Published in final edited form as:
Eat Behav. 2012 January ; 13(1): 42–45. doi:10.1016/j.eatbeh.2011.09.003.

Reliability of a common solution-based taste perception test: implications for validity and a briefer test

S.M. Coulon[a,1], **A. Chellino**[a], **J.M. Reed**[a], and **C.K. Martin**[a]

[a]Pennington Biomedical Research Center, Baton Rouge, LA, 70808, U.S.A

Abstract

The aim of this study was to assess test-retest reliability of a common method for quantifying taste perception and its association with gustatory responses and individual risk for obesity and related health conditions. Forty-six healthy adults rated 20 mixtures comprised of 5 dairy beverages varied in fat content and mixed with sugar concentrations of 0%, 5%, 10%, and 20%, following existing procedures. Individuals rated the sweetness, creaminess, and pleasantness of each mixture during two taste testing sessions occurring 7±2 days apart. Test-retest correlations were of the expected magnitudes ($r \geq .50$) only for the pleasantness ratings of mixtures with higher sugar concentrations. Correlations for sweetness and creaminess taste perception ratings were low, indicating that such ratings may not be reliable over approximately one week, and challenging the validity of such ratings for measuring trait taste perception. A shortened version of the test may be warranted.

Keywords

taste perception; taste preference; reliability; gustatory; sweet; creamy; pleasant

1. Introduction

Investigating psychophysical factors that may contribute to the development of obese and overweight conditions is a high scientific priority (Flegal, Carroll, Ogden, & Curtin, 2010). The consumption of foods and beverages high in sugar and fat is positively associated with developing these conditions (Stubbs, Ferres, & Horgan, 2000), and biological taste perception of sugar and fat varies widely among individuals (Reed & McDaniel, 2006). Taste perception refers to the reported set of oral tactile sensations an individual experiences in response to stimulation of fungiform taste papillae during food or beverage intake

Correspondence concerning this article should be addressed to Corby K. Martin, Pennington Biomedical Research Center, 6400 Perkins Road, Baton Rouge, LA 70808. Corby.Martin@pbrc.edu.
[1]Sandra M. Coulon is now at the Department of Psychology, University of South Carolina, Columbia, SC, U.S.A.

Contributors. Author S. Coulon managed laboratory resources for the study, administered the testing protocol to participants, conducted literature searches, and drafted and edited the manuscript. Author A. Chellino conducted literature searches and provided summaries of previous research studies, contributed to the development and editing of data tables, and edited the manuscript. Author J.M. Reed coordinated the study and administered the testing protocol to participants, and edited the manuscript. Author C.K. Martin designed the study and wrote the protocol, and edited the manuscript. Author B wrote the first draft of the manuscript and all authors contributed to and have approved the final manuscript.

Conflict of Interest. All authors declare that they have no conflicts of interest.

(Bradbury, 2004; Cowart, 1981). Taste perception is often associated with taste sensitivity, which typically refers to tactile acuity, as commonly characterized by sensitivity to substances such as 6-*n*-Propylthiouracil (PROP), though these terms often overlap in the literature (Hayes & Duffy, 2008; Lim, Urban, & Green, 2008; Salbe, DelParigi, Pratley, Drewnowski, & Tataranni, 2004). It is critical that methods for assessing such psychophysical factors are reliable, yet to our knowledge, reliability of a commonly-implemented, solution-based method for assessing taste perception remains unexplored.

Taste perception and other gustatory responses have been associated with consumption of foods high in sugar and fat, as well as phenotypic markers of obesity such as ability to taste PROP (Hayes & Duffy, 2007, 2008; Salbe, et al., 2004). However, the nature of these associations remains unclear and may vary by gender (Hayes & Duffy, 2008), PROP taster status (Hayes & Duffy, 2008;S. V. Kirkmeyer & Tepper, 2003; Salbe, et al., 2004; Yeomans, Tepper, Rietzschel, & Prescott, 2007), body weight (Salbe, et al., 2004), and density of fungiform papillae in the mouth (Hayes & Duffy, 2007, 2008). Additionally, some studies have failed to find associations among taste perception and consumption of foods high in sugar and fat, or phenotypic markers of obesity (Drewnowski, Henderson, Shore, & Barratt-Fornell, 1997; Drewnowski, Henderson, & Barratt-Fornell, 1998; Ly & Drewnowski, 2001). These inconclusive or discrepant findings may be explained in part by the psychometric properties (e.g., reliability) of the methods used to quantify taste perception.

The common, solution-based method involves asking participants to complete a taste test in which 20 beverage mixtures that vary in sugar and fat are sampled and rated on dimensions of sweetness, creaminess, and pleasantness (Drewnowski, 1997; Drewnowski, Brunzell, Sande, Iverius, & Greenwood, 1985; Drewnowski & Greenwood, 1983; Salbe, et al., 2004). Though the method has been associated with phenotypic markers of obesity, procedures have varied widely among studies, with taste tests including a range of 3–20 dairy mixtures or semi-solid dairy products that contain graded amounts of sugar and fats and with water mixtures sometimes used instead (Cooling & Blundell, 2001; Lim, et al., 2008; Ly & Drewnowski, 2001;J. A. Mennella, Pepino, Lehmann-Castor, & Yourshaw, 2010; Yeomans, et al., 2007). For example, taste ratings from a 15-mixture test were positively associated with risk of developing obesity (Salbe, et al., 2004) and taste ratings from a 16-mixture test were positively associated with fungiform papillae density, a genetic trait associated with the obesity phenotype (Hayes & Duffy, 2008). Conversely, taste ratings from 18-mixture and 3-mixture tests were not associated with food selection or PROP taste status (Cooling & Blundell, 2001; Lim, et al., 2008). Thus, while variations of this useful method have been reported in at least 10 studies to assess individual taste perception, which have produced inconsistent findings (Cooling & Blundell, 2001; Drewnowski, et al., 1997; Drewnowski, et al., 1998; Hayes & Duffy, 2007, 2008; Keskitalo et al., 2007;S. V. Kirkmeyer & Tepper, 2003;S.V. Kirkmeyer & Tepper, 2005; Lim, et al., 2008; Ly & Drewnowski, 2001;J. A. Mennella, et al., 2010; J.A. Mennella, Pepino, & Reed, 2005; Salbe, et al., 2004; Yeomans, et al., 2007), the test-retest reliability of the standard 20-solution method has not been reported. Poor reliability might explain such findings, and reliability is a necessary condition for validity. In addition to concerns regarding reliability, a potential liability of the 20-mixture method includes participant burden and satisfaction. Participants anecdotally report that tasting and rating 20 dairy mixtures can be burdensome and unpleasant, a situation which could result in negative subjective states that confound taste perception ratings during the test.

The primary aim of this study was to examine the test-retest reliability of the original 20-mixture method (Drewnowski, et al., 1985; Drewnowski & Greenwood, 1983). Because exclusion of mixtures rated unreliably would decrease the amount of time required to complete method procedures, decrease participant burden, and improve ease of administration and practical utility of the method, a secondary aim was to identify mixtures rated unreliably for possible exclusion from the testing paradigm.

2. Methods and Procedures

2.1 Participants

Forty-six male and pre-menopausal female participants enrolled in the study. Inclusion criteria were 1) healthy men and women not diagnosed with diabetes, cardiovascular illness, or other chronic diseases, 2) between 18 and 60 years old, 3) body mass index (BMI) between 18.5 and 39.9 kg/m^2, and 4) for women, pre-menopausal. Individuals were excluded if they 1) smoked tobacco, 2) took medications with anticholinergic side effects (e.g. antidepressants or antipsychotics), 3) were allergic to foods used in the study, and 4) exhibited attitudes consistent with those of an eating disorder, as assessed by the Multifactorial Assessment of Eating Disorders Symptoms (t score >70 on ≥ 3 subscales; (Anderson, Williamson, Duchmann, Gleaves, & Barbin, 1999). Participants were recruited via advertisement, were screened via telephone and brief on-site evaluations during which weight, height, pulse, blood pressure, and waist circumference were measured, and were monetarily compensated for their time. The study was approved by the Institutional Review Board of the Pennington Biomedical Research Center and participants gave informed consent prior to participation.

2.2 Procedures

Participants completed two test sessions that were 7+/−2 days apart and occurred between the hours of 10:00 AM and 2:00 PM, with the second test session scheduled within 30 minutes of the first. To control for cyclic sex hormones that influence taste perception and gustatory responses in females, the Menstrual Cycle Interview was used to predict the luteal phase of female participants' cycles, and to subsequently schedule both testing sessions during that phase (Geiselman et al., 1998). Immediately prior to testing, research staff confirmed that female participants were in the luteal phases of their cycles, and also that participants were not experiencing cold or allergy symptoms that could affect ability to taste or smell foods.

In private testing rooms, subjects were presented with a tray of 20 randomly ordered dairy mixtures containing graded concentrations of fat and sugar. Specifically, nonfat milk (0.1% fat), whole milk (3.5% fat), half and half (11.3% fat), cream (37.5% fat), and cream with safflower oil (52.6% fat) were each mixed with 0%, 5%, 10%, or 20% sugar by volume, resulting in 20 total mixtures. Each test session took approximately 20 minutes ± 5 minutes. For each mixture participants were instructed first to take a sip of water and swallow it, then sip the mixture, swirl it in their mouths, rate it for sweetness, creaminess, and pleasantness, and finally expectorate the mixture. Electronic Visual Analog Scales (VAS) anchored from "not at all" to "extremely" on a scale ranging 1 to 100 were used to rate sweetness, creaminess, and pleasantness for each mixture. These procedures follow those originally outlined by Drewnowski and Greenwood (Drewnowski & Greenwood, 1983).

3. Statistical Analyses

The test-retest reliability of the method was assessed by computing Pearson correlation coefficients for each dimension[2], with alpha set at .001 for significance testing, to control

[2]PROP taste sensitivity and BMI were included in partial correlation analyses and neither variable meaningfully affected the magnitude or pattern of test-retest correlation coefficients. Therefore, correlation coefficients from analyses that did not include these partial correlations have been reported.

for potential inflation due to multiple testing. Correlation magnitudes were interpreted using standard criteria for which correlations of .10 to .29 are small, .30 to .49 are moderate, and .50 and above are large. The following a priori criteria were established for determining adequacy of test-retest reliability for a given solution: 1) significant ($p < .001$) correlation coefficient, and 2) correlation coefficient of a large magnitude ($\geq .50$).

4. Results

Data from 41 participants were included in the analyses; data from five participants were excluded due to protocol violations that occurred during test administration, or electronic malfunction during collection of VAS ratings. The sample was 73% female (*n*=30), 63% European American (*n*=26), 37% African American (*n*=14), and with a mean age of 33.8 years (*SD*=11.8) and a mean BMI of 27.2 kg/m^2 (*SD*=3.8). Descriptive data for ratings of sweetness, creaminess, and pleasantness for each solution are presented in Table 1.

Correlation coefficients for ratings of sweetness ranged from −.04 to .47, though none were statistically significant with alpha set at .001, or classified as large in magnitude (Table 2). Coefficients for creaminess ratings ranged from −.16 to .74, with 3 of the 20 coefficients statistically significant and classified as large in magnitude. Coefficients for pleasantness ranged from .18 to .71, with 11 of the 20 coefficients statistically significant and classified as large in magnitude. Almost all (9/11) significant coefficients for pleasantness involved solutions with the two highest concentrations of sugar (10% and 20%).

5. Discussion

Taste perception ratings for sweetness and creaminess had poor test-retest reliability over one week, with only 3 of 40 correlation coefficients being large in magnitude and statistically significant. However, ratings for pleasantness of solutions containing higher concentrations of sugar (≥10%) met criteria for satisfactory test-retest reliability. These results call into question the current use of sweetness and creaminess ratings, and of pleasantness ratings for mixtures containing <10% sugar, to measure taste perception. Indeed, the poor test-retest reliability of these ratings challenges the validity of these specific solutions and/or the 20-solution paradigm as a means of quantifying taste perception, and could help explain discrepant and null findings across studies; if the method is not reliable over time, it likely is not a valid measure of trait taste perception, or of changes in perception over time.

These results suggest that a shorter test which elicits ratings only of pleasantness for the 10 solutions that contain 10% and 20% sugar will yield reliable taste sensitivity ratings. This would result in a taste sensitivity test that is 83% shorter than the 20 mixture paradigm that requires participants to make sweetness, creaminess, and pleasantness ratings. A shorter paradigm would have many benefits, including: 1) decreased participant burden, 2) fewer resources required, 3) decreased waste, and 4) improved ease of administration. Further research is warranted to replicate these findings, to determine if use of a shorter, reliable paradigm will more consistently support associations of taste sensitivity, body mass, and obesity phenotypes, or to test other adaptations of the method and other methods of measuring taste perception.

Acknowledgments

Role of Funding Sources. This work was partially supported by National Institutes of Health grant K23 DK068052 (PI: Corby Martin, Ph.D.), and by an NORC Center Grant 1P30 DK072476 entitled “Nutritional Programming: Environmental and Molecular Interactions”. Both grants are sponsored by the NIDDK.

References

Anderson D, Williamson D, Duchmann E, Gleaves D, Barbin J. Development and validation of a multifactorial treatment outcome measure for eating disorders. Assessment. 1999; 6(1):7. [PubMed: 9971879]

Bradbury J. Taste perception: Cracking the code. PLoS biology. 2004; 2(3):e64. [PubMed: 15024416]

Cooling J, Blundell JE. High-fat and low-fat phenotypes: habitual eating of high- and low-fat foods not related to taste preference for fat. Eur J Clin Nutr. 2001; 55(11):1016–1021. [PubMed: 11641752]

Cowart BJ. Development of taste perception in humans: sensitivity and preference throughout the life span. Psychological Bulletin. 1981; 90(1):43. [PubMed: 7267897]

Drewnowski A. Taste preferences and food intake. Annual Review of Nutrition. 1997; 17(1):237–253.

Drewnowski A, Brunzell JD, Sande K, Iverius P, Greenwood M. Sweet tooth reconsidered: taste responsiveness in human obesity. Physiology & behavior. 1985; 35(4):617–622. [PubMed: 4070436]

Drewnowski A, Greenwood M. Cream and sugar: human preferences for high-fat foods. Physiology & Behavior. 1983; 30(4):629–633. [PubMed: 6878464]

Drewnowski A, Henderson S, Shore A, Barratt-Fornell A. Nontasters, tasters, and supertasters of 6-n-propylthiouracil (PROP) and hedonic response to sweet. Physiology & Behavior. 1997; 62(3):649–655. [PubMed: 9272678]

Drewnowski A, Henderson SA, Barratt-Fornell A. Genetic sensitivity to 6-n-propylthiouracil and sensory responses to sugar and fat mixtures. Physiol Behav. 1998; 63(5):771–777. [PubMed: 9617998]

Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999–2008. JAMA. 2010; 303(3):235–241. [PubMed: 20071471]

Geiselman PJ, Smith CF, Williamson DA, Champagne CM, Bray GA, Ryan DH. Perception of sweetness intensity determines women’s hedonic and other perceptual responsiveness to chocolate food. Appetite. 1998; 31(1):37–48. [PubMed: 9716434]

Hayes JE, Duffy VB. Revisiting sugar-fat mixtures: sweetness and creaminess vary with phenotypic markers of oral sensation. Chem Senses. 2007; 32(3):225–236. [PubMed: 17204520]

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]

Keskitalo K, Knaapila A, Kallela M, Palotie A, Wessman M, Sammalisto S, et al. Sweet taste preferences are partly genetically determined: identification of a trait locus on chromosome 16. The American journal of clinical nutrition. 2007; 86(1):55. [PubMed: 17616763]

Kirkmeyer SV, Tepper BJ. Understanding creaminess perception of dairy products using free-choice profiling and genetic responsivity to 6-n-propylthiouracil. Chem Senses. 2003; 28(6):527–536. [PubMed: 12907590]

Kirkmeyer SV, Tepper BJ. Consumer reactions to creaminess and genetic sensitivity to 6-n-propylthiouracil: A multidimensional study. Food quality and preference. 2005; 16(6):545–556.

Lim J, Urban L, Green BG. Measures of individual differences in taste and creaminess perception. Chem Senses. 2008; 33(6):493–501. [PubMed: 18453638]

Ly A, Drewnowski A. PROP (6-n-Propylthiouracil) tasting and sensory responses to caffeine, sucrose, neohesperidin dihydrochalcone and chocolate. Chem Senses. 2001; 26(1):41–47. [PubMed: 11124214]

Mennella JA, Pepino MY, Lehmann-Castor SM, Yourshaw LM. Sweet preferences and analgesia during childhood: effects of family history of alcoholism and depression. Addiction. 2010; 105(4): 666–675. [PubMed: 20148789]

Mennella JA, Pepino MY, Reed DR. Genetic and environmental determinants of bitter perception and sweet preferences. Pediatrics. 2005; 115(2):e216. [PubMed: 15687429]

Reed DR, McDaniel AH. The human sweet tooth. BMC Oral Health. 2006; 6(Suppl 1):S17. [PubMed: 16934118]

Salbe AD, DelParigi A, Pratley RE, Drewnowski A, Tataranni PA. Taste preferences and body weight changes in an obesity-prone population. Am J Clin Nutr. 2004; 79(3):372–378. [PubMed: 14985209]

Stubbs J, Ferres S, Horgan G. Energy density of foods: effects on energy intake. Crit Rev Food Sci Nutr. 2000; 40(6):481–515. [PubMed: 11186237]

Yeomans MR, Tepper BJ, Rietzschel J, Prescott J. Human hedonic responses to sweetness: role of taste genetics and anatomy. Physiol Behav. 2007; 91(2–3):264–273. [PubMed: 17477942]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Highlights

- We assessed test-retest reliability of a common solution-based method for measuring trait taste perceptions.
- Perception ratings were unreliable for most solutions, and were reliable only for perceived pleasantness of solutions high in sugar.
- Findings challenge the validity of the method and indicate that a modified, briefer version should be considered.

Table 1

Mean (SD in parentheses) ratings of sweetness, creaminess, and pleasantness for dairy solutions varied in fat and sugar content on a 1–100 Visual Analog Scale.

	0.0% fat	0.5% fat	1.3% fat	7.5% fat	52.6% fat
Sweetness Ratings					
0% Sugar	26.2 (30.6)	36.1 (34.2)	19.9 (29.3)	34.0 (35.7)	23.2 (29.4)
	23.8 (31.3)	18.8 (24.5)	21.0 (26.5)	23.2 (31.0)	16.6 (24.2)
5% Sugar	49.8 (30.8)	50.3 (28.9)	42.9 (32.1)	44.2 (32.7)	54.3 (29.4)
	52.1 (29.1)	45.0 (27.2)	48.6 (28.2)	47.0 (25.6)	46.0 (26.4)
10% Sugar	54.8 (30.8)	54.8 (31.5)	55.6 (33.1)	65.9 (28.5)	64.0 (28.6)
	64.4 (27.8)	60.1 (26.5)	57.2 (29.5)	64.1 (25.1)	66.9 (25.8)
20% Sugar	60.7 (26.9)	59.1 (33.3)	63.0 (32.5)	65.6 (32.8)	68.0 (30.5)
	67.7 (33.2)	82.4 (19.9)	68.3 (26.8)	75.6 (22.9)	80.4 (20.9)
Creaminess Ratings					
0% Sugar	33.0 (31.9)	38.3 (30.7)	33.2 (28.8)	64.5 (27.0)	57.4 (29.8)
	33.7 (32.4)	29.2 (28.5)	28.5 (24.6)	65.7 (25.3)	66.0 (29.6)
5% Sugar	38.7 (28.4)	43.2 (26.8)	45.0 (30.1)	66.0 (27.1)	67.2 (28.1)
	33.8 (23.8)	33.6 (24.3)	44.6 (27.1)	67.7 (26.6)	69.2 (29.1)
10% Sugar	39.2 (25.2)	49.1 (28.4)	50.3 (26.1)	69.5 (26.3)	63.4 (26.9)
	34.9 (28.7)	37.7 (27.8)	42.1 (26.4)	70.4 (24.9)	72.1 (26.4)
20% Sugar	37.1 (25.6)	46.8 (28.6)	52.6 (27.8)	71.2 (28.9)	73.5 (31.5)
	47.8 (28.5)	43.7 (30.8)	46.8 (27.7)	76.7 (25.4)	76.1 (25.9)
Pleasantness Ratings					
0% Sugar	29.6 (25.0)	28.2 (22.8)	25.3 (28.5)	29.7 (25.6)	22.6 (20.9)
	22.7 (18.5)	24.2 (20.4)	24.5 (21.5)	24.2 (22.0)	21.1 (19.0)
5% Sugar	46.0 (26.4)	39.7 (24.9)	41.4 (29.5)	39.3 (31.4)	45.5 (28.4)
	40.7 (23.7)	39.7 (24.1)	39.2 (26.5)	43.0 (23.8)	43.6 (26.2)
10% Sugar	47.6 (26.7)	44.3 (26.2)	49.2 (26.2)	52.5 (25.9)	49.0 (28.9)
	44.4 (27.3)	48.2 (27.8)	47.2 (31.0)	51.6 (27.8)	51.3 (29.2)
20% Sugar	49.8 (28.2)	49.9 (27.9)	51.1 (26.1)	51.0 (31.2)	51.7 (31.4)
	48.4 (31.8)	55.7 (32.1)	51.5 (28.6)	56.8 (28.2)	53.3 (31.3)

Table 2

Test-retest reliability correlation coefficients for ratings of sweetness, creaminess, and pleasantness of dairy products varied in fat and sugar content (p-values in parentheses)on a 1–100 Visual Analog Scale.

	0.0% fat	0.5% fat	1.3% fat	7.5% fat	52.6% fat
Sweetness Ratings					
0% Sugar	.35 (.027)	.37 (.016)	.15 (.356)	.26 (.094)	.35 (.024)
5% Sugar	.23 (.156)	−.03 (.856)	.41 (.007)	.27 (.083)	.47 (.002)
10% Sugar	.41 (.008)	.39 (.011)	.23 (.156)	.18 (.258)	.04 (.819)
20% Sugar	.17 (.274)	.36 (.020)	.45 (.003)	−.04 (.785)	.33 (.037)
Creaminess Ratings					
0% Sugar	.31 (.046)	**.53* (<.001)**	.39 (.011)	.15 (.363)	.08 (.621)
5% Sugar	−.16 (.307)	.13 (.412)	.21 (.184)	.18 (.269)	.36 (.020)
10% Sugar	.15 (.365)	−.11 (.490)	.31 (.046)	.33 (.034)	.30 (.055)
20% Sugar	.20 (.216)	.10 (.538)	.37 (.018)	**.64* (<.001)**	**.74* (<.001)**
Pleasantness Ratings					
0% Sugar	**.50* (.001)**	.36 (.020)	.18 (.254)	.51* (<.001)	.31 (.045)
5% Sugar	.38 (.016)	.21 (.190)	.43 (.005)	.44 (.004)	.44 (.004)
10% Sugar	**.54* (<.001)**	**.71* (<.001)**	**.52* (<.001)**	**.57* (<.001)**	.47 (.002)
20% Sugar	**.66* (<.001)**	**.60* (<.001)**	**.51* (<.001)**	**.55* (<.001)**	**.70* (<.001)**

Note. Alpha was set at .001 to control for inflation due to multiple statistical tests. Bolded values indicate solutions for which ratings satisfied both reliability criteria.

* $p \leq .001$.