 and B2 regarding sensory and hedonic
310 perceptions in Phase 1 with the expert panel. Therefore, we decided to eliminate sample B2
311 in Phase 2 with consumers in order to simplify the test design and to reduce the
312 respondents' burden while participating in the test.

313 All the test samples were served in plastic cups at a temperature of 8 (\pm 1) °C, which is a
314 common temperature for evaluating beers in sensory studies in Japan. Each product was
315 presented with its own unbranded code using a three-digit random number, according to
316 Williams Latin Square experimental designs.

317

318 2.4. Test environments

319 Each panelist evaluated one beer sample per day in an individual sensory booth in the
320 R&D center at Asahi Breweries, Ltd. During the evaluation session, the panelists were not
321 allowed to drink water to avoid disrupting evaluations of the beer samples. They were only
322 allowed to eat ~~non-topping salt crackers~~ plain crackers with no salt topping while they were
323 drinking the beer samples. All the panelists, including the expert panelists in Phase 1, were
324 instructed to drink the beer samples freely in the way they usually consumed beer. Use of

325 phones or personal computers (PCs) for work was prohibited during the sessions, but
326 magazines were available in the sensory booths for the participants to read. Following the
327 instructions displayed on the PC screens in the isolated sensory booth, each panelist
328 performed sensory and hedonic evaluations throughout the session, and the data were
329 captured using the TimeSens© software (INRA, Dijon, France). After the panelists left the
330 sensory room, the volume of beer actually consumed was calculated by weighing the
331 remaining beer in the cups.

332

333 2.5. TDS evaluation

334 Each panelist took the 1st sip of a beer sample and participated in TDS for 30 seconds.
335 Likewise, they repeated TDS every 3 sips, for the 4th, 7th, 10th sips, and so on. In line with
336 Pineau et al. (2012) recommending not to include more than 10 attributes in a TDS
337 evaluation, the TDS list in Phase 1 included eight attributes: “*malty*” (*roasted, caramel,*
338 *grainy, worty*), “*hoppy*” (*floral, grassy, resinous, spicy, citrus*), “*estery*” (*fruity, banana, apple,*
339 *soapy, solvent-like*), “*sweet taste*” (*primary taste*), “*sour taste*” (*primary taste*),
340 “*bitter/astringent*” (*bitter as primary taste / astringent as the feeling of mouth puckering,*
341 *tannin-like*), “*stimulating feeling*” (*the feeling of pricking and stinging in the mouth and throat*),
342 and “*others*.” These attributes were selected to sufficiently describe the sensory
343 characteristics of the four beer samples based on previously obtained information from the
344 Asahi expert beer sensory panel (Table 1). The expert panelists were all familiar with the
345 definition of these attributes, and they were used to performing TDS with this list. *Others*
346 was included as a choice in the list in case the panelists perceived some flavors that were
347 not applicable to the seven specific attributes. In Phase 2, the naïve panelists received 5
348 minutes of instruction on the procedure of the TDS evaluation. The TDS list in Phase 2
349 included eight attributes: “*malt-like flavor*,” “*sweet flavor*,” “*refreshing feel*,” “*sweet taste*,”
350 “*sour taste*,” “*bitter/astringent*,” “*stimulating feel*,” and “*others*.” These attributes had been
351 commonly used in consumer tests with beer drinkers in Japan. The definition of each
352 attribute was not explained to the naïve panelists as with normal consumer tests in Japan.
353 The *others* category was also included in the attribute list in Phase 2.

354 During the test, the selected TDS attributes were presented on one screen, and only
355 dominance was recorded (intensity was not recorded). The attribute order on the screen
356 was different for each panelist, but each panelist kept his or her own order for all sips and
357 products to avoid attribute order effects while facilitating the task of the tasters (Pineau et al.
358 2012, Thomas et al. 2016). They had to choose at any tasting time the dominant attribute
359 among the eight proposed. The concept of dominant attribute was defined as “the attribute
360 associated with the sensation catching the taster’s attention at a given time.” Panelists were

361 free to choose the same attribute several times during a tasting, or never select a sensation
362 among the eight proposed (Pineau et al. 2009, Thomas et al. 2016). They were also able to
363 unclick the attribute buttons in case they did not feel any sensations during the 30-second
364 period, and those moments were regarded as *No dominance* when analyzing the results.

365

366 *2.6. Dynamic evaluation of wanting*

367 Just before the panelists wanted to have the 3rd sip, they evaluated how much they “want
368 to drink it continuously” on a linear rating scale ranging from “I do not want to drink it
369 continuously at all” to “I want to drink it continuously very much.” The panelist’s indication on
370 the linear rating scale was converted in the TimeSens© software to a score ranging from 0
371 to 10 for data analysis purposes (the same procedure was used for the measurements
372 described in section 2.7.). They repeated this *wanting* evaluation every 3 sips, just before
373 the 6th, 9th, 12th sips, and so on. Since the main focus of this study was to measure “the will
374 of drinking” for beer drinkability, the greatest importance was placed on this dynamic
375 evaluation of “wanting to drink continuously.”

376

377 *2.7. Dynamic evaluation of liking, ease of drinking and residual sensation*

378 Immediately after the panelists completed the TDS (1st, 4th, 7th sips, and so on, every 3
379 sips), they indicated their *liking* for the beer sample on a linear rating scale ranging from “I
380 do not like it at all” to “I like it very much.”

381 Immediately after the subjects took the 2nd, 5th, 8th sips, and so on, every 3 sips, they
382 evaluated its *ease of drinking* on a linear rating scale ranging from “not easy to drink at all” to
383 “very easy to drink.”

384 Subjects evaluated their *residual sensation* (from “not strong at all” to “very strong”) just
385 before they took the 2nd, 5th, 8th sips, and so on, every 3 sips only in Phase 1. This
386 measurement was not included in Phase 2 **with the naïve panel** in order to simplify the test
387 design **and to reduce potential fatigue caused by evaluating many attributes.** ~~with the naïve~~
388 ~~panel.~~

389

390 *2.8. Data analyses*

391 Data was analyzed using SAS 9.4 (SAS Institute Inc., Cary, NC, USA). TDS curves and
392 TDS band plots were generated using TimeSens 1.3 (INRA, Dijon, France).

393

394 *2.8.1. TDS analysis*

395 A TDS curve represents the evolution of the dominance rate (which is the proportion of
396 panelists who selected a certain attribute as dominant) over time. For each product, TDS

397 curves of all the tested attributes are depicted on the same graph. The TDS curves of the
398 differences in the dominance rate can be superimposed to compare two products. The
399 differences are plotted only when significantly different from zero and when the higher
400 dominant rate is significantly higher than the chance level, thereby highlighting the
401 differences between the products over time (Pineau et al. 2009).

402 A TDS band plot is generated from TDS data to simplify interpretation of the TDS curves,
403 and facilitates comparison of the sequence of the dominant attributes in all the products of
404 the study.

405

406 *2.8.2. Analysis of dynamic wanting*

407 The analysis of variance (ANOVA) model used for analyzing the *wanting* data is:

408

409 $Wanting = product + sip + panelist + product*sip + product*panelist + sip*panelist + error$

410

411 In this model, the product, sip and product*sip interaction were treated as fixed effects.
412 The panelist and its interactions with the other factors were random effects. As mentioned in
413 section 2.2., all the panelists drank the beer samples freely. This resulted in a different
414 number of sips for each panelist for each sample, which requires use of a restricted
415 maximum likelihood (REML) algorithm instead of the usual least squares algorithm.
416 Furthermore, successive evaluations of *wanting* were not independent from sip to sip by
417 product within each panelist. Therefore, a repeated mixed model ANOVA was used with a
418 first-order heterogeneous auto-regressive covariance structure among sips within panelist
419 by product. This structure assumes that the covariance between two *wanting* scores from
420 two sips of the same evaluation decreases with the number of sips in between these two
421 sips.

422 The least square mean (LS-mean) and standard error were subsequently computed for
423 each product, for each sip, and for each sip by each product, which enabled interpretation of
424 the statistical differences among products, among sips, and among sips by each product,
425 respectively. This analysis was accomplished using PROC MIXED in SAS 9.4.

426

427 *2.8.3. Analysis of dynamic liking, ease of drinking and residual sensation*

428 The same repeated mixed model ANOVA was used for other dynamic measurements,
429 which included *liking*, *ease of drinking* and *residual sensation*.

430

431 *2.8.4. Evolution of attribute dominance durations*

432 Duration of dominance is obtained at the individual level by adding all the time periods

433 during which the attribute was cited as dominant, regardless of the moment of perception
434 (Galmarini et al. 2017). Data for the dominance duration of each attribute were analyzed by
435 the same repeated mixed model ANOVA as those used for *wanting* and the other
436 measurements.

437

438 [3] Results

439 3.1. Phase 1 results with the expert panel

440

441 3.1.1. Dynamic evaluation of *wanting*, *liking*, *ease of drinking* and *residual sensation*

442 Table 2a and Table 2b show results of the ANOVA with the expert panel. Fixed effects of
443 both Sip and Product were observed for *wanting*, *liking* and *ease of drinking* at a statistically
444 significant level ($p < 0.05$) (Table 2a), but not for *residual sensation*. Sip effect here refers to
445 changes in evaluations over sips while drinking. The result also showed a significant
446 Product by Sip interaction for *wanting* ($p < 0.05$). The p-values of Sip effect by Product
447 (Table 2b) indicate that *wanting* for samples B1 and B3 significantly changed (decreased)
448 while drinking, whereas *wanting* for samples B2 and B4 did not. Interestingly, the product
449 differences in *wanting*, which were almost nonexistent at the beginning of Sip 3 (Sip Cycle 1),
450 gradually became larger while drinking (Figure 2). As a result, the LS-means of *wanting*
451 show a significant decrease in samples B1 ($p = 0.0000$) and B3 ($p = 0.0000$), but not in
452 samples B2 ($p = 0.2367$) and B4 ($p = 0.1576$) (Table 2b). In this way, we found that samples
453 B2 and B4 had higher beer drinkability than samples B1 and B3.

454 Figure 3 shows the evolution of *ease of drinking*. In contrast to the evolution of *wanting*
455 (Figure 2), the product differences were already existent at the beginning of Sip 2 (Sip Cycle
456 1). *In particular, this tendency was clearer when looking at the results of sample B1
457 compared with those of sample B4. Wanting for sample B1 was comparable to that for
458 sample B4 at Sip 3 (Sip Cycle 1); however, ease of drinking for sample B1 was clearly lower
459 than that for sample B4 at Sip 2 (Sip Cycle 1).* The products with lower *wanting* at the end
460 (samples B1 and B3) were lower in *ease of drinking* from the beginning. These results
461 indicate that *ease of drinking* at each sip could lead to “wanting to drink continuously,” that is,
462 beer drinkability.

463 The measurements of *wanting* and *liking* were similar in that both of them detected a
464 significant Sip effect ($p = 0.0000$ for *wanting*, $p = 0.0023$ for *liking*) and Product effect ($p =$
465 0.0037 for *wanting*, $p = 0.0155$ for *liking*) (Table 2a). Indeed, the correlation coefficient
466 between *wanting* and *liking* was high at any sip cycle until Sip Cycle 6, which was 0.79, 0.87,
467 0.82, 0.86, 0.82, 0.83, respectively computed over 2 replicates, 11 panelists and 4 products
468 ($n = 88$). However, as shown in Figure 4, in which the LS-mean scores over products in

469 Table 2b were plotted in order of sip, decreases in evaluation scores became more
470 prominent for *wanting* than for *liking* after the middle of the drinking session (Sip Cycle 4).
471 Also, product by sip interaction was only significant for *wanting* ($p = 0.0310$), but not for
472 *liking* ($p = 0.2672$) (Table 2a). These findings indicate that asking about *wanting* could be
473 more powerful in understanding beer drinkability as well as discriminating beer products,
474 rather than asking about *liking*.

475 Although we found that samples B2 and B4 were significantly higher than samples B1 and
476 B3 for *wanting and ease of drinking* ($p < 0.05$ for both), we could not detect any differences
477 between samples B2 and B4 or between samples B1 and B3 for these measurements
478 (Table 2a).

479

480 3.1.2. Dynamic evaluation (TDS) of flavor characteristics

481 Figure 5 shows the evolution of sensory perceptions. The TDS band plot is a summary of
482 the TDS results for the four beer samples over time while drinking. From the left side, the
483 TDS results are shown every 30 seconds in order of time for the 1st, 4th, 7th sip, and so on.
484 The colors indicate the sensory characteristics that were significantly dominant during each
485 second. For sample B1, “*Hoppy*” (blue) was perceived to be consistently dominant at the
486 beginning of each sip and “*Bitter/Astringent*” (red) became more dominant over time while
487 drinking. For sample B3, “*Bitter/Astringent*” was consistently dominant for most of the time
488 from beginning to the end. Samples B2 and B4 were perceived to be quite similar to each
489 other, with “*Malty*” (purple) and “*No Dominance*” (black) more dominant than the other two
490 products. As previously mentioned in section 2.5. *TDS evaluation*, the moment when a
491 panelist did not feel any sensations was regarded as “*No Dominance*” when analyzing the
492 results. In summary, when comparing the products’ characteristics, sample B1 was
493 characterized as the most *Hoppy*, sample B3 as the most *Bitter/Astringent*, and samples B2
494 and B4 had the longest period of “*No Dominance*”, which can be seen until the end of
495 drinking.

496

497 3.1.3. Link between dynamic wanting and sensory perceptions

498 The relationship between the results of dynamic *wanting* and TDS for the four tested
499 products reveals some indications for beer drinkability and sensory perceptions. The
500 products that scored higher in *wanting*, samples B2 and B4 (Table 2a and Figure 2), were
501 perceived as having less standout flavors such as “*Hoppy*” and “*Bitter/Astringent*” (Figure 5),
502 thereby producing less build-up effects on sensory perceptions. These results indicated that
503 the greater the sensory load produced by a beer, the less one wants to drink it continuously.

504

505 3.2. Phase 2 results with the naïve panel

506

507 3.2.1. Dynamic evaluation of wanting, liking and ease of drinking

508 Table 3a and Table 3b show the results of the ANOVA with the naïve panel. Fixed effects
509 of both Sip and Product were observed for *wanting*, *liking* and *ease of drinking* at a
510 statistically significant level ($p < 0.05$). Again, Sip effect means that the evaluation scores
511 change over sips while drinking. The p-value for product by sip interaction for *wanting* ($p =$
512 0.1767) was relatively lower, but not at a statistically significant level in the naïve panel.
513 Similar to the results from Phase 1 with the expert panel, the p-values of Sip effect by
514 Product indicate that *wanting* for samples B1 ($p = 0.0101$) and B3 ($p = 0.0001$) significantly
515 changed (decreased) while drinking, but not for sample B4 ($p = 0.1893$) (Table 3b). Notably,
516 the product differences in *wanting*, which were almost nonexistent at the beginning,
517 gradually became larger while drinking (Figure 6). This finding is consistent with the results
518 of Phase 1 with the expert panelists. The LS-mean of *wanting* for sample B4 was
519 significantly higher than that for sample B3 ($p < 0.0001$), and almost significantly ~~tended to~~
520 **be** higher than that for sample B1 ($p = 0.1007$) (Table 3a).

521 For the results with the naïve panel, *wanting* and *liking* were similar in that both of them
522 detected a Sip effect ($p < 0.0001$ for *wanting*, $p = 0.0040$ for *liking*) and Product effect ($p <$
523 0.0001 for both *wanting* and *liking*) and differences in product LS-means ($p < 0.05$) at
524 statistically significant levels (Table 3a). The correlation coefficients between *wanting* and
525 *liking* were high throughout all of the sip cycles, which were 0.77, 0.85, 0.87, 0.89, 0.94, 0.89
526 and 0.89, computed over panelist by product ($n = 126$). As shown in Figure 7, in which the
527 LS-mean scores over products in Table 3b are plotted in the order of sip, development of
528 decreases in these scores seem to be in alignment with each other, and no strong evidence
529 was observed in the naïve panel regarding whether asking about *wanting* was more
530 powerful than asking about *liking*. Compared with the expert panelists, the naïve panelists
531 tended to use a narrow range on the linear rating scale while giving higher *wanting* scores
532 overall, and they were less inclined to give lower scores during the entire experiment.

533

534 3.2.2. Dynamic evaluation (TDS) of flavor characteristics

535 Figure 8a shows TDS difference curves between samples B4 and B3, which were
536 respectively the highest and the lowest in *wanting* scores (Table 3a). The sensory attributes
537 over the centerline indicate a significantly higher ($p < 0.1$) dominance rate for sample B4,
538 and those below the centerline indicate a significantly higher ($p < 0.1$) dominance rate for
539 sample B3 over the 30-second period, within each sip. (Note: In Figure 8a and Figure 8b,
540 “within sip” analyses were used to focus on direct comparison of sensory perceptions

541 between two products.) Sample B4 scored significantly higher for “*Refreshing Feel*” in the
542 beginning and for “*No Dominance*” at the end. In contrast, sample B3 scored significantly
543 higher for “*Bitter/Astringent*” from the beginning to the end.

544 Figure 8b shows TDS difference curves between samples B4 and B1, which were
545 respectively the highest and the second highest in *wanting* scores (Table 3a). More
546 dominance could be seen for sample B4 in “*Stimulating Feeling*” and “*Malt-like Flavor*.”
547 However, the difference between samples B4 and B1 was not as clear as that observed by
548 the expert panel, who perceived sample B1 as being more “*Hoppy*.”

549 To grasp a better understanding of the panelists’ responses, we conducted a
550 segmentation analysis for samples B4 and B1 (Figure 9). The left graph shows TDS
551 difference curves among the panelists, who gave a higher *wanting* score for sample B4.
552 Conversely, the right graph shows curves among those that gave a higher *wanting* score for
553 sample B1. Interestingly, perceptual differences were found between these two groups.
554 The panelists who scored sample B4 higher in *wanting* perceived it as less dominant for
555 “*Bitter/Astringent*,” whereas it was the opposite for those who scored sample B1 higher for
556 *wanting*. Because of the small number of panelists ($n = 42$) in this study, these findings from
557 the segmentation analysis should be examined with caution and need to be validated with a
558 larger number of consumers.

559

560 3.2.3. Development of bitter/astringent dominance duration

561 Figure 10 shows the development of dominance duration for “*Bitter/Astringent*” for all
562 three products. Dominance duration refers to the total number of seconds that
563 “*Bitter/Astringent*” was selected as dominant in each sip. Sample B3 was significantly higher
564 ($p < 0.01$) for the dominance duration of “*Bitter/Astringent*” compared to samples B1 and B4,
565 and B3 showed a tendency to accumulate “*Bitter/Astringent*” over time during the drinking
566 session. Considering the fact that sample B3 was lower in *wanting* compared to samples B1
567 and B4, it can be assumed that the build-up of “*Bitter/Astringent*” could lead to low
568 drinkability.

569

570 3.3. Wanting and volume of beer actually consumed

571 Figure 11a shows the volume of beer actually consumed by the expert panelists in Phase
572 1. Samples B2 and B4 tended to be consumed more than samples B1 and B3, but not at a
573 statistically significant level because of the large variance observed among the panelists. By
574 contrast, a significant difference in *wanting* scores was observed between the products ($p <$
575 **0.05**) (Table 2a). Therefore, in this case, asking about *wanting* was more useful for
576 discriminating products than collecting data on the total volume consumed.

577 Similar results were found in Phase 2 with the naïve panelists (Figure 11b). No significant
578 differences in consumed volume of beer were observed, while the trend in consumed
579 volume was in line with the *wanting* scores (Table 3a). However, significant differences ($p <$
580 0.05) were found in *wanting* scores between products (Table 3a), with an indication that
581 asking about *wanting* outperformed measuring the total volume of consumption in terms of
582 discriminating products.

583

584 [4] Discussion

585 The present work aimed at developing a new method of sensory evaluation for beer
586 drinkability, which can be widely and easily applied in consumer tests. On the construct of
587 beer drinkability defined as “the will of drinking,” a new method called the “Multiple-sip
588 Drinkability Test” was developed based on TDS paired with dynamic wanting to drink
589 continuously. In the present study, with no lengthy training, both the expert panelists and the
590 naïve panelists were able to conduct TDS repeatedly while drinking a full portion of beer up
591 to 500 mL or 350 mL, respectively. The panelists were able to understand TDS and
592 implement it properly after receiving only 5 minutes of instruction. This is in line with the
593 suggestions of Monaco et al. (2014) and Schlich (2017). Furthermore, the participants in this
594 study only needed to follow the instructions displayed on PC screens to perform their
595 evaluations readily and intuitively on a web site specifically allocated for each respondent.
596 These aspects of the present study will contribute as part of a widely applicable approach to
597 sensory and consumer research for beer drinkability in CLTs.

598 To develop new, successful products, companies should gain a deep understanding of
599 “the voice of the consumer” in the early stage of product development; however, this step is
600 often ignored or poorly executed (van Kleef et al. 2005). Koster (2003) raised issues
601 concerning researchers’ bias of subject consistency and possibly misleading outputs from
602 consumer studies that rely only on single hedonic evaluation. The results from the present
603 study were in agreement with Koster’s statement. By measuring dynamic *wanting* while
604 drinking a full portion of beer, product differences in *wanting*, which were almost nonexistent
605 at the beginning, were found to gradually become larger while drinking. Investigations were
606 also conducted on the panelists’ responses for both *wanting* and *liking* asked dynamically
607 over sips to find a relationship between these two evaluation scales. Phase 1 results with
608 the expert beer panel showed a superiority of measuring *wanting* over *liking* in terms of
609 identifying product differences over time. In contrast, no large differences were observed
610 between these two evaluations in Phase 2 with the naïve panel, who used a relatively
611 narrow range on the linear rating scale. **We restricted the maximum volume of consumption
612 to 350 mL in Phase 2 with the naïve panel (see section 2.1.3. Rationale behind the test**

613 *designs on maximum consumption volume*); however, increasing the volume of maximum
614 consumption up to 500 mL might lead consumer panelists to use a broader range of the
615 linear rating scale for *wanting*, which could result in better discrimination of beers at the end
616 of drinking. Furthermore, when considering that some consumer groups, such as heavy
617 drinkers are likely to express their “will” of drinking beer continuously in a clearer attitude
618 than to show their appreciation toward a beer, *wanting* would probably be a more valuable
619 evaluation scale if a researcher needs to choose one of these two scales for studying beer
620 drinkability.

621 The test results also indicated that *ease of drinking* at each sip could lead to “wanting to
622 drink continuously,” in other words, beer drinkability. This finding is useful from the
623 perspective of product development, particularly during the early stage when many
624 prototypes need to be tested. By using this knowledge, asking about the *ease of drinking* for
625 just a portion of a beer sample, rather than the full volume every time, could work as a good
626 predictor of beer drinkability for the purpose of screening prototypes, allowing the
627 researcher to move on to further steps with a decreased number of candidates.

628 Since the focus in the present work was mostly methodological and not to investigate
629 details of beer flavors, we conducted the tests with a relatively small number of panelists.
630 However, in addition to the development of the methodology, which enabled the
631 discrimination of beers in terms of drinkability, the conducted study also elicited the flavor
632 characteristics of beer with high drinkability. The results indicated that a product with less
633 standout flavor, thereby producing fewer build-up effects on sensory perceptions, could be a
634 beer with higher drinkability. In other words, the greater the sensory load produced by a beer,
635 the less one wants to drink it continuously. These findings are in agreement with the results
636 of some previous beer studies (Guinard et al. 1998; Mattos & Moretti, 2005, 2006; Davies,
637 2006; Greenhoff & Buck, 2006; Kojima et al. 2009).

638 The ultimate goal for the investigation of beer drinkability was to maximize consumers’
639 enjoyment of drinking beer. Since the present evaluation included only four commercial
640 products marketed in Japan and a limited number of panelists, more tests with designed
641 prototypes sampled by general consumers on a larger scale are necessary in order to clarify
642 further the factors affecting beer drinkability. Since the effectiveness of using the
643 “Multiple-sip Drinkability Test” was validated with consumers, it will be possible to further
644 investigate the impacts of various factors such as ingredients, recipes and brewing
645 technologies by manipulating those variables when making prototypes for consumer tests.
646 This study, particularly in Phase 2 with consumers, also showed the possibilities of
647 perceptual differences and directional differences for highly drinkable beers depending on
648 consumer segments. Applying this method with a large number of target consumers would

649 be beneficial for marketing and product development in the beer industry.

650 Results from the tests with the expert and the naïve panels showed that asking about
651 *wanting* was more useful for discriminating products than collecting data on the total volume
652 consumed. One of the reasons that no significant difference in the total consumption volume
653 was observed among the products was because of the large variance observed among the
654 panelists. Another reason was because of the small variance in the total consumption
655 among the products for each panelist. Possibly, with a limited amount of consumption up to
656 500 mL or 350 mL, “the will of drinking” can be more readily captured by simply asking about
657 *wanting* at a conscious level rather than letting a panelist show it by drinking behavior at a
658 subconscious level. Since the maximum amount of a beer sample was the normal volume
659 consumed daily by a panelist, there might have been psychological barriers to stop drinking
660 a test sample unless its flavors strongly discouraged a panelist’s will of drinking continuously.
661 This may also suggest that there could be inertia in beer drinking behavior; that is, beer
662 could be a type of drink that is consumed continuously without a specific strong feeling of
663 *wanting* while drinking. However, the reason for this finding remains unclear and warrants
664 further investigation.

665 In the present study, the panelists were instructed to only eat ~~non-topping salt crackers~~
666 plain crackers with no salt topping while they were drinking the beer samples. However,
667 considering beer drinking contexts in real life, further investigations of beer drinkability “with
668 food” would add more value for product development and marketing purposes in the future.
669 King et al. (2004) reported that overall liking of a flavored iced tea significantly increased
670 when it was served together with a salad and a pizza compared with the overall liking of the
671 tea alone. Galmarini et al. (2016) conducted a study involving multi-intake TDS to evaluate
672 the influence of cheese on wine perception, and the results of their study showed that the
673 dominance duration of astringency and sourness for a red wine was reduced by cheese
674 consumption between sips of the wine. These previous investigations suggest an
675 effectiveness of taking into consideration the possibility of “meal context” affecting beer
676 drinkability. With the effectiveness of using the “Multiple-sip Drinkability Test” with
677 consumers in an *ad libitum* consumption setting demonstrated in the present study, it would
678 be worth expanding this protocol to search for the optimal pairing with foods in the construct
679 of beer drinkability for consumers.

680

681 [5] Limitations and further research

682 According to Mattos and Moretti (2005, 2006), there are four different effects of beer
683 drinkability: sensory, cognitive, post-ingestive and post-absorptive. Among these, our focus
684 was on the sensory aspect of drinkability. For measuring it at a conscious level, the

685 quantities of consumption in the present study were limited to 500 mL. Consequently, it
686 would not be possible to discuss the results from the aspects of post-ingestive or
687 post-absorptive, which would require participants to drink large quantities of beer in one
688 session. The consumer panelists in Phase 2 drank a maximum of 350 mL of each sample,
689 which is the volume of a full portion of a regular can available in Japan. Therefore, the
690 present method with consumers has a limitation that it could at most estimate the power to
691 evoke *wanting* for the second 350 mL can. As mentioned in section ~~“2.1.1. Test design of~~
692 ~~Phase 2,”~~ **2.1.3. Rationale behind the test designs on maximum consumption volume,**
693 however, the majority of beer consumers in Japan do not consume more than two cans of
694 beer in one session at home. This is why the authors believe that this method is a
695 consumer-relevant approach. **In terms of product discrimination, increasing the maximum**
696 **consumption volume up to 500 mL in future studies could result in better discrimination of**
697 **beers.**

698 In the present study, each panelist conducted their evaluation of beer individually in an
699 isolated sensory booth. They were allowed to read magazines as they wanted during a test
700 so that they could feel at ease; however, the test environment was not ideally designed to be
701 or to make them feel as if they were in a “real” consumption context. As recently reported
702 (Jaeger & Porcherot, 2017; Nijman et al. 2019), consumption context could be one of the
703 key factors affecting beer drinkability. The present method was useful in assessing beer
704 drinkability with consumers, so it would be valuable to evaluate beers in some different
705 consumption contexts or evoked contexts by manipulating test conditions in future studies. **It**
706 **should also be mentioned that the test samples were served in plastic cups, although**
707 **consumers usually drink beer in beer glasses or directly from cans in real life. This might**
708 **have an effect on perceived sensation, possibly caused by mouthfeel when taking a sip,**
709 **amount of beer taken in each sip, changes in beer temperature while drinking, and so forth.**
710 **Depending on the study objectives, this aspect should be taken into consideration when**
711 **designing tests for future research.**

712 As mentioned previously, the focus in this study was on the sensory aspect of beer
713 drinkability. Thus, no list was provided for the panelists to assess their emotions while
714 drinking beer samples. Given that the emotional aspect was proven to be effective in
715 differentiating beers beyond liking (Cardello et al. 2016; Gomez-Corona et al. 2017; Jaeger
716 et al. 2017; Jaeger et al. 2018; Silva et al. 2019), the authors are considering further studies
717 that incorporate measurements of emotions into the current protocol, such as a combination
718 of TDS and Temporal Dominance of Emotions (Jager et al. 2014) paired with dynamic
719 wanting to better understand the construct of beer drinkability for consumers.

720

721 **[6] Conclusion**

722 A new method called the “Multiple-sip Drinkability Test” was developed to evaluate beer
723 drinkability defined as “the will of drinking,” based on Temporal Dominance of Sensations
724 paired with dynamic wanting. To the best of our knowledge, this is the first published
725 methodology with consumers in an *ad libitum* setting, which could be widely and easily
726 applied in central location tests to obtain a better understanding of beer drinkability.

727 The present method enabled the discrimination of beers in terms of drinkability, and the
728 results revealed the importance of monitoring dynamic wanting over sips. It was found that
729 product differences in wanting, which were almost nonexistent at the beginning, gradually
730 became larger while drinking. The results of this study also indicated that a product with less
731 standout flavor, thereby producing fewer build-up effects on sensory perceptions, could be a
732 beer with higher drinkability. In other words, the greater the sensory load produced by a beer,
733 the less one wants to drink it continuously.

734 Since the present work was mostly methodological with a relatively small number of
735 panelists, more work will be required to clarify further the factors that affect beer drinkability
736 for target consumers on a larger scale. This could be achieved by testing designed
737 prototypes with this evaluation method to investigate the impacts of changing variables,
738 such as, ingredients, recipes and brewing technologies.

739

740 **[Acknowledgements]**

741 We thank all the volunteer panelists who participated in our tests.

742 This research was funded by Asahi Breweries, Ltd.

743 **We would also like to thank the reviewers who contributed to improve significantly the**
744 **quality of the manuscript.**

745

746

747

748

749

750

751

752

753

754

755

756

757 **[References]**

- 758 Cardello, AV., Pineau B., Paisley, AG., Roigard, CM., Chheang, SL., Guo, LF., Hedderley DI.,
759 & Jaeger, SR. (2016). Cognitive and emotional differentiators for beer: An exploratory
760 study focusing on “uniqueness.” *Food Quality and Preference*, *54*, 23–38.
- 761 Davies, N. (2006). Malt - a vital part of the brewer’s palette. *European Brewery Convention*,
762 *Monograph 34, Symposium, Drinkability, Edinburgh, United Kingdom, November 2006*.
- 763 Di Monaco, R., Su, C., Masi, P., & Cavella, S. (2014). Temporal dominance of sensations: A
764 review. *Trends in Food Science & Technology*, *38*, 104–112.
- 765 Dickie, KH., Parker, DK., & Muller, RE. (2006). Fancy another beer - same again? *European*
766 *Brewery Convention, Monograph 34, Symposium, Drinkability, Edinburgh, United*
767 *Kingdom, November 2006*.
- 768 Fuji Keizai Co., Ltd. (2019). *Foodstuff Marketing Handbook 2020: Vol. 2*.
- 769 Galmarini, MV., Loiseau, AL., Visalli, M., & Schlich, P. (2016). Use of multi-intake temporal
770 dominance of sensations (TDS) to evaluate the influence of cheese on wine perception.
771 *Journal of Food Science*, *81*, S2566–S2577.
- 772 Galmarini, MV., Visalli, M., & Schlich, P. (2017). Advances in representation and analysis of
773 mono and multi-intake Temporal Dominance of Sensations data. *Food Quality and*
774 *Preference*, *56*, 247–255.
- 775 Greenhoff, K. & Buck, DT. (2006). Consumer measurement of drinkability. *European*
776 *Brewery Convention, Monograph 34, Symposium, Drinkability, Edinburgh, United*
777 *Kingdom, November 2006*.
- 778 Gomez-Corona, C., Chollet, S., Escalona-Buendia HB., & Valentin, D. (2017). Measuring
779 the drinking experience of beer in real context situations. The impact of affects, senses,
780 and cognition. *Food Quality and Preference*, *60*, 113–122.
- 781 Guinard, J., Souchard, A., Picot, M., Rogeaux, M., & Sieffermann, J. (1998). Sensory
782 determinants of the thirst-quenching character of beer. *Appetite*, *31*, 101-115.
- 783 Jaeger, SR., Cardello, AV., Chheang SL., Beresford, MK., Hedderley, DI. & Pineau, B.
784 (2017). Holistic and consumer-centric assessment of beer: A multi-measurement
785 approach. *Food Research International*, *99*, 287–297.
- 786 Jaeger, SR., Xia, Y., Le Blond, M., Beresford, MK., Hedderley, DI., & Cardello, AV. (2018).
787 Supplementing hedonic and sensory consumer research on beer with cognitive and
788 emotional measures, and additional insights via consumer segmentation. *Food Quality*

789 *and Preference*, 73, 117–134.

790 Jaeger, SR. & Porcherot, C. (2017). Consumption context in consumer research:
791 methodological perspectives. *Current Opinion in Food Science*, 15, 30–37.

792 Jager, G., Schlich, P., Tijssen, I., Yao, J., Visalli, M., De Graaf, C., & Stieger, M. (2014).
793 Temporal dominance of emotions: Measuring dynamics of food-related emotions during
794 consumption. *Food Quality and Preference*, 37, 87–99.

795 Kaltner, D. & Mitter, W. (2006). Possible impact of hops on taste and drinkability of beer.
796 *European Brewery Convention, Monograph 34, Symposium, Drinkability, Edinburgh,*
797 *United Kingdom, November 2006.*

798 **Kenkou Nippon 21 (Health Japan 21). Ministry of Health, Labour and Welfare.**

799 King, SC., Weber, AJ., Meiselman, HL., & Lv, N. (2004). The effect of meal situation, social
800 interaction, physical environment and choice on food acceptability. *Food Quality and*
801 *Preference*, 15, 645–653.

802 Kojima, H., Kaneda, H., Watari, J., Nakamura, Y., & Hayashi, T. (2009). Relationships
803 among throat sensation, beer flavor, and swallowing motion while drinking beer. *J. Am.*
804 *Soc. Brew. Chem.* 67(4), 217-221.

805 Köster, E. P. (2003). The psychology of food choice: Some often encountered fallacies.
806 *Food Quality and Preference*, 14, 359–373.

807 Lawless, HT., & Heymann, H. (1998). *Sensory Evaluation of Food: Principles and Practices,*
808 *Chapman & Hall, New York.*

809 Mattos, R. (2006). Recent review on beer drinkability. *European Brewery Convention,*
810 *Monograph 34, Symposium, Drinkability, Edinburgh, United Kingdom, November 2006.*

811 Mattos, R., & Moretti, R. (2005). Beer drinkability - A review. *Tech. Q. Master Brew. Assoc.*
812 *Am.* 42, 13-15.

813 Matuszewska, I., Barylko-Pikielna, N., Szczecinska, A., & Radzanowska, J. (1997).
814 Comparison of three procedures for consumer assessment of fat spreads: Short report.
815 *Polish Journal of Food and Nutrition Sciences*, 6, 139–142.

816 Meilgaard, M., Civille, GV. & Carr, BT. (2016). *Sensory Evaluation Techniques: 5th Edition,*
817 *CRC Press, New York.*

818 Nagao, Y., Kodama, H., Yamaguchi, T., Yonezawa, T., Taguchi, A., Fujino, S., Morimoto, K.,
819 & Fushiki, T. (1999). Reduced urination rate while drinking beer with an unpleasant taste
820 and off-flavor. *Biosci. Biotechnol. Biochem.* 63, 468-473.

- 821 Nagao, Y., Kodama, H., Yonezawa, T., Taguchi, A., Fujino, S., Nakahara, K., Haruma, K., &
822 Fushiki, T. (1998). Correlation between the drinkability of beer and gastric emptying.
823 *Biosci. Biotechnol. Biochem.* 62, 846-851.
- 824 Nijman, M., James, S., Dehrmann F., Smart, K., Ford, R. & Fort, J. (2019). The effect of
825 consumption context on consumer hedonics, emotional response and beer choice. *Food*
826 *Quality and Preference*, 74, 59–71.
- 827 Pineau, N. (2013). Extension of TDS to multi-bite evaluation. In *10th Pangborn sensory*
828 *science symposium, 11–15th August 2013, Rio de Janeiro, Brazil.*
- 829 Pineau, N., Cordelle, S., & Schlich, P. (2003). Temporal dominance of sensations: A new
830 technique to record several sensory attributes simultaneously over time. In *5th Pangborn,*
831 *symposium, July 20–24 (pp. 121), Boston.*
- 832 Pineau, N., Goupil de Bouillé, A., Lepage, M., Lenfant, F., Schlich, P., & Martin, N., et al.
833 (2012). Temporal Dominance of Sensations: What is a good attribute list? *Food Quality*
834 *and Preference*, 26(2), 159–165.
- 835 Pineau, N., Schlich, P., Cordelle, S., Mathonnière, C., Issanchou, S., & Imbert, A., et al.
836 (2009). Temporal Dominance of Sensations: Construction of the TDS curves and
837 comparison with time–intensity. *Food Quality and Preference*, 20(6), 450–455.
- 838 Posri, W., & MacFie, H. (2008). The influence of testing context on tea bag product
839 acceptance in central location tests. *Journal of Sensory Studies*, 23, 835–851.
- 840 Ramsey, I., Ross, C., Ford, R., Fisk, I., Yang, Q., Gomez-Lopez, J., & Hort, J. (2018). Using
841 a combined temporal approach to evaluate the influence of ethanol concentration on
842 liking and sensory attributes of lager beer. *Food Quality and Preference*, 68, 292–303.
- 843 Schlich, P. (2017). Temporal dominance of sensations (TDS): A new deal for temporal
844 sensory analysis. *Current Opinion in Food Science*, 15, 38–42.
- 845 Schlich, P., Urbano, C., & Visalli, M. (2013). Multi-bites or multi-sips TDS with untrained
846 subjects: A live demonstration on chocolates. In *10th Pangborn sensory science*
847 *symposium, 11–15th August 2013, Rio de Janeiro, Brazil.*
- 848 Silva, AP., Voss, H-P., van Zyl H., Hogg, T., de Graaf, C., Pintado, M. & Jager, G. (2019).
849 Effect of adding hop aroma in beer analysed by temporal dominance of sensations and
850 emotions coupled with temporal liking. *Food Quality and Preference*, 75, 54–63.
- 851 Simioni, SCC., Ribeiro, MN., de Souza, VR., Nunes, CA., & Pinheiro, ACM. (2018).
852 Multiple-sip temporal dominance of sensations associated with acceptance test: a study
853 on special beers. *J Food Sci Technol*, 55(3), 1164–1174.

- 854 Taldi, B. (2006). Yeast and fermentation, tools for increasing drinkability. *European Brewery*
855 *Convention, Monograph 34, Symposium, Drinkability, Edinburgh, United Kingdom,*
856 *November 2006.*
- 857 Thomas, A., van der Stelt, AJ., Prokop, J., Lawlor, JB., & Schlich, P. (2016). Alternating
858 temporal dominance of sensations and liking scales during the intake of a full portion of
859 an oral nutritional supplement. *Food Quality and Preference, 53*, 159–167.
- 860 Thomas, A., Visalli, M., Cordelle, S., & Schlich, P. (2015). Temporal drivers of liking. *Food*
861 *Quality and Preference, 40*, 365–375.
- 862 Thomson, D. & Bailey, P. (2006). Psychology of drinkability. *European Brewery Convention,*
863 *Monograph 34, Symposium, Drinkability, Edinburgh, United Kingdom, November 2006.*
- 864 Vandeputte, A., Romans, J., Pineau, N., & Lenfant, F. (2011). Innovative methods to assess
865 the evolution of the sensory characteristics during the tasting of a full product portion
866 (several bites). In *9th Pangborn sensory science symposium, 4–8th September 2011,*
867 *Toronto, Canada.*
- 868 van Kleef, E., van Trijp, HCM., & Luning, P. (2005). Consumer research in the early stages
869 of new product development: a critical review of methods and techniques. *Food Quality*
870 *and Preference, 16*, 181–201.
- 871 van Opstaele, F., Goiris, K., Stryn, E., de Cooman, L., & Aerts, G. (2006). Novel hop-derived
872 ingredients and drinkability. *European Brewery Convention, Monograph 34, Symposium,*
873 *Drinkability, Edinburgh, United Kingdom, November 2006.*
- 874 Vazquez-Araujo, L., Parker, D., & Woods, E. (2013). Comparison of temporal-sensory
875 methods for beer flavor evaluation. *Journal of Sensory Studies, 28*, 387–395.
- 876 Zandstra, EH., de Graaf, C., van Trijp, HCM., & van Staveren, WA. (1999). Laboratory
877 hedonic ratings as predictors of consumption. *Food Quality and Preference, 10*, 411–
878 418.
- 879 Zorn, S., Alcaire, F., Vidal, L., Giménez, A., & Ares, G. (2014). Application of multiple-sip
880 temporal dominance of sensations to the evaluation of sweeteners. *Food Quality and*
881 *Preference, 36*, 135–143.

Figure 1. Test flow of “Multiple-sip Drinkability Test.”

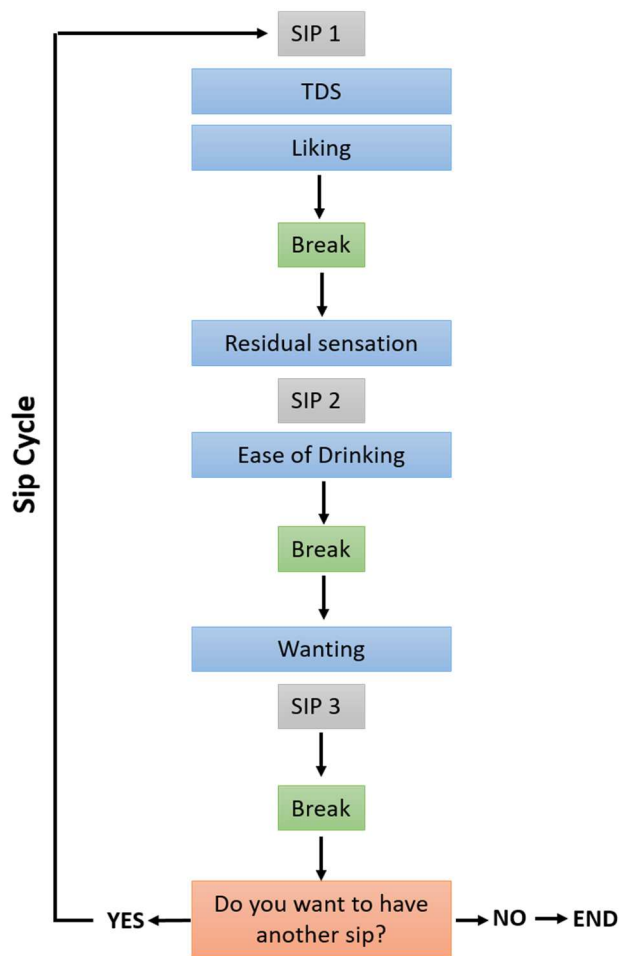


Figure 2. Development of *wanting* - Expert Panel
Product by Sip (Means \pm Standard Error x 2; replicates pooled).

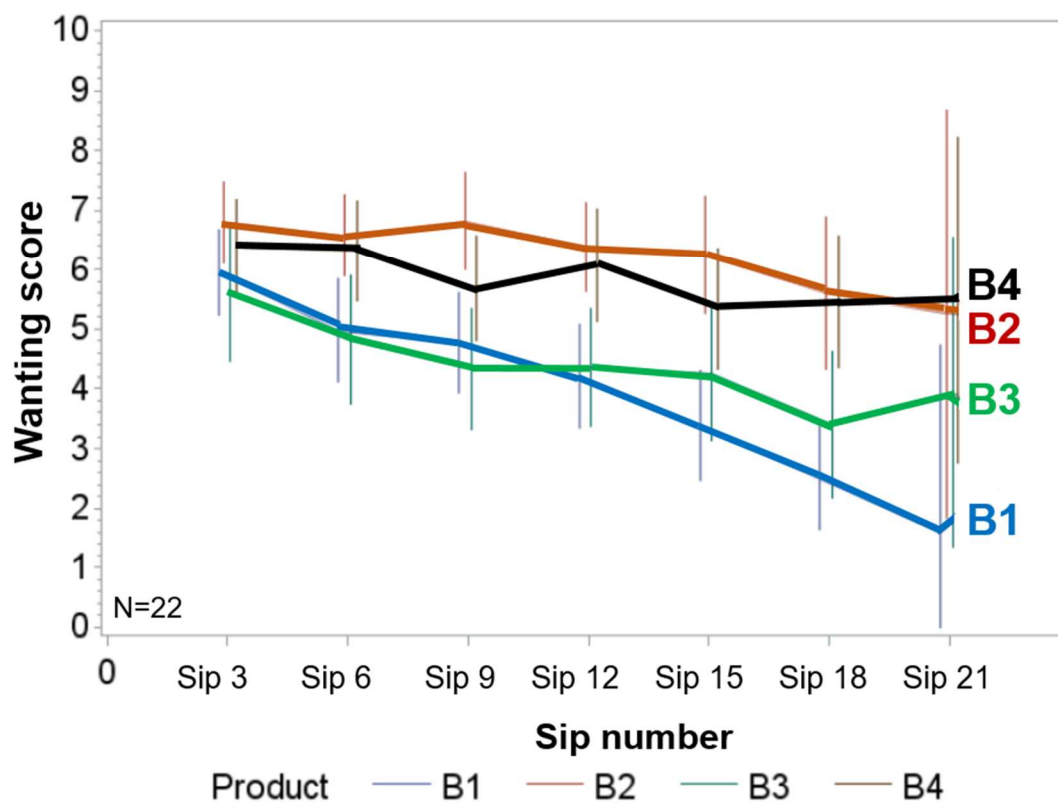


Figure 3. Development of *ease of drinking* - Expert Panel

Product by Sip (Means \pm Standard Error \times 2; replicates pooled).

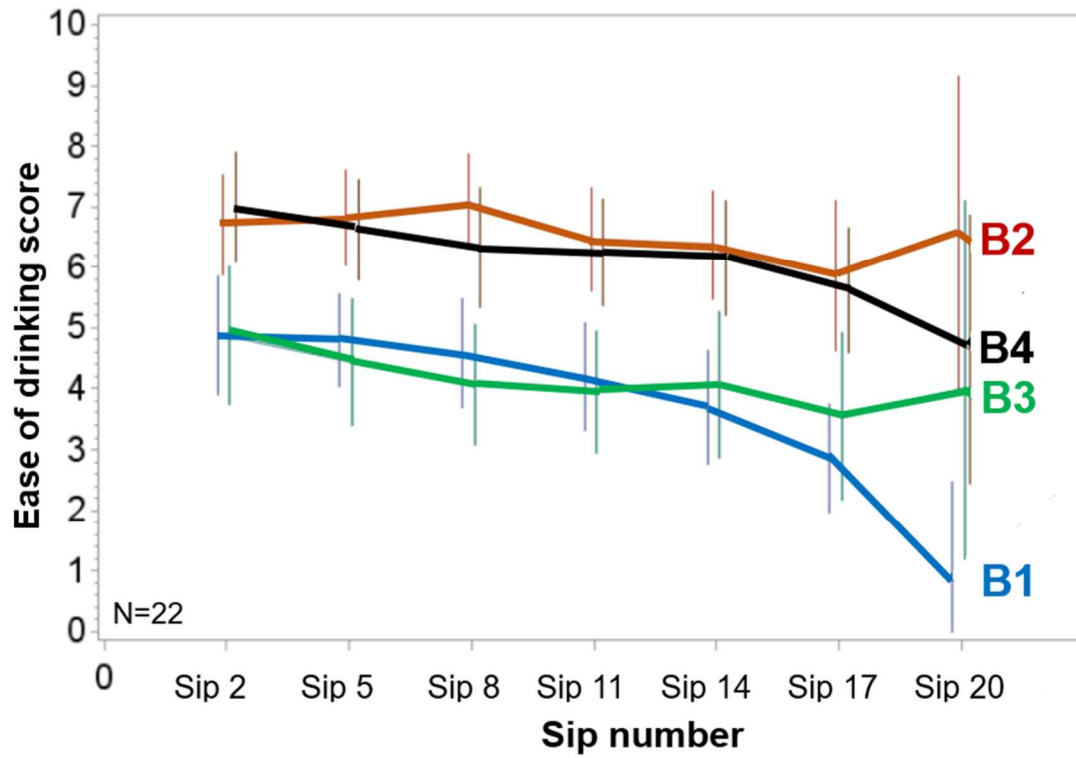


Figure 4. Development of *wanting* compared with *liking*: Sip Cycle LS-Means over products - Expert Panel.

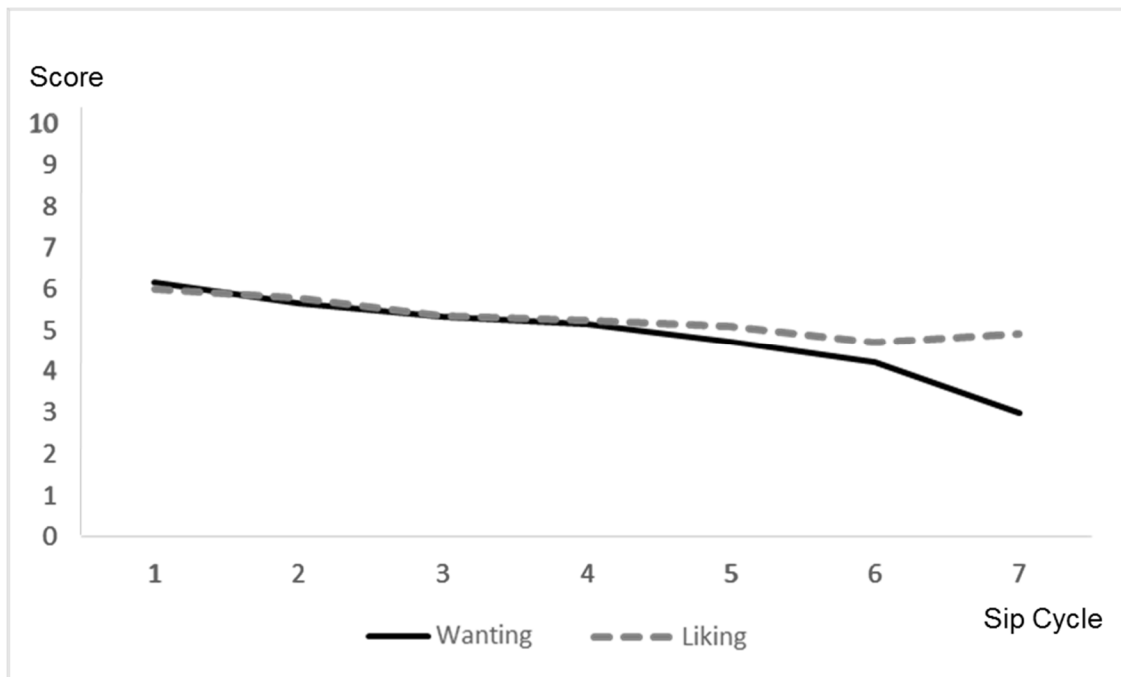


Figure 5. Development of sensory perceptions: TDS Band Plot and Products' characteristics.

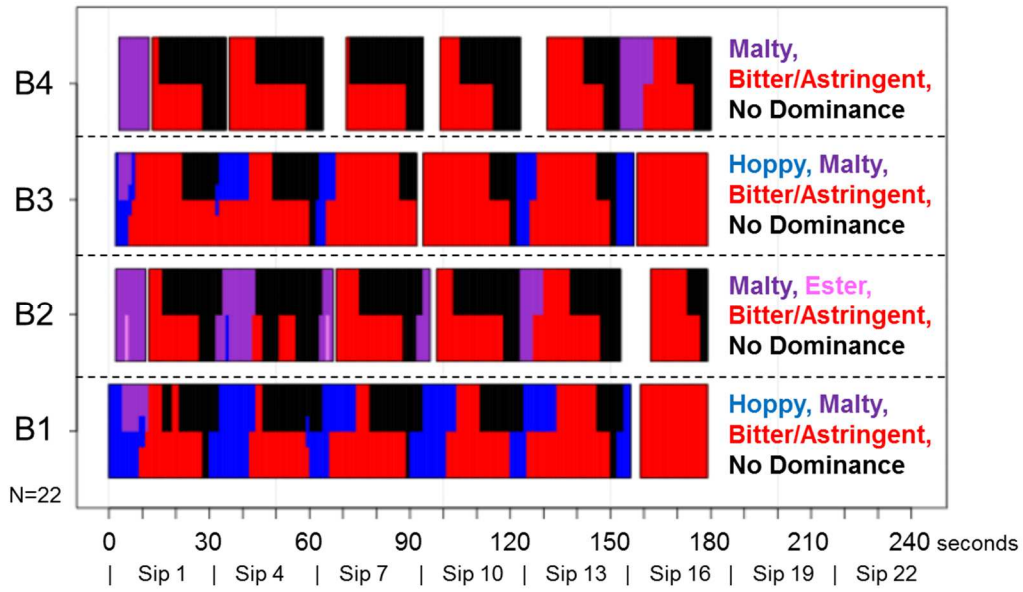


Figure 6. Development of *wanting* - Naïve Panel
Product by Sip (Means \pm Standard Error x 2).

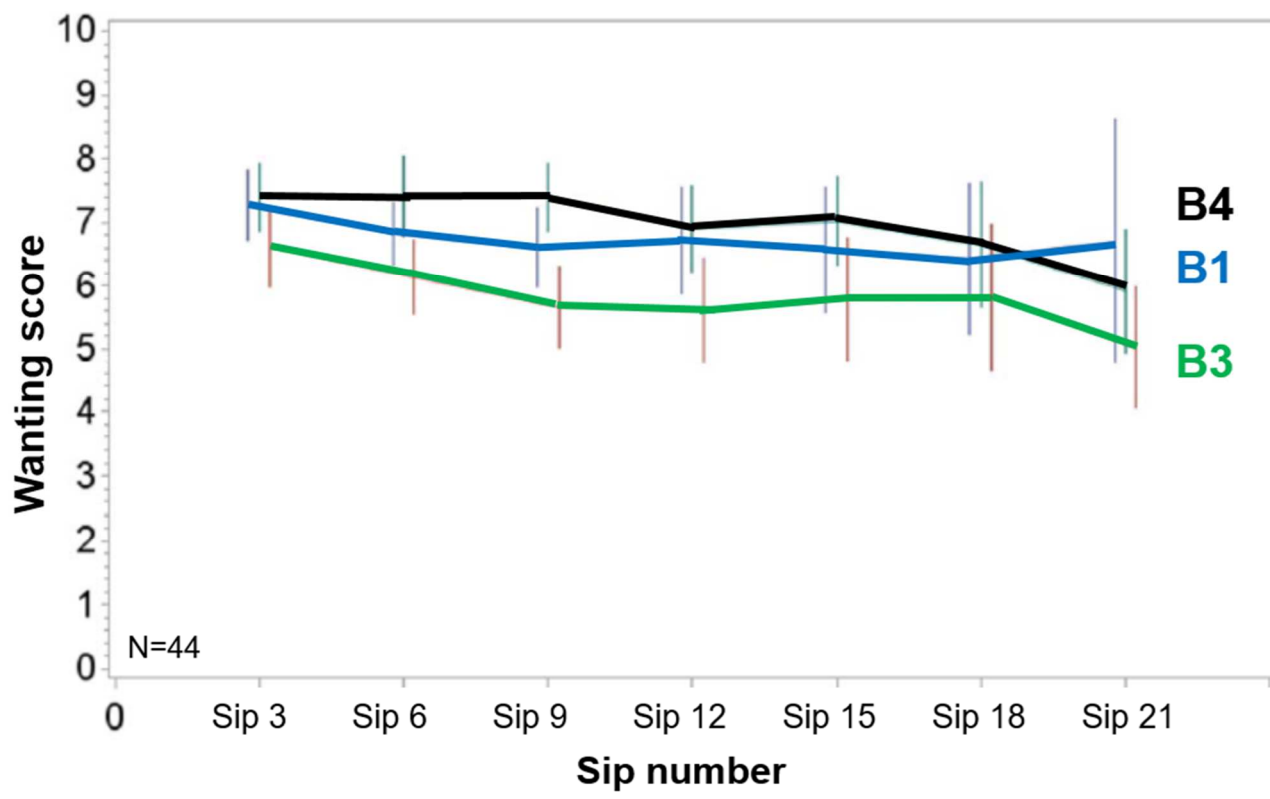


Figure 7. Development of *wanting* compared with *liking*: Sip Cycle LS-Means over products - Naïve Panel.

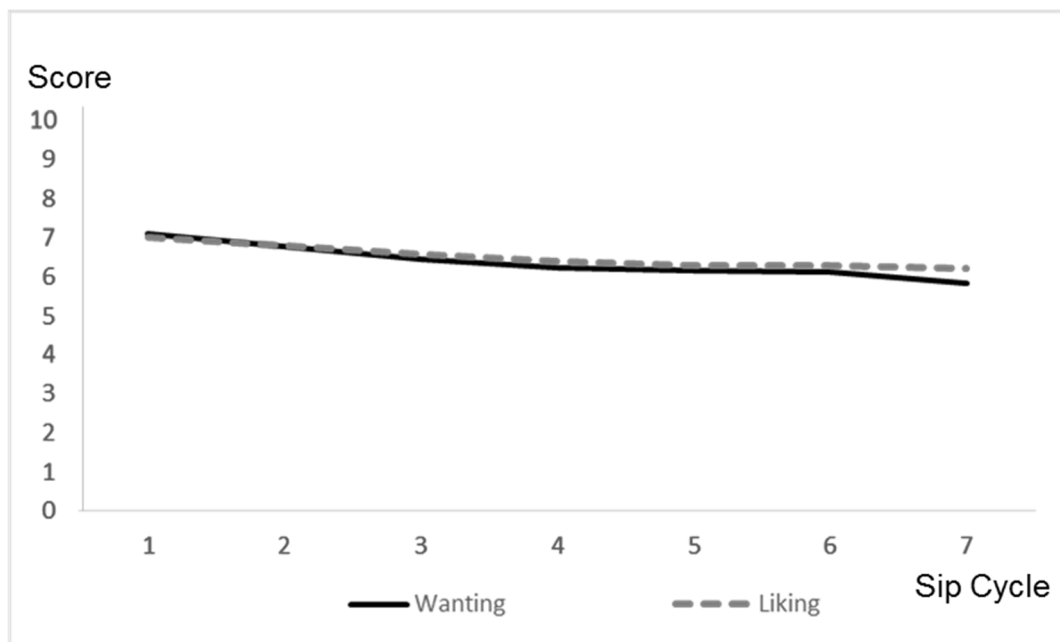


Figure 8a. TDS difference curves of the samples B4 and B3 – Naïve panel.

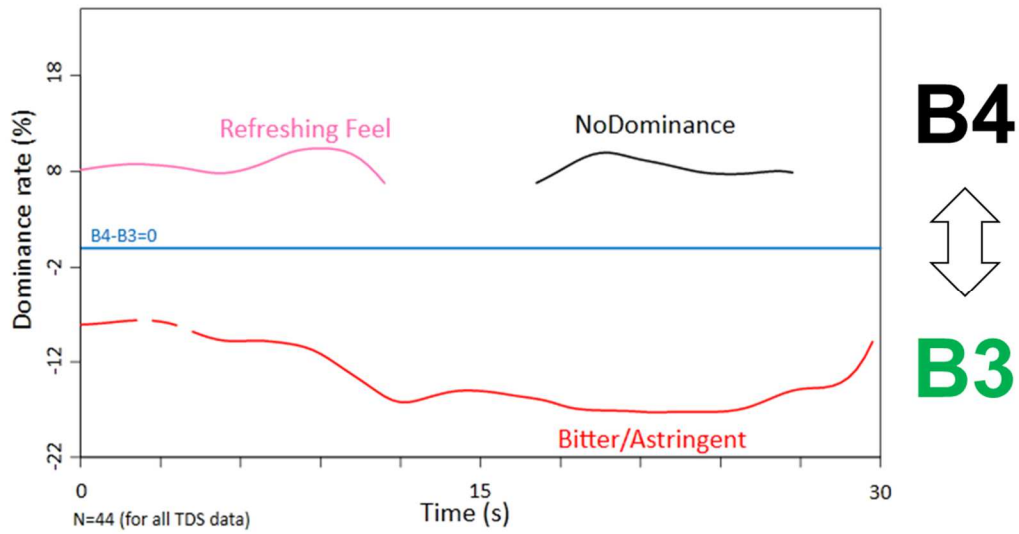


Figure 8b. TDS difference curves of the samples B4 and B1 – Naïve panel.

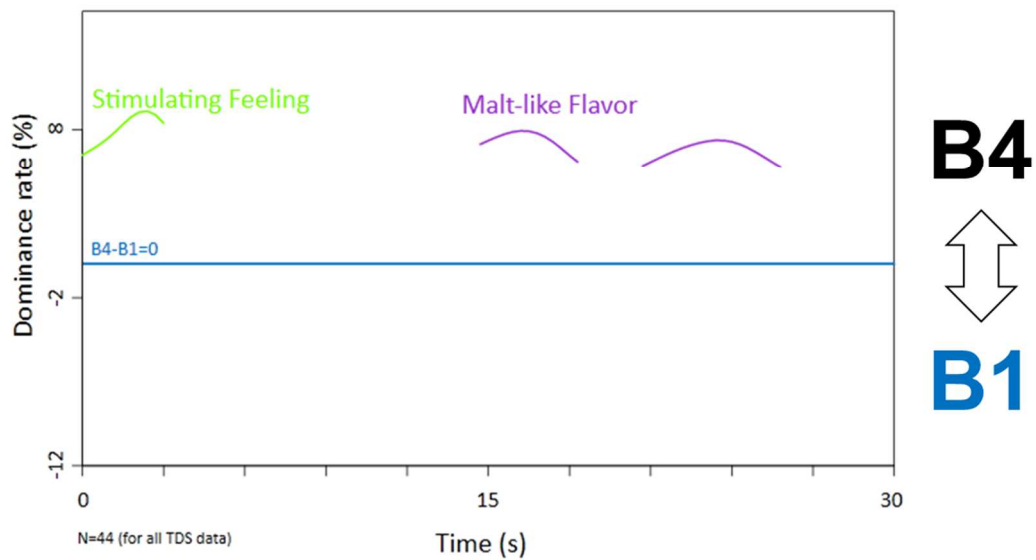


Figure 9. TDS difference curves of the samples B4 and B1 by wanting segmentation – Naïve panel.

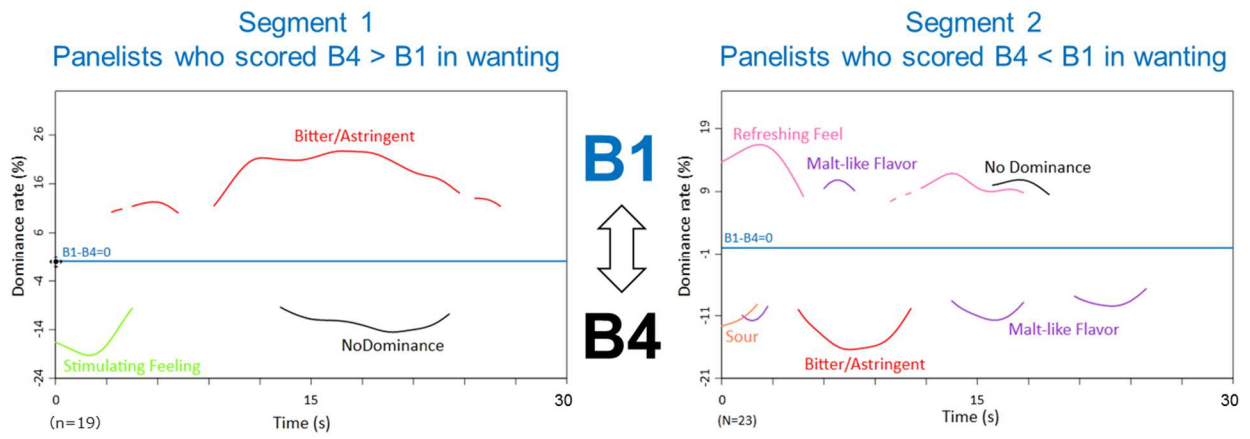


Figure 10. Development of Duration: *Bitter/Astringent* - Naïve Panel
Dominance duration means with their 95% confidence level.

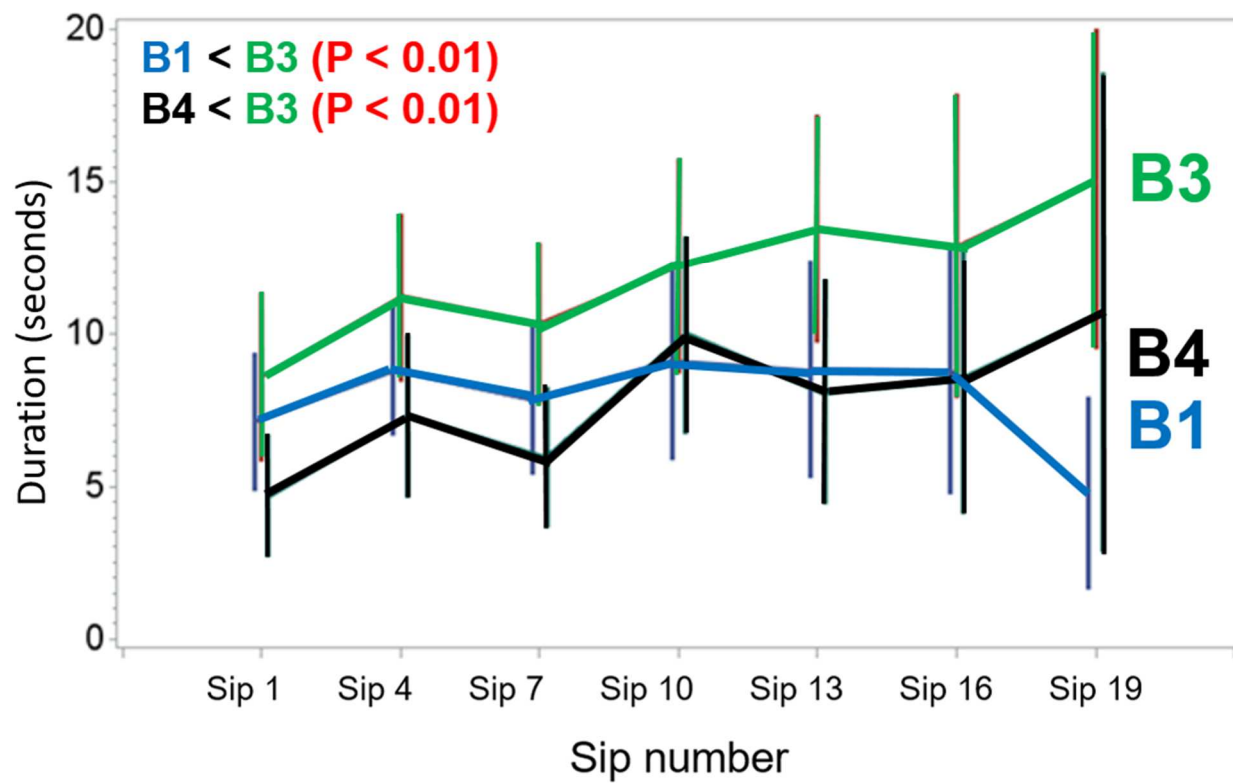


Figure 11a. Volume of beer actually consumed by the expert panel (Mean \pm Standard Error x 2).

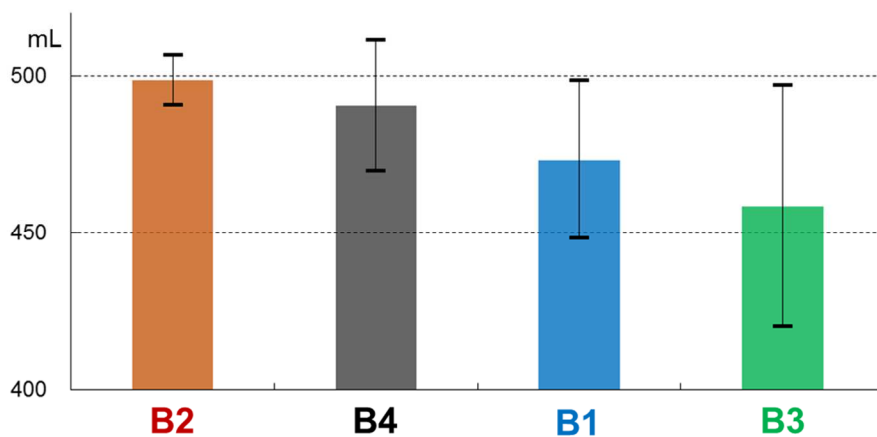


Figure 11b. Volume of beer actually consumed by the naïve panel (Mean \pm Standard Error x 2).

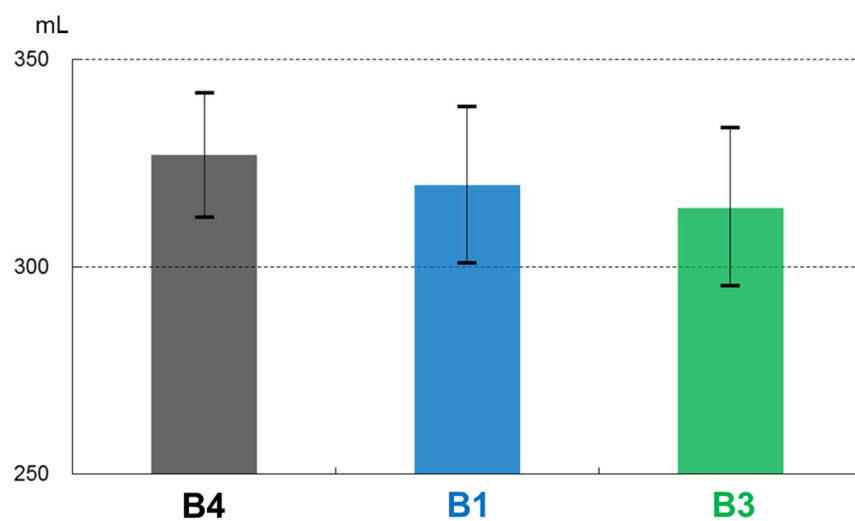


Table 1

ABV, IBU, Linalool and sensory **descriptions** of beer samples used in this study.

Sample	Alcohol content (% ABV)	Bitterness (IBU)	Linalool (ppb)	Sensory descriptions
B1	5	22.5	21.3	<i>Hoppy, Malty, Bitter</i>
B2	5	19.5	1.5	<i>Less standout flavors, Light finish</i>
B3	5	29	2.4	<i>Malty, Bitter, Full body</i>
B4	5	19.5	1.2	<i>Less standout flavors, Light finish</i>

(Note: ABV = alcohol by volume. IBU = international bitterness unit. ppb = parts per billion.)

(Note: Information about alcohol content was obtained from the product package.

Bitterness and linalool were analyzed in the R&D center at Asahi Breweries, Ltd.

Sensory descriptors were obtained from the results of quantitative descriptions and sensory comments, both of which were acquired from the Asahi expert beer sensory panel.)

Table 2a

ANOVA of dynamic wanting, liking, ease of drinking and residual sensation: Repeated Mix Model (1/2) - Expert Panel -
Using Restricted Maximum Likelihood (REML) algorithm with first-order heterogeneous auto-regressive covariance structure among sips

Descriptor	p-values of tests of fixed effects			Product LS-Means			
	Sip	Product	Product*Sip	B1	B2	B3	B4
Wanting	0.0000	0.0037	0.0310	3.65a	5.95b	3.78a	5.59b
Liking	0.0023	0.0155	0.2672	4.46ab	6.29b	4.26a	5.89b
Ease of Drinking	0.0172	0.0003	0.6351	3.88a	6.62b	3.83a	5.98b
Residual Sensation	0.2290	0.4738	0.5048	4.18a	3.98a	4.79a	3.65a

Means with two different letters are significantly different ($p < 0.05$).

Table 2b

ANOVA of dynamic wanting, liking, ease of drinking and residual sensation: Repeated Mix Model (2/2) - Expert Panel -
Using Restricted Maximum Likelihood (REML) algorithm with first-order heterogeneous auto-regressive covariance structure among sips

Descriptor	Sip Cycle LS-Means							p-values of Sip Effect	p-values of Sip Effect by Product			
	1	2	3	4	5	6	7		B1	B2	B3	B4
Wanting	6.18	5.67	5.35	5.15	4.71	4.23	2.99	0.0000	0.0000	0.2367	0.0000	0.1576
Liking	6.02	5.78	5.36	5.26	5.09	4.69	4.92	0.0023	0.0008	0.9161	0.0084	0.2998
Ease of Drinking	5.85	5.67	5.46	5.17	5.08	4.71	3.84	0.0172	0.2475	0.5702	0.0869	0.1434
Residual Sensation	4.18	3.74	3.80	3.60	3.65	4.14	4.93	0.2290	0.5085	0.8558	0.0447	0.4299

Table 3a

ANOVA of dynamic wanting, liking and ease of drinking: Repeated Mix Model (1/2) - Naïve Panel -

Using Restricted Maximum Likelihood (REML) algorithm with first-order heterogeneous auto-regressive covariance structure among sips

Descriptor	p-values of tests of fixed effects			Product LS-Means			p-values for comparing products		
	Sip	Product	Product*Sip	B1	B3	B4	B1-B3	B1-B4	B3-B4
Wanting	<0.0001	<0.0001	0.1767	6.52b	5.69a	6.95b	<0.0001	0.1007	<0.0001
Liking	0.0040	<0.0001	0.5063	6.74b	5.70a	7.11b	<0.0001	0.1346	<0.0001
Ease of Drinking	0.0041	<0.0001	0.6762	6.80b	5.66a	7.23b	<0.0001	0.1037	<0.0001

Means with two different letters are significantly different ($p < 0.05$).

Table 3b

ANOVA of dynamic wanting, liking and ease of drinking: Repeated Mix Model (2/2) - Naïve Panel -
Using Restricted Maximum Likelihood (REML) algorithm with first-order heterogeneous auto-regressive covariance structure among sips

Descriptor	Sip Cycle LS-Means							p-values of Sip Effect	p-values of Sip Effect by Product		
	1	2	3	4	5	6	7		B1	B3	B4
Wanting	7.09	6.79	6.46	6.23	6.15	6.14	5.85	<0.0001	0.0101	0.0001	0.1893
Liking	7.03	6.81	6.58	6.40	6.28	6.28	6.24	0.0040	0.1616	0.0084	0.4305
Ease of Drinking	7.03	6.82	6.67	6.51	6.40	6.42	6.11	0.0041	0.4283	0.0172	0.1261

Figure #	Title
1	Figure 1. Test flow of "Multiple-sip Drinkability Test."
2	Figure 2. Development of "wanting" - Expert Panel Product by Sip (Means \pm Standard Error x 2; replicates pooled).
3	Figure 3. Development of "ease of drinking" - Expert Panel Product by Sip (Means \pm Standard Error x 2; replicates pooled).
4	Figure 4. Development of "wanting" compared with "liking": Sip Cycle LS-Means over products - Expert Panel.
5	Figure 5. Development of sensory perceptions: TDS Band Plot and Products' characteristics.
6	Figure 6. Development of "wanting" - Naïve Panel Product by Sip (Means \pm Standard Error x 2).
7	Figure 7. Development of "wanting" compared with "liking": Sip Cycle LS-Means over products - Naïve Panel.
8a	Figure 8a. TDS difference curves of the samples B4 and B3 – Naïve panel.
8b	Figure 8b. TDS difference curves of the samples B4 and B1 – Naïve panel.
9	Figure 9. TDS difference curves of the samples B4 and B1 by wanting segmentation – Naïve panel.
10	Figure 10. Development of Duration: "Bitter/Astringent" - Naïve Panel Dominance duration means with their 95% confidence level.
11a	Figure 11a. Volume of beer actually consumed by the expert panel (Mean \pm Standard Error x 2).
11b	Figure 11b. Volume of beer actually consumed by the naïve panel (Mean \pm Standard Error x 2).