Published in final edited form as:

Food Qual Prefer. 2012 October 1; 26(1): 128–133. doi:10.1016/j.foodqual.2012.04.009.

Rejection Thresholds in Chocolate Milk: Evidence for Segmentation

Meriel L. Harwood^{1,2}, Gregory R. Ziegler², and John E. Hayes^{1,2,*}

¹Sensory Evaluation Center, College of Agricultural Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA

²Department of Food Science, College of Agricultural Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA

Abstract

Bitterness is generally considered a negative attribute in food, yet many individuals enjoy some bitterness in products like coffee or chocolate. In chocolate, bitterness arises from naturally occurring alkaloids and phenolics found in cacao. Fermentation and roasting help develop typical chocolate flavor and reduce the intense bitterness of raw cacao by modifying these bitter compounds. As it becomes increasingly common to fortify chocolate with 'raw' cacao to increase the amount of healthful phytonutrients, it is important to identify the point at which the concentration of bitter compounds becomes objectionable, even to those who enjoy some bitterness. Classical threshold methods focus on the presence or absence of a sensation rather than acceptability or hedonics. A new alternative, the rejection threshold, was recently described in the literature. Here, we sought to quantify and compare differences in Rejection Thresholds (RjT) and Detection Thresholds (DT) in chocolate milk spiked with a food safe bitterant (sucrose octaacetate). In experiment 1, a series of paired preference tests was used to estimate the RjT for bitterness in chocolate milk. In a new group of participants (experiment 2), we determined the RjT and DT using the forced choice ascending method of limits. In both studies, participants were segmented on the basis of self-declared preference for milk or dark solid chocolate. Based on sigmoid fits of the indifference-preference function, the RjT was ~2.3 times higher for those preferring dark chocolate than the RjT for those preferring milk chocolate in both experiments. In contrast, the DT for both groups was functionally identical, suggesting that differential effects of bitterness on liking of chocolate products are not based on the ability to detect bitterness in these products.

Keywords

bitterness; consumer rejection threshold; food preference; psychophysics; methodology

1. Introduction

Bitterness is generally considered a negative attribute in food, yet many individuals enjoy some amount of bitterness in products such as coffee, beer, or dark chocolate. In chocolate,

© 2012 Elsevier Ltd. All rights reserved

*Corresponding author. Tel.: +1-814-863-7129; jeh40@psu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

bitterness arises from naturally occurring alkaloids and phenolic compounds. Cacao undergoes many processing steps to develop the typical flavor and aroma of chocolate (Hoskin 1994; Adriaenssens 2010). Many of the crucial steps in processing, such as fermentation and heat treatments (e.g. roasting, conching) significantly reduce the intense bitterness of 'raw' cacao by transforming a large portion of the bitter compounds. The standard processing that most chocolate undergoes today is a result of consumers' sensitivity to and rejection of highly bitter products. Multiple studies suggest taste is the primary driver of consumers' food selection behaviors, eclipsing cost, convenience and even potential health value (e.g. Glanz, Basil et al. 1998; IFIC 2011).

Notably, many of the aversive bitter compounds found in chocolate have also been studied for their apparent cardiovascular health benefits (e.g. Corti, Flammer et al. 2009; Visioli, Bernaert et al. 2009; Djousse, Hopkins et al. 2011) – cacao is abundant in potentially healthy compounds (e.g. the monomer epicatechin (flavan-3-ols) and its polymer, procyanidin) that elicit very bitter and also often astringent sensations (Drewnowski and Gomez-Carneros 2000; Stark, Bareuther et al. 2006). However, the same processing that cacao undergoes to develop its well-loved flavor and aroma destroys up to 85% of these compounds (Visioli, Bernaert et al. 2009). For this reason, some manufacturers are working to develop processes to make chocolate that preserve more of these compounds, or alter formulations in an effort to boost the amount of these compounds in the finished product (Visioli, Bernaert et al. 2009). For example, Kealey and colleagues (Kealey, Snyder et al. 2011) at Mars recently patented processes and formulations that preserve polyphenol content in chocolate. This provides a challenge in formulation, however, as these changes may significantly alter the sensory profile of the product in a way that is aversive to the consumer. To meet the demands of consumers for health enhancing products, it will be important to identify the point at which the concentration of bitter compounds in chocolate becomes objectionable, even to those who enjoy some bitterness.

The measurement of detection thresholds dates back to Fechner and the dawn of psychophysics 150 years ago (Engen 1972; Duffy, Hayes et al. 2009). Unfortunately, detection thresholds often fail to predict suprathreshold intensity (e.g. Bartoshuk 1978; Frijters 1978) or food liking (e.g., Duffy, Peterson et al. 2004; Lucas, Riddell et al. 2011). In the case of off flavors (taints) or for attributes where only low levels are desired, it is more critical to determine the point at which liking is adversely affected. The question becomes not can the consumer sense the attribute, but do they care? Recently, Prescott and colleagues elegantly adapted classical threshold methods to directly address the question of acceptability of cork taint (TCA) in white wine (Prescott, Norris et al. 2005). Subsequently, this approach has been applied to eucalyptol in red wine (Saliba, Bullock et al. 2009), caffeine in coffee, citric acid in orange juice, and salt in beef broth (Lee, Prescott et al. 2008). Notably, rejection never reached 100% in any of these studies. Also, the statistical analysis strategies in these prior studies have been based on linear or segmental fits and binomial distributions tables for preference tests (Prescott, Norris et al. 2005; Lee, Prescott et al. 2008; Saliba, Bullock et al. 2009).

Here, we sought to quantify and compare differences in Rejection Thresholds (RjT) and Detection Thresholds (DT) across individuals who differed in their preference for solid chocolate (dark chocolate versus milk chocolate). The present study aimed to answer three questions. First, was it possible to reach universal rejection? Second, would solid chocolate preferences generalize to flavored fluid milk? Third, if segmentation was present in fluid milk, could it be explained by differences in bitterness detection thresholds? Finally, we sought to use a sigmoid fit in place of a linear fit to be consistent with the expected shape of the underlying psychometric function.

2. Methods

2.1. Overview of Methods

Chocolate milk was chosen as the model system, and sucrose octaacetate (SOA) was selected to be the model bitter compound. SOA is a highly bitter acetylated sugar that is generally recognized as safe (GRAS) by the FDA as a direct and indirect food additive (21 CFR 172.515 and 21 CFR 175.105, respectively), and it shows little or no evidence of genetic variation in humans (Boughter and Whitney 1993; Hansen, Reed et al. 2006). Compusense *five* software v5.2 (Guelph ONT) was used for sample randomization and data collection. All tests occurred in individualized test booths in the Sensory Evaluation Center in the Department of Food Science at Penn State. Procedures were exempted from Institutional Review Board review by the Penn State Office of Research Protections under the wholesome foods/approved food additives exemption in 45 CFR 46.101(b)(6). Participants provided informed consent and were paid for their time. In experiment one, we determined group rejection thresholds in 55 participants in a single session. In experiment two, we recruited a new group of participants and determined detection thresholds and rejection thresholds on separate days (80 recruits; 70 completers).

2.2 Experiment One

2.2.1. Participants—Fifty-five reportedly healthy, non-smoking individuals were recruited from the campus of and community around the Pennsylvania State University (State College, PA) via email for their willingness to consume chocolate milk. Forty-seven women and eight men, aged between 18 – 45 years participated in the study. Of these participants, 24 endorsed preferring dark chocolate and 31 endorsed preferring milk chocolate when asked about their preference when eating solid chocolate at the conclusion of the test session.

2.2.2. Stimuli—The chocolate milk samples were spiked with varying concentrations of SOA: 0 (blank), 3.2, 5.6, 10, 18, 32, and 56 μM . These concentrations were selected based on published threshold values for SOA in water (Boughter and Whitney 1993), which ranged from 0.25 – 16 μM , and adjusted for milk via informal piloting at the lab bench. Each spiked sample was paired with a blank containing only chocolate milk. Samples (10mL) were pre-poured in small clear plastic soufflé cups, kept refrigerated until testing (4C), and served cold.

2.2.3. Procedure—This first test was a two-alternative forced choice (2-AFC) preference test carried out according to the American Society for Testing and Materials (ASTM) method E-2263 (ASTM 2004). In this test, six pairs of samples were presented to each participant, with each pair containing a spiked and blank sample. The pairs were presented in order of ascending concentration, with three pairs on the first tray, and 3 pairs on the second tray. Within a pair, the order of the samples was randomized. The participants were asked to move through the sample pairs one at a time by tasting both of the samples in the pair and then indicating which of the samples in that pair that they preferred by clicking the appropriate button on the computer screen before moving on to the next pair of stimuli.

A 'no preference' option was not provided for two reasons: recent work suggests a 'no preference' option does not clarify an ambiguous 50/50 result as was previously argued (see Chapman & Lawless, 2005), and it would have complicated the present analysis unnecessarily. In contrast to earlier 'identical sample' studies on the incidence of no preference, the rejection threshold model covers the entire range of the response function, beginning with samples that are essentially identical (e.g. below detection threshold) before progressing to samples that are obviously different. For more information on the recent

debate over use of a 'no preference' option, the interested reader is referred to (Marchisano, Lim et al. 2003; Chapman and Lawless 2005; Chapman, Lovelace et al. 2010; Ennis and Ennis 2012).

2.3 Experiment Two

2.3.1. Participants—Eighty reportedly healthy, non-smoking chocolate consumers aged 18–45 were recruited as above (of the eighty, twenty five had participated in Experiment 1). Fifty-six women and twenty-four men participated in the detection threshold portion of the experiment (2A). Of these participants, 39 endorsed preferring dark chocolate and 41 endorsed preferring milk chocolate when asked about their preference when eating solid chocolate. The participants in this experiment were randomly assigned to two different groups, stratifying for chocolate preference (Group A, total $n = 40$, milk $n = 20$, dark $n = 20$; and Group B, total $n = 40$, milk $n = 21$, dark $n = 19$).

Seventy of these individuals (52 women and 18 men) returned on another day for determination of SOA rejection thresholds in chocolate milk (Experiment 2B). Of these, 38 endorsed preferring dark chocolate and 32 endorsed preferring milk chocolate.

2.3.2. Stimuli—For experiment 2A (detection thresholds), the chocolate milk samples were spiked with increasing concentrations of SOA: 0.056, 0.18, 0.56, 1.8, 5.6, and 18 μM . For experiment 2B (rejection thresholds), the concentrations were revised based on Experiment 1, resulting in 1, 2.1, 4.6, 10, 21, and 46 μM samples. Each spiked sample was paired with a blank sample containing only chocolate milk. Ten mL samples were prepared and presented as above.

2.3.3. Procedure—In Experiment 2A, detection thresholds were determined in accordance with ASTM method E-679 Forced Choice Ascending Method of Limits (ASTM 2011). Briefly, participants execute a series of triangle tests, where every triad of samples contains one spiked sample and two blank samples. The triads were presented in order of ascending concentration, and the order of samples within a triad was randomized. To limit fatigue, the participants were randomized into two groups, where each group tasted three of the six concentrations, as described in E-679. Therefore, each participant tasted three triads of samples total. Group A tasted the concentrations 0.056, 0.56 and 5.6 μM , while Group B tasted the concentrations 0.18, 1.8 and 18 μM . The participants were asked to taste each of the samples in the triad and indicate of the three samples tasted, which of the samples was different before moving on to the next triad. If the participants were unsure of their answer, they were forced to choose a sample; degree of certainty was not assessed. For Experiment 2B (rejection thresholds), the procedure was identical to that used in Experiment 1.

We considered using a 2-AFC task instead of a triangle test, as it is a more powerful method able to find smaller differences at the same number of participants (Ennis 1993). However, after much discussion, we chose to use a triangle here instead of a 2-AFC to avoid having to specify a particular attribute, given the complex matrix we were working with. Specifically, adding SOA might alter the chocolate milk samples in attributes beyond bitterness. For example, perithreshold levels of the bitterant might reduce sweetness via mixture suppression (Lawless 1979; Hayes, Wallace et al. 2011), or increase perceived cocoa flavor via learned associations (e.g. Keast 2008), without increasing perceived bitterness. Thus, we felt a 2-AFC was too narrow a task for this exploratory work, in spite of the additional power and sensitivity it provides over a triangle test.

2.4 Statistical Analysis

To determine the group rejection thresholds (RjT) in experiment 1 and 2B, we adapted the method described by Lawless (2010). That is, we plotted the proportion of responses (i.e., preference for the blank) against concentration and fit a curve through the points. The point at which performance exceeds chance by 50% can then be determined via Abbott's formula (Lawless 2010), resulting in chance corrected probabilities of 75% for 2AFC tasks. Instead of using OLS regression or a logit model as recommended by Lawless (2010), we chose to use a sigmoid model, as we expected preference to oscillate around chance before climbing steeply to an asymptote near universal preference. Specifically, we fit a sigmoid variable slope dose-response function using the following four parameter logistic equation:

$$Y = \text{Min} + \left[\frac{(\text{Max} - \text{Min})}{1 + 10^{(\text{LogRjT}_{50} - X) \times \text{HillSlope}}} \right]$$

Adapted from the Hill equation commonly used in pharmacology, this model describes the top of the curve (max), the bottom of the curve (min), the spot halfway between min and max (classically EC₅₀; called RjT₅₀ here), and the slope of the curve (the Hill coefficient). Notably, the EC₅₀ in the traditional Hill equation corresponds to the traditional definition of a sensory threshold used in psychophysics (Engen 1972).

Here, we fit the rejection functions using GraphPad Prism 5.0C for OSX (GraphPad Software, San Diego CA) and RjT₅₀ values were calculated directly in software. In experiment 1 and 2B, we compared the rejection thresholds between those preferring milk chocolate and those preferring dark chocolate via the 'compare two data sets' option in Prism, which generates an F-statistic and p-value for the difference between the EC₅₀ values. Prism generates this statistic by fitting the data twice: once for each group separately (two curves), and once for a global fit (i.e. a common EC₅₀) across all of the data. Because the global model is nested within the two group (separate curve) model, the two models can be tested with a standard extra sum-of-squares F test (Motulsky and Christopoulos 2004). The equation for this F value is:

$$F = \frac{(SS_{\text{global}} - SS_{\text{2curve}}) / (DF_{\text{global}} - DF_{\text{2curve}})}{SS_{\text{2curve}} / DF_{\text{2curve}}}$$

In experiment 1, the dose range did not include chance performance (near 50%) for one group, so we constrained the fit in Prism and refit the curve (described below). Finally, we used the ECx feature in Prism to determine the RjT₂₀ value for both preference segments (i.e., the threshold where performance was 20% above chance.)

To determine the group detection thresholds in Experiment 2A, we again adapted the method of Lawless (2010). The proportion of individuals correctly identifying the odd sample in each triangle test was plotted against concentration. For this, we used linear regression to fit the data, after excluding the lower concentrations that oscillated around chance. The group detection threshold was defined as the concentration at which the chance corrected proportion was 66% (i.e., halfway between chance (0.33) and perfect performance (1.0)).

3. Results

3.1 Experiment One

The rejection threshold (RjT) reflects the concentration at which the participants preferred the un-spiked sample. As shown in Figure 1, as concentration increased, more participants chose the blank over the spiked sample, since bitterness is generally considered a negative attribute in food products. This was true here, as the participants unanimously rejected the spiked sample at the top two concentrations tested. Figure 1 also shows that the rejection thresholds for individuals who report preferring solid milk chocolate was noticeably lower than those who report preferring solid dark chocolate. Due to the lack of points near chance for the milk chocolate group, the original sigmoid fit had an ambiguous RjT₅₀ value (see Supplemental Figure 1). Therefore, we constrained two of the four parameters in the Hill equation, minimum and maximum, to their theoretically indicated values of 0.5 and 1.0, and refit the curve. When we did so, the rejection threshold for the dark chocolate group was 2.37 times higher [$F(1,8)=25.1$; $p=0.001$] than the milk chocolate group (9.35 μM versus 3.95 μM , respectively). As a cross check, we also analyzed these data using OLS regression and determined the group thresholds at 75% of chance via the Lawless (2010) method; the values (not shown) were not substantially different from the results above.

3.2 Experiment Two

As shown in Figure 2, the group detection thresholds (DTs) were functionally identical across both groups (5.86 μM versus 5.38 μM for dark and milk chocolate groups, respectively). Regarding rejection thresholds, Figure 3 shows that the participants almost unanimously rejected the spiked sample at the highest concentration tested. Extension of the concentration range downward relative to experiment 1 resulted in proportions near the theoretical minimum and maximum (0.5 and 1.0, respectively) at either end of the curve, allowing for an unconstrained fit. Similar to Experiment 1, the rejection thresholds for individuals who report preferring dark chocolate were 2.3 times higher [$F(1,4)=11.90$; $p=0.026$] than those preferring milk chocolate. For completeness, the constrained and unconstrained RjTs from experiments 1 and 2 are summarized in Table 1.

4. Discussion

The present work extends prior findings in several ways. First, we were able to demonstrate that it is possible to reach universal rejection (right side of Figures 1 and 3). We were also able to see clear segmentation based on preference for solid chocolate despite the fact that we were working with a model system of chocolate milk. This indicates that solid chocolate preference may generalize to preferences for other chocolate products. For those preferring dark chocolate, the rejection threshold was approximately 2.3 times higher than the rejection threshold for those preferring milk chocolate. In contrast, the detection threshold (DT) for both groups was functionally identical. This suggests that tolerance for bitterness in chocolate products is not based on differential ability to detect bitterness in those products. Finally, we were able to describe the dose-response function for bitterness in chocolate milk using a sigmoid model as opposed to the linear fits that had been used previously.

Although experiments 1 and 2 contained a few of the same participants, the consistency of the data across both experiments suggests these effects are robust and may generalize to other populations. In experiment 2B, we were able to clearly see the sigmoid shape of the entire function with the proportion preferring the control starting near chance, and plateauing once the concentration reached a level that was objectionable to all (unlike prior work on rejection thresholds). Notably, while the detection thresholds from experiment 2 were slightly higher than the RjT for the milk chocolate groups in experiments 1 and 2, they

were not significantly different as the DTs fell within the 95% confidence intervals for the milk chocolate group RjTs. In contrast, within each experiment, the dark chocolate group RjTs were significantly higher than those for the milk chocolate group. This result suggests that market segmentation in solid chocolate preference is not driven by differences in perception.

Use of a four parameter logistic equation (the Hill equation) to describe the group rejection function has multiple advantages over prior attempts to model rejection thresholds. First, unlike earlier examples that use binomial tables to determine where to put the rejection threshold (Prescott, Norris et al. 2005; Lee, Prescott et al. 2008; Saliba, Bullock et al. 2009), the cutoff of interest here does not change with the number of participants, facilitating comparisons across groups of different size. Also, this neatly avoids the issue that binomial models may not accurately reflect human choice behavior (see Chapman & Lawless, 2005). Second, it is theoretically attractive, as the expected psychometric function is a sigmoid, not linear. The use of a sigmoid captures the entirety of the relationship, where at levels around the detection threshold, we expect to see preference oscillate around chance (due to momentary differences in perception (see Chapman et al., 2010)), then climb steadily upwards as more people begin to reject the spiked sample at higher concentrations and eventually plateau at universal rejection. Third, using the Hill equation precludes the need to adjust for chance via Abbott's formula, as the definition of EC_{50} is the halfway point between chance and perfect performance (e.g. the classical definition of a threshold (Engen 1972)). Fourth, it appears to be robust to small errors in dose-finding (e.g. Experiment 1) as the minimum and maximum values are easily constrained to their theoretical values. Fifth, this type of model allows for the calculation of an 'acceptable concentration limit' (ACL) for any desired percentage of the respondents (i.e., RjT_x / EC_x). For example, based on the results of experiment 2B, the RjT_{20} for the milk chocolate preferring group is 2.175 μM and for the dark chocolate preferring group is 4.325 μM . This may be an advantage in quality assurance or brand management situations where one might want to be more conservative than a binomial distribution table, by setting the proportion of participants rejecting a product at a lower value.

This technique is most applicable for attributes that show a monotonic liking function. Here, we show a typical example of this by using bitterness, which shows a sigmoid relationship as a function of concentration. In contrast, this method may not work for stimuli that have a clear optimum. When considering an attribute with an optimum, one could plot the proportion preferring the spiked sample, which would result (initially) in a curve similar to those seen here. However, once the optimum was exceeded, the proportion preferring the spike would drop back toward chance as the respondents then switch their preference back to the blank sample. While it would be possible to analyze the data up to the point of optimum by treating the optimum as a local maximum, difficulty in fitting the complex curve shape would limit the utility of this approach, especially if the drop was rapid, resulting in an asymmetric curve. Thus, this approach is best suited to negative attributes whose psychohedonic (concentration-liking; (Hayes and Duffy 2008)) function is monotonic. Notably, this method could also be used for variables other than physical concentration, where one might also expect to see a monotonic relationship. For example, it might be possible to test for shelf life, with time on the x-axis.

The concept of a preference threshold has been used in animal psychophysics since the 1930's (e.g. (Richter 1936)). However, the rejection threshold method here is quite different in spite of the similar name and goals. In a preference threshold task, an animal is allowed to feed from two bottles. When faced with two bottles that are indistinguishable, the animals will drink from both bottles equally. This is similar to what we see as respondents choose a sample randomly when the levels of the spike are below the detection threshold. In a

preference task, the animal is then presented an ascending series of concentrations paired against a blank. At the end of each session, the amount consumed from each bottle is measured, and relative preference is *inferred* from the amount consumed. Thus, the preference threshold is defined as the concentration at which the animal consumes significantly more of one solution than the other (Richter 1936). However, this method implicitly conflates the ability to tell a difference with a preference (e.g. Fregly 1973). That is, an animal (or nonverbal humans) could conceivably be able to distinguish between stimuli while remaining indifferent. In adult humans, we have the luxury of defining the task more narrowly, as was done here. Thus, while the two approaches have similar goals, they have different underlying assumptions. Finally, while prior reports have referred to the present technique as a 'consumer rejection threshold', we suggest dropping the prefix consumer in the future, as the technique need not be limited to consumer-oriented research. For example, it could be used to characterize different hedonic phenotypes in genetic research (e.g. Duffy, Hayes et al. 2009).

4.1 Limitations and Conclusions

Here, we used a triangle test method to determine the group detection threshold using ASTM method E-679, in spite of the fact that n-AFC tasks are more powerful than triangle tests (Ennis 1993). The triangle method and the n-AFC methods require different cognitive strategies on the part of the participant: in a triangle test, the participant must pick out the odd sample, whereas in the n-AFC approach, participants can skim across the samples, seeking to identify the weakest or strongest sample. Thus, the n-AFC approach is an easier cognitive task, which makes it more sensitive with the same number of participants and samples. (Astute readers may note ASTM E-679 refers to 3-AFC in places; however, this refers solely to the sample presentation – one test and two blanks – and not the cognitive strategy to be used. The E-679 method explicitly indicates panelists should identify the *different* sample.) As noted by an anonymous reviewer, the rejection threshold task used here is much more analogous to a directed n-AFC task than a triangle test, so it may have been more appropriate to determine our detection thresholds using a 2-AFC or 3-AFC task for a more direct comparison. When designing our study, we decided against this experimental approach for the reasons outlined in section 2.2.3, but we agree that additional work is needed to explore how the results might differ with an alternative detection threshold method. Because the group detection threshold may be lower with a more sensitive task, present results may be overly conservative, underestimating the gap between detection and rejection thresholds. In addition, more work is needed to determine whether inclusion of a 'no preference' option alters rejection threshold estimates.

Another potential limitation to the study is the study population. The sample was limited to consumers from central Pennsylvania, the majority of whom were women. More work is needed to explore any potential sex differences. Also, SOA is not an endogenous compound found within chocolate. With regard to validation of the method itself, this is not an issue, but inferences about the absence of perceptual differences in chocolate bitterness should not be made from present data; it remains possible that endogenous bitter compounds found in chocolate could exhibit genetic variation in perception (e.g., Hayes, Wallace et al. 2011). Here, SOA was chosen for its ease of acquisition and use, potency, and GRAS status. We also note that "which do you prefer" and "which do you reject" are two different cognitive tasks. Here, we only asked our participants which sample they preferred, to be consistent with prior studies. Additional work is needed to verify that the curves obtained under both tasks would be symmetric. Finally, further work is also needed to see if this method will be effective for quantifying segmentation in solid products, as all work to date on this method have occurred in liquid food systems. In conclusion, we demonstrated that rejection thresholds can be used to study market segmentation and that the differences that we

observed in rejection thresholds were not due to differences in the ability of our participants to detect the bitter compound.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This manuscript was completed in partial fulfillment of the requirements for a Master of Science degree at the Pennsylvania State University by MLH. The authors thank Dr. David DePasquale and the Sensory Evaluation Center staff for their assistance with data collection; we also thank our study participants for their time and participation.

Funding MLH is supported by funds from the Pennsylvania Manufacturing Confectioners' Association (PMCA) and JEH and GRZ receive support from the Pennsylvania State University. JEH received additional support from National Institutes of Health grant number DC010904.

References

- Adriaenssens M. Impact of Ingredients on Chocolate Flavor. *The Manufacturing confectioner*. 2010; 90(9):87.
- ASTM. Standard E-2263: Paired Preference Test. ASTM International; West Conshohocken, PA: 2004.
- ASTM. Standard E-679: Forced Choice Ascending Method of Limits. ASTM International; West Conshohocken, PA: 2011.
- Bartoshuk LM. The psychophysics of taste. *Am J Clin Nutr*. 1978; 31(6):1068–1077. [PubMed: 352127]
- Boughter JD, Whitney G. Human taste thresholds for sucrose octaacetate. *Chemical Senses*. 1993; 18(4):445–448.
- Chapman KW, Lawless HT. Sources Of Error And The No-Preference Option In Dairy Product Testing. *Journal of Sensory Studies*. 2005; 20(5):454–446.
- Chapman KW, Lovelace E, et al. Preference for One of Two Identical Stimuli: Expectations, Explicit Instructions and Personal Traits. *Journal of Sensory Studies*. 2010; 25:35–53.
- Corti R, Flammer AJ, et al. Cocoa and cardiovascular health. *Circulation*. 2009; 119(10):1433–1441. [PubMed: 19289648]
- Djousse L, Hopkins PN, et al. Chocolate consumption is inversely associated with prevalent coronary heart disease: The National Heart, Lung, and Blood Institute Family Heart Study. *Clinical Nutrition*. 2011; 30(2):182–187. [PubMed: 20858571]
- Drewnowski A, Gomez-Carneros C. Bitter taste, phytonutrients, and the consumer: a review. *The American journal of clinical nutrition*. 2000; 72(6):1424–1435. [PubMed: 11101467]
- Duffy, VB.; Hayes, JE., et al. Taste: Vertebrate Psychophysics. In: Larry, RS., editor. *Encyclopedia of Neuroscience*. Academic Press; Oxford: 2009. p. 881-886.
- Duffy VB, Hayes JE, et al. Surveying food and beverage liking: a tool for epidemiological studies to connect chemosensation with health outcomes. *Ann N Y Acad Sci*. 2009; 1170:558–568. [PubMed: 19686193]
- Duffy VB, Peterson JM, et al. Associations between taste genetics, oral sensation and alcohol intake. *Physiology & Behavior*. 2004; 82(2–3):435–445. [PubMed: 15276808]
- Engen, T. Psychophysics. In: Kling, JW.; Riggs, LA., editors. *Woodworth & Schlosberg's Experimental Psychology*. Rinehart and Winston, Inc.; New York, Holt: 1972.
- Ennis DM. The Power Of Sensory Discrimination Methods. *Journal of Sensory Studies*. 1993; 8(4): 353–370.
- Ennis DM, Ennis JM. Accounting for no difference/preference responses or ties in choice experiments. *Food Quality and Preference*. 2012; 23(1):13–17.

- Fregly MJ. Effect of an oral contraceptive on NaCl appetite and preference threshold in rats. *Pharmacol Biochem Behav.* 1973; 1(1):61–65. [PubMed: 4775584]
- Frijters JER. Critical Analysis of Odor Unit Number and Its Use. *Chemical Senses & Flavour.* 1978; 3(2):227–233.
- Glanz K, Basil M, et al. Why Americans Eat What They Do: Taste, Nutrition, Cost, Convenience, and Weight Control Concerns as Influences on Food Consumption. *Journal of the American Dietetic Association.* 1998; 98(10):1118–1126. [PubMed: 9787717]
- Hansen JL, Reed DR, et al. Heritability and genetic covariation of sensitivity to PROP, SOA, quinine HCl, and caffeine. *Chem Senses.* 2006; 31(5):403–413. [PubMed: 16527870]
- Hayes JE, Duffy VB. Oral sensory phenotype identifies level of sugar and fat required for maximal liking. *Physiol Behav.* 2008; 95(1–2):77–87. [PubMed: 18538361]
- Hayes JE, Wallace MR, et al. Allelic Variation in TAS2R Bitter Receptor Genes Associates with Variation in Sensations from and Ingestive Behaviors toward Common Bitter Beverages in Adults. *Chem Senses.* 2011; 36(3):311–319. [PubMed: 21163912]
- Hoskin JC. Sensory Properties of Chocolate and Their Development. *American Journal of Clinical Nutrition.* 1994; 60(6):S1068–S1070.
- IFIC. Food & Health Survey: 452 Consumer Attitudes Toward Food Safety, Nutrition & Health. International Food Information Council Foundation; Washington, DC: 2011.
- Kealey, KS.; Snyder, RM., et al. Chocolates and Chocolate Liquor Having an Enhanced Polyphenol Content. U. S. Patent. USA, Mars, Incorporated. **US 7,968,140 B2** 2011. p. 1-40.
- Keast RS. Modification of the bitterness of caffeine. *Food Qual Pref.* 2008; 19(5):465–472.
- Lawless HT. Evidence for neural inhibition in bittersweet taste mixtures. *Journal of Comparative & Physiological Psychology.* 1979; 93(3):538–547. [PubMed: 479396]
- Lawless HT. A simple alternative analysis for threshold data determined by ascending forced-choice methods of limits. *Journal of Sensory Studies.* 2010; 25(3):332–346.
- Lee YM, Prescott J, et al. PROP Taster Status and the Rejection of Foods with Added Tastants. *Food Science and Biotechnology.* 2008; 17(5):1066–1073.
- Lucas L, Riddell L, et al. The Influence of Sodium on Liking and Consumption of Salty Food. *Journal of Food Science.* 2011; 76(1):S72–S76. [PubMed: 21535718]
- Marchisano C, Lim JY, et al. Consumers report preferences when they should not: A cross-cultural study. *Journal of Sensory Studies.* 2003; 18(6):487–516.
- Motulsky, H.; Christopoulos, A. Fitting models to biological data using linear and nonlinear regression : a practical guide to curve fitting. Oxford University Press; Oxford; New York: 2004.
- Prescott J, Norris L, et al. Estimating a “consumer rejection threshold” for cork taint in white wine. *Food Quality and Preference.* 2005; 16(4):345–349.
- Richter CP. Increased Salt Appetite in Adrenalectomized Rats. *American Journal of Physiology.* 1936; 115(1):155–161.
- Saliba AJ, Bullock J, et al. Consumer rejection threshold for 1,8-cineole (eucalyptol) in Australian red wine. *Food Quality and Preference.* 2009; 20(7):500–504.
- Stark T, Bareuther S, et al. Molecular definition of the taste of roasted cocoa nibs (*Theobroma cacao*) by means of quantitative studies and sensory experiments. *Journal of Agricultural and Food Chemistry.* 2006; 54(15):5530–5539. [PubMed: 16848542]
- Visioli F, Bernaert H, et al. Chocolate, Lifestyle, and Health. *Critical Reviews in Food Science and Nutrition.* 2009; 49(4):299–312. [PubMed: 19234942]

Highlights

- A new psychophysical method, the rejection threshold, was recently described
- We refine the models used in this approach by applying sigmoidal curve fitting
- We demonstrate attributes can reach universal rejection in contrast to earlier work
- This method successfully characterized different market segments for chocolate milk
- In spite of clear segmentation, preference was not driven by perceptual differences

Figure 1.

Figure 2.

Figure 3.

Table 1

Rejection Threshold Values for Each Group- from experiments 1 and 2B

RjT50				
	Milk Chocolate Group	Dark Chocolate Group	Fold Difference	p-value
Experiment 1 (Constrained Fit)	3.95	9.35	2.37	0.001
Experiment 2B (Unconstrained Fit)	3.90	9.00	2.30	0.026