1
2
3
4
5
6
7
8
9
10
11

12 **Sensory nutrition: The role of taste in the reviews of commercial food products**

13

14 Danielle. R. Reed, Joel D. Mainland, and Charles J. Arayata

15 Monell Chemical Senses Center, Philadelphia, PA 19104

16

17

18

19

20

21 Please address correspondence to:

22 Danielle R. Reed, Ph.D.
23 Associate Director
24 Monell Chemical Senses Center
25 Philadelphia PA 19104
26 a

2

27   Abstract

28   Many factors play a role in choosing what to eat or drink. Here we explored the role of

29   sensation to explain these differences, drawing on consumer reviews for commercially available

30   food products sold through an online retailer. We analyzed 393,568 unique food product

31   reviews from Amazon customers with a total of 256,043 reviewers rating 67,553 products. Taste-

32   associated words were mentioned more than words associated with cost, food texture, customer

33   service, nutrition, smell, or those referring to the trigeminal senses, e.g., spicy. We computed the

34   overall number of reviews that mentioned taste qualities: the word *taste* was mentioned in over

35   30% of the reviews (N= 142,768), followed by *sweet* (10.7%, N=42,315), *bitter* (2.9%, N=11,424),

36   *sour* (2.1%, N=8,252), and *salty* (1.4%, N=5,688). We identified 38 phrases used to describe the

37   evaluation of sweetness, finding that 'too sweet' was used in nearly 0.8% of the reviews and

38   oversweetness was mentioned over 25 times more often than under-sweetness. We then focused

39   on 'polarizing' products, those that elicited a wide difference of opinion (as measured by the

40   ranges of the product ratings). Using the products that had more than 50 reviews, we identified

41   the top 10 most polarizing foods (i.e., those with the largest standard deviation in ratings) and

42   provide representative comments about the polarized taste experience of consumers. Overall,

43   these results support the primacy of taste in real-world food ratings and individualized taste

44   experience, such as whether a product is 'too sweet'. Analysis of consumer review data sets can

45   provide information about purchasing decisions and customer sensory responses to particular

46   commercially available products and represents a promising methodology for the emerging

47   field of sensory nutrition.

48 Keywords: taste, sweetness, bitterness, consumers, segmentation

49

50   'Sensory nutrition' is a research area examining how sensation affects what an animal (person)

51   chooses to eat or drink and how these sensory-motivated choices affect their nutritional health.

52   However, many studies that fall into the 'sensory-nutrition' study arena analyze these processes

53   under artificial circumstances, e.g., one meal or one snack in the laboratory. Here we capitalize

54   on real human behavior in a large arena of food choice, an on-line retailer that offers thousands

55   of choices, examining the importance of sensory experience in the written comments of those

56   who have purchased foods and drinks.

57   Sensory nutrition as a research area has become more essential because of the increasing

58   realization that human foods have become 'hyperpalatable', engineered to make them so

59   desirable from a sensory perspective that they are hard to resist and the overeating of these

60   foods leads to obesity and other disease associated with overconsumption. However, while

61   taste is often studied in simplified foods systems, like sugar dissolved in plain water [1], real-

62   world foods are rarely evaluated for taste in a choice context in laboratories studies, although

63   these types of studies are commonly done to evaluate products for the marketplace. Most

64   broadly, the assertion is that the taste of foods drives overconsumption but that people differ in

65   their perception. The larger question is whether personal differences in taste experience for

66   different foods drive overconsumption of that food, and can ultimately predict who will not

67   only like a certain food, but perhaps who cannot resist that food and why.

68   We took a step toward this larger question by analyzing how the sense of taste factors into food

69   ratings by examining reviews of commercial foods written by customers of an online retailer.

70 These ratings contain both a text narrative about the food and a 'star' rating, from one to five,

71 with five stars representing the highest score. We were interested in how often taste was

72 mentioned by reviewers and which of the taste qualities were mentioned most often. We were

73 also interested in the idea that certain products were more polarizing among reviewers, with

74 extreme diversity—love it or hate it—in responses, reasoning that a list of polarizing products

75 might be a tool for future research to understand food choice. We extract and present reviewer

76 comments about polarizing food products to illustrate what role taste played in creating the

77 diverse viewpoints. We contrasted taste-specific works with those for texture (e.g., lumpy,

78 creamy, soft, hard) and odor (e.g., smell) as well as price and customer service to see what

79 sensory words were more common among reviewers.

80 **Methods**

81 **Data and its structure**. We obtained the data set through the open-source data competition site

82 Kaggle (www.kaggle.com), where the data are offered freely to all under a Creative Commons

83 public license. We performed all analyses in R (version 3.5.2) [2] and made this R script

84 available on Github (https://github.com/joelmainland/Taste_is_king). The data contained ten

85 variables: *Id* (each review has a unique identifier), *ProductId* (unique identifier for the product),

86 *UserId* (identifier for the user), *ProfileName* (the self-assigned user name), *HelpfulnessNumerator*

87 (the number of people who found the review helpful), *HelpfulnessDenominator* (the number of

88 people who found the review helpful or not), *Score* (rating of the product on a 1 to 5 scale, with

89 5 being best and 1 being worst), *Time* (time of day the review was submitted), *Summary* (brief

90 review), and *Text* (full review). The reviews were submitted over a ten-year period ending in

91 October 2012. There were 568,454 reviews but some were duplicated (same text for similar

92   products) and were removed prior to analysis. Analysis of the de-duplicated data indicated

93   there were a total of 393,568 reviews of 67,553 products by 256,043 unique reviewers (as defined

94   by unique reviewer IDs; however, this does not ensure absolute uniqueness because the same

95   person might have more than one ID. All reviewers were 'verified purchasers,' meaning the

96   online retailer (Amazon) had a record that the reviewer purchased the food item being

97   reviewed.

98   **Word use analysis.** We used the Word2Vec package to create a vector representation of words

99   based on the distributional hypothesis, namely that words that appear in the same context share

100   semantic meaning [3]. Likes and dislikes can arise because of food tastes, textures or odors, cost,

101   nutrition, or quality of service. Thus, the vector representation was used to identify clusters of

102   semantically similar words based on seed words from these categories (**Supplementary Table**

103   **1**). We counted the numbers of reviews that contained words from these clusters. In some cases

104   we elected not to use individual odor descriptors because odors are often described using the

105   word for the object producing the odor [4]., e.g., honey smells like honey, peaches smell like

106   peaches.

107   Building on the results that we explain below, we also probed for more details about sweet

108   (because it was the most commonly used taste word). We extracted all phrases using the word

109   'sweet' and cleaned the data by eliminating common but irrelevant uses of the word sweet, e.g.,

110   'sweet potato'. We next extracted 38 phrases that captured the majority of ways sweetness was

111   discussed, and tallied the number of times the phrase was used, e.g., 'too sweet' versus 'not too

112   sweet' and used to calculate the percent of the reviews which contained that phrase. The

113   phrases were placed into one of three categories: oversweet, (e.g., cloying, sickeningly sweet),

114   under-sweet (e.g., not sweet enough for me) and neutral. We then tabulated the percentage of

115   comments within each of the three categories.

116   **Polarization**. We extracted all food products that had 50 or more reviews and computed the

117   standard deviation of the ratings (on the star scale). We refer to the foods with the largest

118   standard deviation as 'polarizing' foods. We chose the top 10 polarizing foods for a more in-

119   depth analysis. As we mentioned above, each product is identified by a unique ID number. To

120   find out which foods were associated with which ID number, we automatically extracted data

121   about the Amazon product page for each ID to obtain the product title. For each of the 10 most

122   polarizing products, two readers evaluated the narrative portion of each review (*Summary* plus

123   *Text*) and extracted representative comments about taste.

124   **Results**

125   **Overview.** We examined several global categories using seed words: taste and related words,

126   price and related words, likewise customer service, texture, smell and trigeminal (e.g.,

127   spiciness). In this analysis, the predominant word used in reviews was 'taste' with over 30% of

128   reviews using this word; fewer than those who mentioned 'price' (**Table 1**). To examine

129   whether this method of generating words was valid, we compared this list of words obtained to

130   those words used by sensory panels in the food industry to describe texture [5], finding

131   substantial agreement, e.g., hard, rough. For smell, we elected not to use individual odor

132   descriptors as seed words because odors are often named for the physical source [4] and we

133   cannot easily differentiate when the word 'coffee' is used as a product description from when it

134   is used to describe an odor. When the results are aggregated over the five categories, the results

135   show that 'taste' is mentioned more often than texture, customer service, cost, health, smell, and

136   trigeminal sensations (**Figure 1**). Reviewers mentioned *sweet* far more often than any other taste

137   quality (10.75%), followed by *bitter* (2.90%), *savory* (0.27%), *sour* (2.10%), and *salty* (1.45%).

138   *Umami* is a synonym for *savory* but a word that is rarely used by reviewers (0.02%).

139   We also probed in more detail about sweet and sweetness and identified 38 phrases used to

140   indicate this property. We considered each phrase and its negation, e.g., 'sweet enough' versus

141   'not sweet enough' and parsed each phrase into one of three categories, over-sweet, under-

142   sweet or neutral. The results were striking. Almost 1 percent of all reviews, regardless of food

143   type, used the phrase 'too sweet' indicating that excessive sweetness is often mentioned. When

144   evaluating the pattern of reviews that mention sweetness, over-sweetness was mentioned more

145   than 25 times more often than under-sweetness (**Table 2**). See **Supplemental Table 2** for a list of

146   the 38 phrases, the counts, and percentage of time each phrase was used in all reviews.

147   **Polarizing products**. For this analysis, we excluded products with fewer than 50 reviews,

148   reducing the number of reviews by roughly half (N=109,698) and the number of products from

149   67,533 to 908. We computed the standard deviation for each remaining product and ranked the

150   products from highest to lowest standard deviation. After excluding products not intended for

151   human consumption (e.g., pet food), we selected the top 10 products (**Table 3**). Standard

152   deviations ranged from 1.82 (most polarizing) to 0.21 (least polarizing).

153   The top two factors for polarization were (a) formulation changes in which a product initially

154   liked was changed and got negative reviews and (b) diverse views about the taste of the

155   product. For instance, for formulation change, consumers objected to increases in sugar in a

156   formerly beloved cereal. Illustrative phrases that highlight the opinion diversity were manually

157   extracted and are listed in **Table 4**.

158   <div align="center">**Discussion**</div>

159   The field of 'Sensory nutrition' brings together knowledge and methods from sensory,

160   physiology and nutrition sciences to understand the key drivers of nutrient choices so that we

161   can modulate diet to promote human health. Taste is often described as a primary influence on

162   human food selection and intake in nutrition and biopsychology research, e.g., [6]. Studies such

163   as the one just referenced rely on data from several thousand people, but here we demonstrate

164   the primacy of taste among nearly half a million respondents who are unaware of the import of

165   their commentary, essentially catching consumers responding in an unself-conscious way.

166   These results demonstrate that when consumers write about food, rather than cost or its

167   nutritional benefits, they write about taste. Taste is often applied generically to the flavor of

168   food, which has more inputs from different sensory systems, such as the somatosensory system

169   (texture) and the olfactory system (smell)[7]. Therefore, these results could be construed broadly

170   to apply to food flavor, not to taste as narrowly defined by sensory biologists.

171   **Taste qualities.** We learned that *sweet* was the taste quality mentioned most often, almost three

172   times more often than *bitter*, the next closest word used, followed distantly by *sour* and *salty*.

173   This result was a surprise because the opinions about bitterness would be complex and worthy

174   of mention in food product reviews, either as a desirable feature, perhaps in coffee, or an

175   undesirable feature in foods that do not normally taste bitter. However, perhaps bitterness is so

176   rarely present in commercial foods that it is rarely mentioned. Likewise, it was surprising how

177   little saltiness was mentioned by reviewers, given the global attention to salt reduction for

178   health [8]. While the overconsumption of sugar and salt are common public health concerns ([8-

179   10]), consumers have much more to say about sweetness than saltiness, at least in this particular

180   venue. We have learned from analysis of the sweet receptor that different versions arise from

181   inborn genotype, and some people are more sensitive to sweet taste than others because of their

182   genotype [11-15]. With that point in mind, it is interesting that consumers complained more

183   about products being 'too sweet' rather than 'not sweet enough', indicating that the over-

184   sugaring of processed foods is undesirable for some people and that offering a range of

185   sweetness of products might be even more important that previously realized.

186   **Polarizing products.** We were also interested in polarizing products—those rated variably by

187   different reviewers. While reading these reviews, we noted several trends that appeared to

188   account in part for the polarized ratings. One issue was formulation change—if consumers had

189   bought a product in the past and been satisfied with it, only to find on repurchasing that the

190   formulation had changed (e.g., increasing the product's sugar content), they down-rated the

191   product. In some ways, these formulation changes muddy the analysis, because the consumers

192   are rating two different products listed with the same product ID.

193   Often the diversity in viewpoint appeared to arise from different perspectives on a product's

194   taste. Some reviewers extolled the desirable taste of a particular product whereas others

195   disliked it, sometimes going so far as to berate other reviewers for their opinions. One

196   prominent example was the diversity of viewpoint on product sweetness, which is supported

197   by laboratory-based studies of sweet likers and dislikers, e.g., [16]. Recent studies of personal

198   differences in the liking of intensely sweetened foods suggest this may be an inborn trait [17]

199   Smell also contributed to the ratings of a few polarizing foods. There are other long-standing

200   debates about the desirable odor of some foods, a common Internet trope being the dislike of

201   cilantro [18, 19]. We also noted that smell figured prominently as a polarizing agent for some

202   products with a fishy odor, and we wonder whether the inborn variation in the ability to smell

203   the fishy odor trimethylamine might account for this diversity of viewpoint [20]. Although this

204   study does not allow us to match genotype to reviewer to understand whether these taste

205   disagreements owe to genetics, it does suggest that there is polarization and has identified a

206   handful of products that might be most profitably explored further for genetic effects.

207   **Limitations.** This study has several limitations. It is an analysis of data offered by the online

208   retailer that are freely available to all via a website that encourages exploratory data analysis of

209   large data sets (Kaggle). As such, we had no control over the collection of the data, the number

210   and type of variables included, or the accuracy of the data itself. Thus, all results must be

211   interpreted with this limitation in mind. Further, additional information such as item category

212   or other classifiers would have been useful to limit the analysis to only certain types of foods or

213   to add food type in the analysis. We also learned that foods can have the same identifier but

214   when the manufacturer changes the formulation (e.g., adds sugar) it may lead to polarization

215   because people who preferred the previous version of the product are now dissatisfied. This

216   type of polarization does not arise from diversity of viewpoint about the same food item and

217   these instances dilute the true polarizing response.

218   A second limitation was the limited choice of words and phrases to count in the reviews, which

219   capture few instances of related speech. While we used a variety of common phrases to capture

220   the concept of sweetness (too much or not enough), many reviewers might have used different

221   words to convey the concept of 'too sweet'. Also, some phrases about sweetness we classified as

222   'neutral' might be interpreted in one of several ways, e.g., 'on the sweet side'. Capturing the

223   intended meaning in text strings from real-world situations is imperfect, even using a large

224   palette of terms to describe a certain situation, and those limitations are present here.

225   Finally, there is an imprecision in the focus on taste, which encompasses several qualities for the

226   average person [7]. The reviewers can use this word both strictly, to evaluate the taste but not

227   smell or texture of the product, or more generally as a holistic quality (for example,

228   '…wonderful smell and flavor makes my coffee syrup taste like hazelnuts…' describes a smell,

229   but not a taste, and is flagged as both in our analysis). This limitation, the imprecision of the

230   word 'taste', is offset by the value of capturing real-world perceptions of foods, by people who

231   can report on whatever features they consider to be most important.

232   **The future.** This first study of food reviews from an on-line retailer, taste and polarizing foods

233   portends several avenues of the future research in sensory nutrition. Here we analyzed the

234   content of almost half a million reviews, but it is clearly possible and desirable to perform a

235   similar study on larger cohort, but one in which other information was available about the

236   reviewers, such as demographics, e.g., age and sex, other social and demographic information

237   (e.g., amount of formal education), medical history (e.g., diabetes, hypertension), genotype, and

238   other biological information, e.g., hormone concentrations in the blood or brain imaging. Large

239   scale studies that make the marketplace a laboratory become technically more possible with the

240   potential for linking food purchasing information with electronic medical records. While the

241   social, political and ethical barriers to these types of studies may be insurmountable, studies at

242   this scale are increasingly technically possible. These studies could reveal previously

243   overlooked patterns of food consumption and disease and point to new avenues of biology to

244   explain why people choose the foods they do.

245   **Conclusion.** We learned from these data that, when it comes to commercially available food

246   products, taste matters and that there are diverse viewpoints about some products that may

247   stem in part from differences in basic biology. Looking ahead and drawing on the research steps

248   from the preceding paragraph, it may be possible to find genotypes (for instance, in taste

249   receptors) that predict who will or will not like the taste of a given product. This idea may

250   translate into healthier foods, if producers can reduce sugar or salt in ways that appeal to

251   groups that prefer those products, by linking genetics (taste-related genotypes) to behavior

252   (food-purchasing habits). This study is a step toward realizing this idea.

253

254

255   Electronic resources

256   https://www.kaggle.com/snap/amazon-fine-food-reviews

257   https://github.com/joelmainland/Taste_is_king

258

265

266 **Table 1**: Word counts by category

Word	Count	Percent	Category
taste	121447	30.86	Taste
price	53327	13.55	Price
order	52279	13.28	Customer Service
sweet	42315	10.75	Taste
light	27329	6.94	Texture
ship	26640	6.77	Customer Service
hard	23253	5.91	Texture
rough	22058	5.60	Texture
healthy	19674	5.00	Health
smell	17457	4.44	Smell
organic	16239	4.13	Health
texture	15544	3.95	Texture
diet	13514	3.43	Health
fine	13475	3.42	Texture
arrived	12368	3.14	Customer Service
deal	12259	3.11	Price
bitter	11424	2.90	Taste
cheaper	9481	2.41	Price
sour	8252	2.10	Taste
value	7506	1.91	Price
spicy	6945	1.76	Trigeminal
aroma	6724	1.71	Smell
salty	5688	1.45	Taste
service	5613	1.43	Customer Service
condition	5264	1.34	Customer Service
sale	5053	1.28	Price
healthier	4375	1.11	Health
safe	3941	1.00	Health
bland	3889	0.99	Trigeminal
kick	3868	0.98	Trigeminal
scent	3232	0.82	Smell
fruity	2045	0.52	Smell
odor	1803	0.46	Smell
tabasco	572	0.15	Trigeminal
spiciness	395	0.10	Trigeminal

267 'Trigeminal' refers to the common chemical sense, e.g., burning, stinging, cooling.

268   **Table 2:** Sweet taste phrase use in reviews

Count	Sweetness	Percent
7230	Over	56.2
5370	Neutral	41.7
268	Under	2.1

269   'Over' are phrases about excessive sweetness, e.g., 'too sweet'; 'Under' are phrases indicating
270   the product is not sweet enough. 'Neutral' indicates the item is not either too sweet or not sweet
271   enough. See Supplemental Table 2 for a list of all tallied sweet-relevant phrases

272

273

274   **Table 3:** Products with the highest standard deviation in rating (polarization)

	Product ID	Product name	n	Mean	SD
1	B001M08YZA	Special K Cereal, Protein	109	2.83	1.82
2	B00507A02Q	blk Premium Alkaline Water Infused w/ Fulvic Trace	126	2.84	1.78
3	B002OMV09W	Ghost Chili Pepper	58	2.90	1.77
4	B000HDKZKU	Enjoy Life Chewy Bars	68	2.75	1.77
5	B000EM9E2Y	Just the Cheese Popped Cheese, Butter Flavor	58	2.60	1.76
6	B000F6SNPS	Good Earth Herbal Tea, Sweet & Spicy	143	3.69	1.76
7	B000CRIBCA	High Protein Bars by Think Thin-On the Go	69	3.07	1.76
8	B000AQJRWG	Tofu Shirataki Noodles Spaghetti Shape	75	3.19	1.74
9	B002CENRLG	Shirataki Noodles	101	3.54	1.72
10	B002EDEMLY	Red Vines Red Original Licorice Twists	77	3.69	1.73

275   n = number of reviews per product. Mean = mean star rating for each product. SD = standard
276   deviation for product ratings.
277

278   **Table 4:** Examples of polarized taste comments about the same product

Abbreviated Product Name	Positive Taste	Negative Taste
Special K Cereal, Protein	'taste[s] pretty darn good'	'it tastes totally horrible!!'
blk Premium Alkaline Water	'it taste[s] just like water with no funny after taste'	'this stuff tastes like dirty coins'
Enjoy Life Chewy Bars	'the bars have a good taste'	'tastes like cardboard. zero flavor'
Just the Cheese Popped Cheese	'I think they taste great'	'worst tasting stuff that I have ever eaten'
Good Earth Herbal Tea	'it tastes fantastic'	'sick with disappointment over degraded and truly unrecognizable taste'
High Protein Bars	'They taste more like a de[s]sert'	'tastes of a mixture of cardboard and chemicals'
Tofu Shirataki Noodles	'Tastes identical to real pasta'	'gross in tast[e] and texture'
Shirataki Noodles	'tasted fresh and enticing each time'	'taste like rotten fish'
Red Vines	'tasted delicious'	'they have a horrible after taste'

References

[1] Coldwell, S. E., Mennella, J. A., Duffy, V. B., Pelchat, M. L., Griffith, J. W., Smutzer, G., et al. Gustation assessment using the NIH Toolbox. Neurology. 2013,80:S20-4.

[2] R Core Team. R: A language and environment for statistical computing. In: Computing RFfS, editor. Vienna, Austria2013.

[3] Mikolov, T., Chen, K., Corrado, G., Dean, J. Efficient Estimation of Word Representations in Vector Space. arXiv:1301.37812013.

[4] Dravnieks, A., ASTM Committee E-18 on Sensory Evaluation of Materials and Products. Section E-18.04.12 on Odor Profiling. Atlas of odor character profiles. Philadelphia, PA: ASTM; 1985.

[5] Szczesniak, A. S. Consumer Awareness of Texture and of Other Food Attributes, Ii. J Texture Stud. 1971,2:196-206.

[6] Glanz, K., Basil, M., Maibach, E., Goldberg, J., Snyder, D. A. N. Why Americans Eat What They Do. J. Am. Dietetic Assoc. 1998,98:1118-26.

[7] Bartoshuk, L. M. Taste, smell and pleasure. The hedonics of taste. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc; 1991. p. 15-28.

[8] World Health Organization. Global action plan for the prevention and control of NCDs 2013-2020. 2013. p. 55.

[9] World Health Organization. Guideline: Sugars intake for adults and children. Geneva2015.

[10] WHO. Diet, nutrition and the prevention of chronic diseases. In: WHO/FAO, editor.: World Health Organization; 2003.

[11] Fushan, A. A., Simons, C. T., Slack, J. P., Drayna, D. Association between common variation in genes encoding sweet taste signaling components and human sucrose perception. Chem Senses. 2010,35:579-92.

[12] Fushan, A. A., Simons, C. T., Slack, J. P., Manichaikul, A., Drayna, D. Allelic polymorphism within the TAS1R3 promoter is associated with human taste sensitivity to sucrose. Curr Biol. 2009,19:1288-93.

[13] Joseph, P. V., Reed, D. R., Mennella, J. A. Individual differences among children in sucrose detection thresholds: Relationship with age, gender, and bitter taste genotype. Nurs Res. 2016,65:3-12.

[14] Mennella, J. A., Finkbeiner, S., Reed, D. R. The proof is in the pudding: children prefer lower fat but higher sugar than do mothers. Int. J. Obesity. 2012,36:1285-91.

[15] Mennella, J. A., Reed, D. R., Mathew, P. S., Roberts, K. M., Mansfield, C. J. 'A spoonful of sugar helps the medicine go down': bitter masking by sucrose among children and adults. Chem Senses. 2015,40:17-25.

[16] Looy, H., Callaghan, S., Weingarten, H. P. Hedonic response of sucrose likers and dislikers to other gustatory stimuli. Physiol Behav. 1992,52:219-25.

[17] Hwang, L. D., Lin, C., Gharahkhani, P., Cuellar-Partida, G., Ong, J. S., An, J., et al. New insight into human sweet taste: a genome-wide association study of the perception and intake of sweet substances. Am J Clin Nutr. 2019.

[18] Tullo, A. Vile weed or essential ingredient? Chemical and Engineering News. 2010,88:72.

[19] Eriksson, N., Wu, S., Do, C. B., Kiefer, A. K., Tung, J. Y., Mountain, J. L., et al. A genetic variant near olfactory receptor genes influences cilantro preference
. 2012.

[20] Amoore, J. E., Forrester, L. J. Specific anosmia to trimethylamine: the fishy primary odor. J Chem Ecol. 1976,2:49-56.

Supplemental Table 1. Word list from Word2Vec analysis

Word	Count	Percent	Category
taste	121447	30.86	Taste
sweet	42315	10.75	Taste
bitter	11424	2.9	Taste
sour	8252	2.1	Taste
salty	5688	1.45	Taste
sugary	1830	0.46	Taste
acidic	1194	0.3	Taste
savory	1075	0.27	Taste
metallic	527	0.13	Taste
umami	69	0.02	Taste
light	27329	6.94	Texture
hard	23253	5.91	Texture
rough	22058	5.6	Texture
texture	15544	3.95	Texture
fine	13475	3.42	Texture
smooth	12998	3.3	Texture
dry	12500	3.18	Texture
soft	10871	2.76	Texture
lean	9714	2.47	Texture
thick	7028	1.79	Texture
crunchy	6492	1.65	Texture
moist	5318	1.35	Texture
chewy	5250	1.33	Texture
creamy	4742	1.2	Texture
consistency	4345	1.1	Texture
heavy	3451	0.88	Texture
firm	2547	0.65	Texture
wet	2541	0.65	Texture
airy	2517	0.64	Texture
tough	1960	0.5	Texture
tender	1730	0.44	Texture
dense	1681	0.43	Texture
gummy	1666	0.42	Texture
delicate	1597	0.41	Texture
oily	1557	0.4	Texture
greasy	1552	0.39	Texture
watery	1545	0.39	Texture

mushy	1264	0.32	Texture
sounds	1234	0.31	Texture
crumbly	1017	0.26	Texture
grainy	1006	0.26	Texture
soggy	980	0.25	Texture
fluffy	978	0.25	Texture
gooey	971	0.25	Texture
melts	930	0.24	Texture
gritty	924	0.23	Texture
juicy	772	0.2	Texture
chunky	763	0.19	Texture
runny	739	0.19	Texture
silky	618	0.16	Texture
brittle	531	0.13	Texture
chalky	508	0.13	Texture
glue	505	0.13	Texture
powdery	462	0.12	Texture
rubbery	427	0.11	Texture
biting	419	0.11	Texture
slimy	347	0.09	Texture
flaky	343	0.09	Texture
noise	325	0.08	Texture
lumpy	317	0.08	Texture
pasty	221	0.06	Texture
crispness	154	0.04	Texture
slippery	148	0.04	Texture
stringy	141	0.04	Texture
puffy	137	0.03	Texture
slick	115	0.03	Texture
pulpy	51	0.01	Texture
binding	37	0.01	Texture
crackling	32	0.01	Texture
crispy	18	0	Texture
feathery	7	0	Texture
sponginess	3	0	Texture
squashy	1	0	Texture
order	52279	13.28	Customer Service
ship	26640	6.77	Customer Service
arrived	12368	3.14	Customer Service

service	5613	1.43	Customer Service
condition	5264	1.34	Customer Service
seller	4346	1.1	Customer Service
broken	3776	0.96	Customer Service
vendor	2018	0.51	Customer Service
refund	1994	0.51	Customer Service
promptly	1543	0.39	Customer Service
damaged	1497	0.38	Customer Service
timely	1369	0.35	Customer Service
expired	722	0.18	Customer Service
refunded	393	0.1	Customer Service
fulfilled	159	0.04	Customer Service
price	53327	13.55	Price
deal	12259	3.11	Price
cheaper	9481	2.41	Price
value	7506	1.91	Price
sale	5053	1.28	Price
pricey	3343	0.85	Price
offers	1630	0.41	Price
discount	1623	0.41	Price
bargain	1461	0.37	Price
pricing	1131	0.29	Price
savings	986	0.25	Price
healthy	19674	5	Health
organic	16239	4.13	Health
diet	13514	3.43	Health
healthier	4375	1.11	Health
safe	3941	1	Health
vitamins	2665	0.68	Health
nutritious	2027	0.52	Health
nutrients	1252	0.32	Health
wholesome	984	0.25	Health
unhealthy	844	0.21	Health
harmful	499	0.13	Health
smell	17457	4.44	Smell
aroma	6724	1.71	Smell
scent	3232	0.82	Smell
fruity	2045	0.52	Smell
odor	1803	0.46	Smell

burnt	1410	0.36	Smell
jasmine	982	0.25	Smell
fragrance	892	0.23	Smell
smoky	743	0.19	Smell
minty	646	0.16	Smell
chocolaty	496	0.13	Smell
chamomile	490	0.12	Smell
spearmint	485	0.12	Smell
spicy	6945	1.76	Trigeminal
bland	3889	0.99	Trigeminal
kick	3868	0.98	Trigeminal
tabasco	572	0.15	Trigeminal
spiciness	395	0.1	Trigeminal

Supplemental Table 2: Sweet phrases

Phrase	Count	Percent	Type
not very sweet	163	0.04	Under
not sweet enough	94	0.02	Under
not sweet enough for me	11	0	Under
too sweet	3121	0.79	Over
very sweet	1297	0.33	Over
to sweet	1020	0.26	Over
overly sweet	616	0.16	Over
so sweet	378	0.1	Over
really sweet	230	0.06	Over
sickly sweet	141	0.04	Over
sickeningly sweet	114	0.03	Over
cloyingly sweet	108	0.03	Over
pretty sweet	106	0.03	Over
overwhelmingly sweet	38	0.01	Over
excessively sweet	21	0.01	Over
overpoweringly sweet	20	0.01	Over
disgustingly sweet	13	0	Over
ultra sweet	6	0	Over
couldn't stand how sweet it was	1	0	Over
not too sweet	2550	0.65	Neutral
slightly sweet	855	0.22	Neutral
not overly sweet	700	0.18	Neutral
sweet enough	644	0.16	Neutral
on the sweet side	212	0.05	Neutral
not to sweet	135	0.03	Neutral
not so sweet	83	0.02	Neutral
sweet enough for me	36	0.01	Neutral
not overwhelmingly sweet	28	0.01	Neutral
not cloyingly sweet	27	0.01	Neutral
not sickly sweet	26	0.01	Neutral
not sickeningly sweet	22	0.01	Neutral
not overpoweringly sweet	18	0	Neutral
not really sweet	15	0	Neutral
not excessively sweet	10	0	Neutral
not disgustingly sweet	4	0	Neutral
sweet as i would prefer	2	0	Neutral
not ultra sweet	2	0	Neutral

bioRxiv preprint doi: https://doi.org/10.1101/662585; this version posted June 7, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

not on the sweet side	1	0	Neutral

The count refers to the number of reviews that contain the relevant phrase and the percent of all reviews that contain that phrase. Type refers to undersweetness, oversweetness or neutral comments.