

Characterizing individual differences in sweet taste hedonics: Test methods, locations, and stimuli

May M. Cheung¹, Matthew Kramer², Gary K. Beauchamp¹, Paul M. Wise¹

1. Monell Chemical Senses Center, 3500 Market St, Philadelphia, PA 19104
2. Beltsville Agricultural Research Center, United States Department of Agriculture, 10300 Baltimore Ave, Beltsville, MD 20705

Corresponding author: May M. Cheung, Ph.D., Room 319, 3500 Market St, Philadelphia, PA 19104; email: mcheung@monell.org

Funding sources: Monell Institutional Fund, NIDCD T32 Institutional Training Fund

Abstract: Sweetness drives consumption of added sugars, so understanding how individuals differ is important for developing strategies to lower sugar intake. However, methods to assess hedonic response to sweetness vary, making results across studies difficult to integrate. We compared methods to measure optimal sucrose concentration in 21 healthy adults (1) using paired-comparison preference tracking vs. ratings of liking, (2) with participants in the laboratory vs. at home, and (3) using aqueous solutions vs. vanilla milk. Tests were replicated on separate days to assess test-retest reliability. Test-retest reliability was similar between laboratory and home testing, but tended to be better for vanilla milk and preference tracking. Optimal sucrose concentration was virtually identical between laboratory and home, slightly lower when estimated via preference tracking, and about 50% lower in vanilla milk. However, individual optimal sucrose concentration correlated strongly between Methods, test Locations, and Stimuli. More than 50% of the variability in optimal sucrose concentration could be attributed to consistent differences among individuals while much less variability was attributable to differences in Methods, test Locations or Stimuli. These results demonstrate convergent validity between measures of preference and liking, support testing at home to lower participant burden, and suggest that aqueous solutions can be useful proxies for some commonly consumed beverages for measuring individual differences.

Keywords: Sweet taste; hedonics; individual differences; methodology; sugar

1. Introduction

The appeal (hedonic response) of sweetness is a key driver of consuming foods high in added sugars^{1,2}. Overconsumption of added sugars contributes to increased risk of obesity and related chronic illnesses such as type 2 diabetes mellitus and cardiovascular disease³⁻⁵. A number of leading health agencies have recommended a reduction in intake of added sugars to improve public health and prevent chronic disease^{6,7}. To develop successful treatments or policies to reduce intake of sugar, it is vital to understand individual differences in sweet taste and how they interact with diet to drive consumption. Yet, research on the relationship between sweet taste, diet and health has been equivocal⁸⁻²³. Some researchers have found significant associations between perception of sweetness and intake of sugars¹⁹, carbohydrates^{9,10,20}, energy from sweetened beverages^{16,21}, total energy intake²⁰, and body composition¹⁵. Other researchers have found no associations^{12-14,18,22,23}. Differences in outcome may be due in part to differences in methods used to measure sweet perception, and several recent reviews have identified a need to standardize procedures^{2,11,15,24-26}.

Measures of sweet perception range from tests of sensitivity (e.g., minimum amount of sweetener detected or recognized as sweet), supra-threshold intensity (sweetness rated on various scales), and hedonic response (liking, preference, and related constructs)¹⁸⁻²⁰. Though associations with consumption of sugar have been observed for all these types of measures, association with hedonic response is the most consistent and most promising^{2,11,19}. The current study focuses primarily on two common classes of hedonic measures: preference and liking, though sweetness intensity was also measured.

Measures of preference typically involve comparisons among stimuli that differ in sweetness²⁷⁻³⁰. The Monell Forced-Choice, Paired-Comparison Preference Tracking test included in The NIH Toolbox for Assessment of Neurological and Behavioral Function is a widely used example³¹. Each trial, participants taste a pair of aqueous solutions of sucrose selected from five total concentrations (0.09 to 1.05 M), and they must choose the more preferred concentration. Concentrations of stimulus pairs vary

over an experimental session to find the most preferred concentration as described in the methods (2.5.2, below). Henceforth, this method is called preference tracking (PT). This PT test was chosen as a representative measure of preference.

Measures of liking typically involve subjective ratings, though the sweet stimuli, concentration(s), liking scales, and methods to summarize data vary¹⁸⁻²⁰. To facilitate direct comparisons with preference tracking, the current study includes ratings of liking for the same five concentrations of sucrose used in preference tracking test. A visual analogue scale (VAS) was chosen for ratings of liking since most VASs are reliable and easy for participants to use^{24,26,32}. Henceforth, this method is called the rating method (RM). If the two techniques measure the same underlying variable, as we hypothesized, the most liked (highest rated) concentration should correspond to the most preferred concentration measured using PT. However, few studies have directly compared the two approaches in the same participants to determine whether there is convergent validity between the methods.

The advent of the COVID-19 pandemic in early 2020 greatly heightened interest in testing without person-to-person contact, with some encouraging results³³⁻³⁵. Conducting tasting sessions with participants at home could lower participant burden, use of fossil fuel, and risks associated with repeated visits to the laboratory (including risk of infectious illness). Testing at home or in the field has yielded useful data³³⁻³⁵, but some researchers suggest that environmental factors can influence emotional and sensory responses to test samples^{33,36-39}. However, to the best of our knowledge, laboratory and home testing environments have not previously been compared for the preference tracking test.

A variety of stimuli have been used to determine sweetness preference and liking. Aqueous solutions of sucrose (AS) are perhaps the most commonly used because they are easy to prepare and yield highly reproducible data⁴⁰. However, AS take sweetness out of its natural context in foods and beverages. More realistic model beverages have also been used to study sweet hedonics, e.g., vanilla milk (VM)⁴¹ and

lemonade^{13,42}, and do not always yield identical results to AS^{19,43}. Regardless, there have been few direct comparisons between more realistic beverages and AS in the same participants using multiple procedures and test settings to determine how the various methodological factors interact.

The current methods study was undertaken to inform the effort to compare and standardize methods. Three main manipulations were conducted: (1) Hedonic response measurement Methods (paired-comparison judgments of preference vs. rated liking); (2) test Locations (in the laboratory vs. at home, supervised via video conferencing); (3) two Stimulus matrices (simple aqueous solutions of sucrose vs. sucrose in vanilla milk). All participants were tested under all conditions.

2. Materials and methods

2.1 Ethics statement

Study procedures were reviewed by an institutional review board (IRB) at the University of Pennsylvania (protocol # 844423). The study was conducted in accordance with the guidelines of the Declaration of Helsinki, and participants provided written, informed consent prior to engaging in study procedures.

2.2 Participants

We recruited women and men between the ages of 21 and 65 years old from the Greater Philadelphia area between January and June 2021. Participants were recruited using flyers and from a pool of previous Monell participants who opted to be contacted again for future studies. We included generally healthy adults free of chronic diseases. We excluded individuals with current and chronic illnesses (e.g., heart diseases, diabetes, HIV/AIDS, kidney diseases), those with a history of food allergies or sensitivities, and those who medication daily (except for birth control). Participants were screened via a phone interview for inclusion and exclusion criteria prior to enrollment. Informed consent was obtained from each participant prior to all data collection.

2.3 Design

All participants were tested under all conditions. After an initial laboratory session dedicated to instructions and practice as described below, participants completed four sensory testing sessions. Two replicate sessions were conducted in a sensory testing facility at the Monell Chemical Senses Center (Lab), and two were conducted with participants at home, guided via video conferencing (Home). Eleven participants alternated test setting as follows: Lab, Home, Lab, Home. Ten participants alternated Home, Lab, Home, Lab. Assignment to the two orders was counter-balanced. In each session, participants completed both sensory tasks, i.e., preference tracking (paired comparisons, PT) and ratings (ratings of liking and intensity of sweetness, RM). The order of the sensory tasks (PT and RM) was counter-balanced across replicate sessions. Within each sensory task, participants tasted both stimuli (AS and VM). Order of the stimuli within sensory task was randomized within sessions.

Safety measures were developed in consultation with the Monell Human Subjects Committee. Participants were screened for symptoms and known or suspected exposure the night before and the day of each visit. One participant was tested at a time, with at least 30 minutes between participants for aerosols to clear via building ventilation (with UV filtration) and to disinfect the testing room. Participants wore masks while not tasting. The experimenter wore a dental gown, N-94 or N-95 mask, face shield, gloves, and hair covering. Social distancing was practiced to the extent possible. Neither experimenters nor participants reported symptoms or positive COVID-19 tests during this study.

2.4 Stimulus materials

Stimuli were prepared using food-grade sucrose (Fisher Chemical, crystalline/NF, catalog # S3-500) dissolved in either MilliporeTM filtered, distilled, deionized water (aqueous solutions, or AS) or in vanilla milk (VM). AS was prepared at the five concentrations used in the preference tracking test²⁷: 0.09, 0.18, 0.35, 0.70, and 1.05 M. VM consisted of 2% fat milk (GiantTM brand, SKU# 688267008634), and vanilla extract (1 ml per liter; Nature's Promise Organic, SKU# 688267156502). Milk and vanilla were purchased from a local grocery store. Milk was stored under refrigeration, and never

used past use-by dates. Sucrose concentrations in VM were 0.03, 0.12, 0.23, 0.47, and 0.70 M, selected based on pilot work to approximately match the sweetness of AS at each step. Solutions were prepared using sterile glassware. Samples were presented as 10 mL aliquots in 30 mL plastic medicine cups and served cold at approximately 4°C.

For Lab tests, experimenters poured samples out of participants' sight. For Home tests, participants poured 10 mL samples themselves, using graduated plastic medicine cups provided. Home test kits included 20, 120 mL bottles (two instances of each of the five concentrations of each stimulus, sufficient for one test session). Bottles were labeled with random alpha-numeric codes and packed in cooler bags with ice packs for transport. Lab and Home sessions alternated, as described below, so participants received a test kit to take home at the end of Lab sessions. Participants were instructed to take the kits directly home and refrigerate samples until scheduled sessions. Samples were used within 5 days or less.

2.5 Procedures:

2.5.1 Training. The training session was dedicated to measurement of participant height/weight, collection of demographic information, and task instructions. Participants received standard instructions on the use of the general Labeled Magnitude Scale (gLMS) for measuring sweetness intensity, then practiced by rating the intensities of real and imagined sensations⁴⁴. The gLMS is a vertical scale with intensity descriptors as follows: barely detectable, weak, moderate, and strong, with the spacing of the descriptors on the scale determined empirically to be proportional to strength of sensation⁴². Next, participants were instructed in the use of a 100-point visual analogue scale (VAS) for rated liking (horizontal scale, anchored with “dislike extremely” on the left and “like extremely” on the right)³¹. Participants practiced by rating liking for remembered or imagined sensations.

2.5.2 Preference tracking. Procedures for the Monell Forced-Choice, Paired Comparison Tracking Procedure have been published elsewhere²⁷. Participants completed the preference tracking procedure for each stimulus during each session,

with breaks of 5 min between preference tracking runs for the two stimuli. Participants began each session by rinsing the mouth four times with distilled water (Good & Gather™, SKU#: 085239047675). Each trial, participants tasted pairs of 10 mL samples of liquid presented in 30 mL plastic medicine cups, with a 1-minute break between pairs. Participants then chose which stimulus they preferred. The first pair of samples were the from the middle range (0.18 M vs. 0.70 M). Each subsequent pair contained the participant's preceding preferred concentration paired with an adjacent stimulus concentration. This pattern continued until the participant chose two consecutive times either the same concentration of sucrose paired with both a higher and lower concentration or the highest (1.05 M) or lowest (0.09 M) concentration. The entire task was repeated after a 5-minute break, with the stimulus pairs presented in reverse order. The procedure was repeated twice within a session, and the geometric mean of the two trials was defined as the most preferred concentration.

2.5.3 Ratings of sweetness and liking. Participants sampled by taking the entire contents of the cup into the mouth, moving the liquid around in the mouth for several seconds, then rating liking and sweetness intensity in that order^{24,44}. Stimulus presentation and tasting followed procedures in 2.5.2, above. Next, participants expectorated the sample and rinsed with water at least twice to begin a 1-minute pause before the next sample. Each sensory test session included 20 trials, separated into two blocks of 10 (all five concentrations for a given stimulus in random order, then again in random order). A 5-minute break separate blocks for the two stimuli.

2.6 Data Analysis:

2.6.1. Sensory endpoints. Measures of main interest included optimal sucrose concentration estimated via PT (most preferred) and optimal sucrose concentration estimated via RM (most liked) at different locations (Lab vs Home) in two Stimulus matrices (AS vs VM). Methods for calculating most preferred concentration have been described previously²⁷. Most liked was defined as the concentration associated with the maximum (among presented concentrations) in rated liking. To find maxima, rated liking was plotted against the cube root of sucrose concentration (to space concentrations

approximately equal distances apart). Functions of liking vs. transformed concentration were fitted using stepwise regression (lowest BIC criterion), up to a cubic polynomial. If the resulting model was intercept only (flat), the most liked concentration was defined as the geometric mean of the two concentrations associated with the highest ratings (see **Supplementary Table 1** for fit parameters). These parametric estimates of most liked concentration correlated strongly ($r = 0.86$ for AS, $r = 0.94$ for VM) with the concentration associated with the highest rating of liking; because the two methods for estimating most liked concentration supported the same conclusions, only values from model fits are reported.

Most preferred and most liked concentrations were both positively skewed. The best common Box-Cox power transformation ($\lambda = 0.22$) was used for both PT and RM prior to inferential analysis. Mean values of optimal sucrose concentration, defined as the most preferred or the most liked concentrations from PT or RM, reported below were back-transformed and reported in the original units (molar concentration of sucrose). Rated sweetness intensity (gLMS) was of secondary interest. Ratings made using the gLMS were also positively skewed. The optimal Box-Cox power transform ($\lambda = 0.47$) was used for inferential analysis for sweet intensity. Mean values were back-transformed and reported in units ranging from 0 (“no sensation”) to 100 (“strongest imaginable sensation of any kind”).

2.6.2. Statistics. Test-retest reliability between replicate sessions was evaluated using Pearson’s correlation coefficients (r). To examine differences in test-retest reliability between Locations (Lab vs. Home), Stimuli (AS vs. VM), and Methods (PT vs. RM) for estimating optimal sucrose concentration, we tested pairs of partial correlation coefficients using the Fisher r -to- z transformation. For example, to assess whether optimal concentration is more reliable in the laboratory than at home, the partial correlation (across participants, adjusting for Stimulus and Method) between session 1 and session 2 in the laboratory was compared to the corresponding partial correlation between session 1 and session 2 at home. Partial correlations were tested in the same way for Method and Stimulus. Repeated measures analysis of variance (ANOVA) was

used to assess the effects of experimental conditions on average values of sensory measures. Overall variance in optimal concentration was decomposed using a linear mixed model with all effects random (to estimate variances). Main effects were Method, Stimulus, Location, and Participant. All second order interactions were also included. The residual term contained within subject variability, unexplained variability (not due to included terms, such as higher order interactions or unmeasured independent variables), and model error. Analyses were conducted using the lme4 package⁴⁵ in R (R Development Core Team, 2021) and Microsoft Excel (Version 16.52, Microsoft Corporation). All results, unless otherwise indicated, were reported in mean \pm standard deviation or (95% confidence interval).

3. Results

3.1 Participant characteristics

Twenty-one healthy adults between the ages of 21 and 49 (33.1 ± 10.0 years) participated. Eleven self-identified as female, the rest self-identified as male. Self-identified racial demographics were as follows: 16 Caucasian; 3 African-American/Black; 1 Asian; 1 multi-racial. Average body mass index (BMI) was 29.6 ± 7.8 kg/m². Four participants were normal weight (BMI < 25 kg/m²); ten were overweight (BMI between 25 to 30 kg/m²), and seven were obese (BMI ≥ 30 kg/m²).

3.2. Hedonic measures

3.2.1 Test-retest reliability. For rated liking, averaged across concentrations, test-retest reliability coefficients were $r = 0.66 \pm 0.13$ and 0.64 ± 0.11 for AS tested in Lab and at Home, respectively. Average test-retest reliability coefficients were $r = 0.67 \pm 0.16$ and 0.72 ± 0.11 for VM tested in Lab and at Home, respectively. Test-retest reliability for optimal concentrations derived from ratings of preference tracking (most preferred) and liking (most liked, as described in 2.7.1, above) were of comparable strength (**Table 1**), though test-retest reliability was lower for most liked concentration of AS.

Differences in test-retest reliability between test locations were not statistically significant ($p = 0.72$). VM was more reliable than in AS ($p = 0.02$). PT tended to be more reliable than RM ($p = 0.06$).

3.2.2 Mean values of optimal concentration. Individual participants displayed different patterns of rated liking over concentrations (**Supplementary Figures 1 and 2**). A 4-way ANOVA was performed on optimal concentration: Method (PT vs. RM) X Location X Stimulus X Session, with Participant as a random effect. Optimal concentration estimated by RM was slightly but significantly higher than PT (**Figure 1**); $F(1, 305) = 10.03$, $p = 0.002$. There was a significant main effect of Stimulus; $F(1, 305) = 69.35$, $p < 0.001$. Optimal concentration tended to be lower in VM than in AS (**Figure 1**). Finally, there was a significant Stimulus X Session interaction; $F(1, 305) = 4.90$, $p = 0.028$. Averaged across other conditions, optimal concentrations in AS differed little between the first (0.39 M; 95% CI: 0.29, 0.51) and second sessions (0.37 M; 95% CI: 0.26, 0.51). However, for VM, optimal concentration increased slightly from the first session (0.23 M; 95% CI: 0.16, 0.31) to the second session (0.27 M; 95% CI: 0.19, 0.37). No other effects reached significance ($0.10 < p < 0.94$). Thus, mean optimal concentration did not differ between Location (Lab and Home), nor was testing Location or its interactions with other variables significant. Optimal concentration differed by Method and Stimulus, but the only interaction was a tendency for optimal concentration to increase from session 1 to session 2 and only for VM.

3.2.3 Correlations between conditions across individuals. Correlations between optimal concentration measured under the various experimental conditions (using PT vs. RM, in the Lab vs. Home, and using AS vs. VM) were positive. The average correlation between optimal concentration measured with at least one difference in conditions ($r = 0.68 \pm 0.11$) was comparable to the average test-retest reliability ($r = 0.71 \pm 0.13$). Averaging across replicate sessions, correlations between most preferred and most liked concentrations ranged from 0.83 to 0.89, suggesting that both methods provided overlapping information on individuals (**Figure 2**). Correlations between Lab and Home ranged from 0.84 to 0.87, suggesting that tests conducted at the two

Locations provided overlapping information on individuals (**Figure 3**). Finally, correlations between optimal concentrations measured using AS and VM ranged from 0.72 to 0.85, suggesting that both model stimuli provided overlapping information on individual differences in hedonic judgments (**Figure 4**).

In the variance decomposition analysis, the Participant X Method, Participant X Stimulus, and Participant X Session two-way interactions accounted for between 3 and 6% of variance, suggesting that individuals differed slightly in their response to testing methods (**Table 2**). However, the individual-to-individual variability by itself (Participant) accounted for by far the largest proportion of total variance, about 51%. Accordingly, though optimal concentration differed between Methods and Stimuli (3.1.2, above), consistent differences among individuals had a larger effect on optimal sucrose concentration than experimental main effects.

3.3. Rated sweetness intensity

3.3.1 Test-retest reliability. For AS, average (across concentrations) test-retest reliability coefficients were similar between test Locations (Lab $r = 0.65 \pm 0.08$; Home $r = 0.64 \pm 0.14$). For VM, average test-retest reliability were also comparable between Locations (Lab $r = 0.50 \pm 0.30$; Home $r = 0.56 \pm 0.16$). A 4-way, repeated measures ANOVA was performed on transformed ratings: Location X Stimulus X Sweetness Level (the five concentrations per Stimulus) X Session, with Participant as a random effect. There was a significant main effect of Session; $F(1, 1637) = 11.53$, $p < 0.001$. Ratings were lower overall in session 1 (19.29; 95% CI = 16.35, 22.36) than in session 2 (19.76; 95% CI 16.94, 22.83). A significant Location X Session interaction; $F(1, 1637) = 6.27$, $p < 0.02$, reflected a difference between sessions in the laboratory: Ratings were slightly lower in session 1 (18.75; 95% CI = 15.53, 22.29) than in session 2 (20.13; 95% CI 17.11, 23.42). There were no differences between session 1 (19.53; 95% CI = 16.24, 23.14) and session 2 (19.54; 95% CI 16.56, 22.76) at Home. Thus, differences between sessions were small in practical terms, and suggest greater stability at Home. No other effects involving Location reached significance. In general, there was good agreement between Lab and Home tests (agreement between Locations also held at the individual participant level, **Supplementary Figures 3 and 4**).

The effect of Sweetness Level was significant; $F(1, 1637) = 1422.86$, $p < 0.001$, an expected dose-response relationship (**Figure 5**). The effect of Stimulus was also significant; $F(1, 1637) = 39.42$, $p < 0.001$. Ratings were slightly lower overall for AS (18.14; 95% CI = 14.99, 21.62) than for VM (20.90; 95% CI 17.81, 24.25). Furthermore, there was a significant Stimulus X Sweetness Level interaction; $F(1, 1637) = 12.37$, $p < 0.001$. Rated intensity covered a similar range for the two stimuli, but VM was sweeter for concentration step 2 (**Figure 5**). Other effects were not significant. Thus, sweetness covered a similar range for AS and VM, and sweetness was approximately, though not perfectly, matched at individual concentration steps.

3.3.2 Correlations between sweetness intensity and optimal concentration. To determine if individual differences in sweetness intensity were associated with individuals' optimal concentration, average (across concentrations) sweet intensity was calculated as a measure of sweet taste sensitivity. Rated intensity for AS was not significantly ($p > 0.05$) correlated with either most preferred ($r = 0.07$) or most liked ($r = 0.07$) concentration of AS. Similarly, rated intensity for VM was not significantly correlated with either most preferred ($r = -0.10$) or most liked ($r = -0.09$) concentration. Accordingly, the data analyses provide no evidence that individual differences in sweetness intensity are associated with individual estimates of optimal concentration of sucrose.

4. Discussion

4.1. Summary of major findings

The current study assessed the effects of Method (PT vs. RM), test Location (Lab vs. Home), and Stimulus (AS vs. VM) on optimal concentration of sucrose in model beverages. Between Methods, test-retest reliability for most preferred (PT) and for most liked (RM) concentration did not differ although the former tended to be greater than the latter. Optimal concentration across participants tended to be slightly higher for RM than for PT, but the two measures were strongly correlated across individuals. In terms of Location, results were essentially identical between test Locations in terms of test-retest reliability, mean values, and patterns of individual differences. For Stimulus, the optimal sucrose concentration averaged across participants was higher in AS than in VM (by

about 50% on average), but values for the two stimuli were strongly correlated across individuals. Rated sweetness intensity was also similar between Lab and Home. Overall, our results suggested that all methods tested in the current study captured the same underlying variance in optimal sucrose concentration in model beverages. Remote testing at home to lower participant burden is feasible and reliable. Furthermore, aqueous solutions can be useful proxies for some commonly consumed beverages for measuring individual differences in sweet hedonic response.

4.2. Preference tracking vs. rating method

Judgments of preference and ratings of liking are different tasks for measuring sweet hedonics, but they captured similar person-to-person differences. Results from the two methods (PT vs. RM) correlated strongly ($0.83 \leq r \leq 0.89$), and the variance decomposition suggested that Method accounted for only about 1% of total variance in selecting the optimal concentration, compared to consistent differences among individuals, which accounted for about 51% of the total variance, Method was not a significant contributor to the differences in optimal concentration. These findings confirm those of Asao and colleagues, who also found a strong association between most preferred and most liked concentration of sucrose in aqueous solutions using similar methods⁴⁶. The current study extends those findings to show associations of comparable strength across two stimuli and testing environments. Other studies, which did not focus on optimal concentration for both tasks, found that most preferred concentration differed across sweet liker phenotypes, defined according to the overall shapes of functions of rated liking vs. concentration²⁸. Another study found that rejection thresholds, or the concentration differences at which people choose a low sugar formulation of orange juice over a higher sugar formulation in preference judgements, were only measurable for people whose ratings of liking decreased at high concentrations⁴⁷. Our results add to the accumulating evidence that judgments of preference and of ratings of liking measure the same underlying variable of hedonic response. Here, tests used the same stimulus, with the same participants, tested at the same time. Agreement between PT and RM might be weaker if the tests also differed in stimuli, concentration range, number of concentration steps, or other factors. However,

these findings suggest that studies of individual differences in sweet hedonics that use paired comparison preference judgements and ratings of liking could be meaningfully integrated in reviews or meta-analyses.

Ratings of liking may yield higher optimal concentration compared to judgements of preference. Note that the difference between the methods was small and consistent across test Locations and Stimuli, which suggests that an adjustment factor could be applied if one desired to use one test as a proxy for the other. PT was slightly more reliable and may provide practical advantages in some situations. As for both methods, test-retest reliability in the current study fell well within the range previously reported using similar methods^{20,27,46,48,49}. The small difference in test-retest reliability between the methods may mean more replicate measures using RM might be required to match the test-retest reliability of PT.

4.3. In laboratory vs. at home tests

Consistent with other studies, our findings show that sensory test conducted outside a carefully controlled laboratory setting are feasible and generally comparable to laboratory tests^{33-35,50-54}. For example, a study by Seo and colleagues compared data from drive-in sensory booths to data from the laboratory in 106 consumers³⁴. Hedonic and emotional responses to model beverages were not statistically different between testing conditions³⁴. Another study conducted by the Italian Sensory Science Society compared laboratory and remote (at work and at home) sensory tests, including ratings of liking with multiple food items, and found similar results between test locations³³. With adequate guidance by experimenters via video conferencing, most remote tests (including ratings of liking) yielded data comparable to those from laboratory tests³³. Consistent with these findings, test-retest reliability, mean values, and patterns of individual differences in optimal sucrose concentration were essentially identical between test Locations in the current study, as was rated sweetness. Furthermore, variance decomposition suggested that test Location and second order interactions involving Location accounted for little or no variance in optimal concentration. The current study extends previous findings to paired-comparison preference tracking of sweetness in two model stimuli. Thus, results suggest that remote evaluation of hedonic

response to sweetness is feasible and comparable to laboratory tests. Limiting or eliminating visits to the laboratory could reduce participant burden, use of fossil fuels, and risk of exposure to infectious illnesses such as COVID-19. Remote testing would also expand the pool of potential participants beyond those available locally, which could in turn ease recruitment efforts and facilitate studies with more generally representative samples.

4.4. Aqueous solutions vs. vanilla milk

The sample matrix used in sensory tests can affect hedonic responses^{19,28,43,55,56}. The most commonly used stimuli for measuring individual differences in hedonic response to sweetness are simple aqueous solutions of sucrose^{26,57}. However, AS take sweetness out of the natural context of foods and beverages, which might affect patterns of individual differences. For example, Bertino and colleagues found that Taiwan-born students studying in the US tended to show higher preference for sucrose solutions than students of European descent, but preferred lower concentrations in cookies⁴³. Similarly, Holt and colleagues found differences in rated liking for orange juice and biscuits between Malaysian-born and Australian-born students, but hedonic response to sucrose solutions did not predict these differences¹⁹. Thus, experience and culture may play a role in determining optimal sweetness in particular foods⁴⁰. Furthermore, individual differences in sensitivity to sourness, bitterness, and astringency (sensations which are often ameliorated by sweetness) might interact with liking for sweetness to drive preference in foods and beverages in which these generally negative sensations are prominent^{55,58}. However, the current study found the optimal concentrations of sucrose were correlated between AS and VM ($0.77 \leq r \leq 0.85$). Linear model analysis suggested that stimulus accounted for about 10% of total variance in optimal concentration (consistent with small mean difference between methods), and second order interactions involving stimulus accounted for up to only 6%. This variance accounted for is likely due to the mean difference for optimal concentrations between the stimuli, not to variations in individual differences. Thus, within the sample of participants living in the Philadelphia area of the United States, AS and VM provided comparable information on individual differences in optimal sucrose concentration. This

result suggests that individual differences in optimal sucrose concentration in AS can be generalized to some more realistic beverages²⁸, though based on past work, we expect that the results may not generalize to all model foods⁵⁹.

Although hedonic results from AS and VM are strongly correlated, the overall optimal concentration of sucrose was substantially lower in VM than in AS, an effect that was consistent across sensory methods and test Locations. Previous studies have suggested that sweetness might interact with other flavor components of foods and beverages, including fat and aroma, to shape hedonic response⁵⁹⁻⁶¹. In particular, vanilla flavor can enhance sweetness in both simple model solutions and dairy beverages in cultures and ethnic groups where vanilla is associated with sweet taste^{41,62,63}. Also, the slightly higher test-retest reliability in VM compared to AS measured at home might be because it is easier for participants to choose an optimal concentration in a more familiar beverage context. Consistent with sweetness enhancement by vanilla, participants rated VM as equally sweet or even sweeter than AS at each concentration step despite lower sucrose concentrations. It would be interesting to repeat the comparison between stimuli using nose-clips to block retro-nasal aroma to see if the optimal concentrations still differ between stimuli. Results might be relevant to the hypothesis that aroma can be used to partially compensate for reduced sugar levels in beverages⁶⁴.

4.5. Rated sweetness intensity

Rated sweetness was not associated with individual differences in optimal sucrose concentration. This result is consistent with many studies on sweet hedonics, which have found that sweet liking is weakly associated with perceived intensity of sweetness^{16,24,28,65,66}. Despite good matching at each sweetness concentration step in a pilot study, the current sample of participants rated the second and third concentration steps of VM as sweeter than corresponding steps of AS. Though perceived sweetness covered a similar range for the two stimuli, this imbalance at two intermediate steps is consistent with the lower optimal concentrations found in VM, as discussed above (section 4.4). However, the imbalances in intensity did not affect the correlation between individual optimal concentration between the two stimuli.

4.6. Limitations

The study several has limitations. First, this study has a modest sample size relative to many past studies of hedonic response to sweetness. However, the statistical power of this study was sufficient to demonstrate associations between Methods, test Locations, and Stimuli in characterizing individual differences. The sample size was also sufficient to discern modest differences in optimal sucrose concentration between test conditions. However, conclusions regarding apparent differences in correlations between methods and stimuli require confirmation in a larger sample. It is also possible that different sweet hedonic patterns could be seen in a larger sample that includes a wider range of ethnic and cultural backgrounds, ages, or more varied health status. Another limitation is that the study included a limited array of methods. Only one realistic model beverage was compared to AS. More beverages and foods, particularly solid foods, would be useful to evaluate the generality of the findings, but it would be difficult to include enough model foods in a single study to reach very broad conclusions. Also, other endpoints related to hedonic response, e.g., wanting and purchase intent, could be incorporated in future studies. Furthermore, we only used five sucrose concentrations, considerably fewer than in some studies^{16,24,46}. This might limit resolution of individual hedonic response, which could in principle obscure some differences between methods. Although the range of concentrations used was comparable to some other studies^{13,19,24,28,46}, different results might be obtained if higher concentrations of sucrose were used. Test-retest reliability was measured over a short period of time. Previous researchers have shown that preference for sweetness may interact with age^{15,27,67} and dietary factors^{15,68}. Studies with many replicate sessions over time would be useful in this regard. Regarding the conclusion that home tests provide data comparable to laboratory tests, it should be noted that participants attended one in-person training session prior to their home testing sessions. Further work would be required to determine if remote training is equally effective. Finally, our study did not measure changes in sweet hedonics over time. Previous studies reported that sweet preference changes throughout lifespan^{27,67}. Diet may also influence sweet taste sensitivity and preference^{8,69,70}. Most of our participants completed all study

sessions within eight weeks without any prescribed dietary interventions. It will be important to extend this work to a study in a larger sample size with dietary manipulations to see if changes in sugar intake can affect sweet hedonics outcomes.

5. Conclusions

Optimal concentration of sucrose in model beverages was virtually identical between in laboratory and at home tests and varied only slightly by method and stimulus. Individuals' estimates of optimal concentration correlated strongly among the various conditions, suggesting that all methodological conditions ultimately captured a common underlying trait – sweet hedonic response. These results suggest that reviews and meta-analyses on individual differences in optimal sucrose concentration can be conducted with studies using different methods. The choice of which method to use may depend on other factors, including target population. Regarding remote testing, we show that preference tracking via video conferencing is feasible. However, preference tracking requires focused, one-on-one interaction between an experimenter and participants, because the stimuli presented each trial depend on previous stimuli and responses. Therefore, some variant of the rating method, which can use pre-ordered sets of stimuli, might be more practical and efficient for remote testing.

References:

1. Drewnowski A. Taste preferences and food intake. *Annu Rev Nutr.* 1997;17:237-253.
2. Cheon E, Reister EJ, Hunter SR, Mattes RD. Finding the Sweet Spot: Measurement, Modification, and Application of Sweet Hedonics in Humans. *Adv Nutr.* 2021.
3. Macdonald IA. A review of recent evidence relating to sugars, insulin resistance and diabetes. *Eur J Nutr.* 2016;55(Suppl 2):17-23.
4. Yang Q, Zhang Z, Gregg EW, Flanders WD, Merritt R, Hu FB. Added sugar intake and cardiovascular diseases mortality among US adults. *JAMA Intern Med.* 2014;174(4):516-524.
5. Stanhope KL. Sugar consumption, metabolic disease and obesity: The state of the controversy. *Crit Rev Clin Lab Sci.* 2016;53(1):52-67.
6. WHO. Information note about intake of sugars recommended in the WHO guideline for adults and children. 2015.
7. White JR, Jr. Sugar. *Clin Diabetes.* 2018;36(1):74-76.
8. Appleton KM, Tuorila H, Bertenshaw EJ, de Graaf C, Mela DJ. Sweet taste exposure and the subsequent acceptance and preference for sweet taste in the diet: systematic review of the published literature. *The American Journal of Clinical Nutrition.* 2018;107(3):405-419.
9. Mattes RD. Gustation as a determinant of ingestion: methodological issues. *Am J Clin Nutr.* 1985;41(4):672-683.
10. Mattes RD, Mela DJ. Relationships between and among selected measures of sweet-taste preference and dietary intake. *Chemical Senses.* 1986;11(4):523-539.
11. Tan SY, Tucker RM. Sweet Taste as a Predictor of Dietary Intake: A Systematic Review. *Nutrients.* 2019;11(1).
12. Drewnowski A, Henderson SA, Levine A, Hann C. Taste and food preferences as predictors of dietary practices in young women. *Public Health Nutr.* 1999;2(4):513-519.
13. Pangborn RM, Giovanni ME. Dietary intake of sweet foods and of dairy fats and resultant gustatory responses to sugar in lemonade and to fat in milk. *Appetite.* 1984;5(4):317-327.
14. Weizenbaum F, Benson B, Solomon L, Brehony K. Relationship among reproductive variables, sucrose taste reactivity and feeding behavior in humans. *Physiol Behav.* 1980;24(6):1053-1056.
15. Iatridi V, Armitage RM, Yeomans MR, Hayes JE. Effects of Sweet-Liking on Body Composition Depend on Age and Lifestyle: A Challenge to the Simple Sweet-Liking-Obesity Hypothesis. *Nutrients.* 2020;12(9).
16. Garneau NL, Nuessle TM, Mendelsberg BJ, Shepard S, Tucker RM. Sweet liker status in children and adults: Consequences for beverage intake in adults. *Food Qual Prefer.* 2018;65:175-180.
17. Smith SL, Ludy MJ, Tucker RM. Changes in taste preference and steps taken after sleep curtailment. *Physiol Behav.* 2016;163:228-233.

18. Stevenson RJ, Boakes RA, Oaten MJ, Yeomans MR, Mahmut M, Francis HM. Chemosensory Abilities in Consumers of a Western-Style Diet. *Chem Senses*. 2016;41(6):505-513.
19. Holt SHA, Cobiac L, Beaumont-Smith NE, Easton K, Best DJ. Dietary habits and the perception and liking of sweetness among Australian and Malaysian students: A cross-cultural study. *Food Quality and Preference*. 2000;11(4):299-312.
20. Jayasinghe SN, Kruger R, Walsh DCI, et al. Is Sweet Taste Perception Associated with Sweet Food Liking and Intake? *Nutrients*. 2017;9(7).
21. Turner-McGrievy G, Tate DF, Moore D, Popkin B. Taking the bitter with the sweet: relationship of supertasting and sweet preference with metabolic syndrome and dietary intake. *J Food Sci*. 2013;78(2):S336-342.
22. Leong CS, Forde CG, Tey SL, Henry CJ. Taste perception and diet in people of Chinese ancestry. *Asia Pac J Clin Nutr*. 2018;27(2):478-486.
23. Pangborn RM, Simone M. Body size and sweetness preference. *J Am Diet Assoc*. 1958;34(9):924-928.
24. Iatridi V, Hayes JE, Yeomans MR. Quantifying Sweet Taste Liker Phenotypes: Time for Some Consistency in the Classification Criteria. *Nutrients*. 2019;11(1).
25. Armitage RM, Iatridi V, Yeomans MR. Understanding sweet-liking phenotypes and their implications for obesity: Narrative review and future directions. *Physiol Behav*. 2021;235:113398.
26. Trumbo PR, Appleton KM, de Graaf K, et al. Perspective: Measuring Sweetness in Foods, Beverages, and Diets: Toward Understanding the Role of Sweetness in Health. *Adv Nutr*. 2021;12(2):343-354.
27. Mennella JA, Lukasewycz LD, Griffith JW, Beauchamp GK. Evaluation of the Monell forced-choice, paired-comparison tracking procedure for determining sweet taste preferences across the lifespan. *Chem Senses*. 2011;36(4):345-355.
28. Kim J-Y, Prescott J, Kim K-O. Patterns of sweet liking in sucrose solutions and beverages. *Food Quality and Preference*. 2014;36:96-103.
29. Desor JA, Beauchamp GK. Longitudinal changes in sweet preferences in humans. *Physiol Behav*. 1987;39(5):639-641.
30. Bobowski N, Mennella JA. Personal Variation in Preference for Sweetness: Effects of Age and Obesity. *Child Obes*. 2017;13(5):369-376.
31. Coldwell SE, Mennella JA, Duffy VB, et al. Gustation assessment using the NIH Toolbox. *Neurology*. 2013;80(11 Suppl 3):S20-24.
32. Hayes JE, Allen AL, Bennett SM. Direct comparison of the generalized visual analog scale (gVAS) and general labeled magnitude scale (gLMS). *Food Quality and Preference*. 2013;28(1):36-44.
33. Dinnella C, Pierguidi L, Spinelli S, et al. Remote testing: sensory test during Covid-19 pandemic and beyond. *Food Quality and Preference*. 2021:104437.
34. Seo H-S, Buffin K, Singh A, Beekman TL, Jarma Arroyo SE. Stay safe in your vehicle: Drive-in booths can be an alternative to indoor booths for laboratory sensory testing. *Food Quality and Preference*. 2021;94:104332.
35. Vaira LA, Salzano G, Petrocelli M, Deiana G, Salzano FA, De Riu G. Validation of a self-administered olfactory and gustatory test for the remotely evaluation of COVID-19 patients in home quarantine. *Head Neck*. 2020;42(7):1570-1576.

36. Hannum ME, Simons CT. Development of the engagement questionnaire (EQ): A tool to measure panelist engagement during sensory and consumer evaluations. *Food Quality and Preference*. 2020;81:103840.
37. Bangcuyo RG, Smith KJ, Zumach JL, Pierce AM, Guttman GA, Simons CT. The use of immersive technologies to improve consumer testing: The role of ecological validity, context and engagement in evaluating coffee. *Food Quality and Preference*. 2015;41:84-95.
38. Samant SS, Seo H-S. Influences of sensory attribute intensity, emotional responses, and non-sensory factors on purchase intent toward mixed-vegetable juice products under informed tasting condition. *Food Research International*. 2020;132:109095.
39. Kong Y, Sharma C, Kanala M, et al. Virtual Reality and Immersive Environments on Sensory Perception of Chocolate Products: A Preliminary Study. *Foods*. 2020;9(4).
40. Venditti C, Musa-Veloso K, Lee HY, et al. Determinants of Sweetness Preference: A Scoping Review of Human Studies. *Nutrients*. 2020;12(3).
41. Wang G, Hayes JE, Ziegler GR, Roberts RF, Hopfer H. Dose-Response Relationships for Vanilla Flavor and Sucrose in Skim Milk: Evidence of Synergy. *Beverages*. 2018;4(4):73.
42. De Graaf C, Zandstra EH. Sweetness intensity and pleasantness in children, adolescents, and adults. *Physiol Behav*. 1999;67(4):513-520.
43. Bertino M, Beauchamp GK, Jen K-IC. Rated taste perception in two cultural groups. *Chemical Senses*. 1983;8(1):3-15.
44. Bartoshuk LM, Duffy VB, Green BG, et al. Valid across-group comparisons with labeled scales: the gLMS versus magnitude matching. *Physiol Behav*. 2004;82(1):109-114.
45. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*. 2015;67(1):1 - 48.
46. Asao K, Miller J, Arcori L, Lumeng JC, Han-Markey T, Herman WH. Patterns of Sweet Taste Liking: A Pilot Study. *Nutrients*. 2015;7(9):7298-7311.
47. Methven L, Xiao C, Cai M, Prescott J. Rejection thresholds (RjT) of sweet likers and dislikers. *Food Quality and Preference*. 2016;52:74-80.
48. Coulon SM, Miller AC, Reed JM, Martin CK. Reliability of a common solution-based taste perception test: implications for validity and a briefer test. *Eat Behav*. 2012;13(1):42-45.
49. Iatridi V, Hayes JE, Yeomans MR. Reconsidering the classification of sweet taste liker phenotypes: A methodological review. *Food Quality and Preference*. 2019;72:56-76.
50. Petrocelli M, Ruggiero F, Baietti AM, et al. Remote psychophysical evaluation of olfactory and gustatory functions in early-stage coronavirus disease 2019 patients: the Bologna experience of 300 cases. *J Laryngol Otol*. 2020;134(7):571-576.
51. Lee N, Lee J. Comparison of Home Use Tests with Differing Time and Order Controls. *Foods*. 2021;10(6).

52. Zhang MT, Jo Y, Lopetcharat K, Drake MA. Comparison of a central location test versus a home usage test for consumer perception of ready-to-mix protein beverages. *J Dairy Sci.* 2020;103(4):3107-3124.
53. Boutrolle I, Delarue J, Arranz D, Rogeaux M, Köster EP. Central location test vs. home use test: Contrasting results depending on product type. *Food Quality and Preference.* 2007;18(3):490-499.
54. Schouteten JJ, Gellynck X, Slabbinck H. Influence of organic labels on consumer's flavor perception and emotional profiling: Comparison between a central location test and home-use-test. *Food Res Int.* 2019;116:1000-1009.
55. Ervina E, Almli VL, Berget I, Spinelli S, Sick J, Dinnella C. Does Responsiveness to Basic Tastes Influence Preadolescents' Food Liking? Investigating Taste Responsiveness Segment on Bitter-Sour-Sweet and Salty-Umami Model Food Samples. *Nutrients.* 2021;13(8):2721.
56. Szczygiel EJ, Cho S, Tucker RM. The Effect of Sleep Curtailment on Hedonic Responses to Liquid and Solid Food. *Foods.* 2019;8(10):465.
57. Mennella JA, Bobowski NK. Psychophysical Tracking Method to Measure Taste Preferences in Children and Adults. *J Vis Exp.* 2016(113).
58. Spinelli S, Prescott J, Pierguidi L, et al. Phenol-Rich Food Acceptability: The Influence of Variations in Sweetness Optima and Sensory-Liking Patterns. *Nutrients.* 2021;13(3):866.
59. Drewnowski A, Shrager EE, Lipsky C, Stellar E, Greenwood MR. Sugar and fat: sensory and hedonic evaluation of liquid and solid foods. *Physiol Behav.* 1989;45(1):177-183.
60. Lavin JG, Lawless HT. Effects of color and odor on judgments of sweetness among children and adults. *Food Quality and Preference.* 1998;9(4):283-289.
61. 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.
62. Isogai T, Wise PM. The Effects of Odor Quality and Temporal Asynchrony on Modulation of Taste Intensity by Retronasal Odor. *Chemical Senses.* 2016;41(7):557-566.
63. Bertelsen AS, Mielby LA, Alexi N, Byrne DV, Kidmose U. Individual Differences in Sweetness Ratings and Cross-Modal Aroma-Taste Interactions. *Foods.* 2020;9(2):146.
64. Ponzio V, Pellegrini M, Costelli P, et al. Strategies for Reducing Salt and Sugar Intakes in Individuals at Increased Cardiometabolic Risk. *Nutrients.* 2021;13(1).
65. Thompson DA, Moskowitz HR, Campbell RG. Effects of body weight and food intake on pleasantness ratings for a sweet stimulus. *J Appl Physiol.* 1976;41(1):77-83.
66. Kampov-Polevoy A, Lange L, Bobashev G, Eggleston B, Root T, Garbutt JC. Sweet-liking is associated with transformation of heavy drinking into alcohol-related problems in young adults with high novelty seeking. *Alcohol Clin Exp Res.* 2014;38(7):2119-2126.
67. Desor JA, Beauchamp GK. Longitudinal changes in sweet preferences in humans. *Physiology & Behavior.* 1987;39(5):639-641.
68. Drewnowski A, Mennella JA, Johnson SL, Bellisle F. Sweetness and food preference. *J Nutr.* 2012;142(6):1142s-1148s.

69. Ebbeling CB, Feldman HA, Steltz SK, Quinn NL, Robinson LM, Ludwig DS. Effects of Sugar-Sweetened, Artificially Sweetened, and Unsweetened Beverages on Cardiometabolic Risk Factors, Body Composition, and Sweet Taste Preference: A Randomized Controlled Trial. *J Am Heart Assoc.* 2020;9(15):e015668.
70. Wise PM, Nattress L, Flammer LJ, Beauchamp GK. Reduced dietary intake of simple sugars alters perceived sweet taste intensity but not perceived pleasantness. *Am J Clin Nutr.* 2016;103(1):50-60.

Table 1. Test-retest reliabilities over two sessions (Pearson's r) for ratings of liking and hedonically optimal sucrose concentration

AS	Lab	Home
0.09 M	0.82	0.79
0.18 M	0.76	0.57
0.35 M	0.52	0.58
0.70 M	0.54	0.54
1.05 M	0.64	0.71
Pref ^a	0.71	0.73
Liked ^b	0.60	0.42
VM	Lab	Home
0.03 M	0.88	0.75
0.12 M	0.47	0.54
0.23 M	0.64	0.70
0.47 M	0.79	0.79
0.70 M	0.65	0.81
Pref ^a	0.79	0.86
Liked ^b	0.72	0.79

^a Most preferred, preference tracking^b Most liked, rating method

Table 2. Results of variance decomposition of hedonically optimal concentration

Source	Percent of variance
Location	0.00
Method	1.24
Stimulus	11.11
Session	0.00
Participant	51.00
Location X Method	0.00
Location X Stimulus	0.00
Location X Session	0.21
Location X Participant	0.42
Method X Stimulus	0.00
Method X Session	0.46
Method X Participant	2.82
Stimulus X Session	0.91
Stimulus X Participant	6.10
Session X Participant	3.73
Residual	22.00

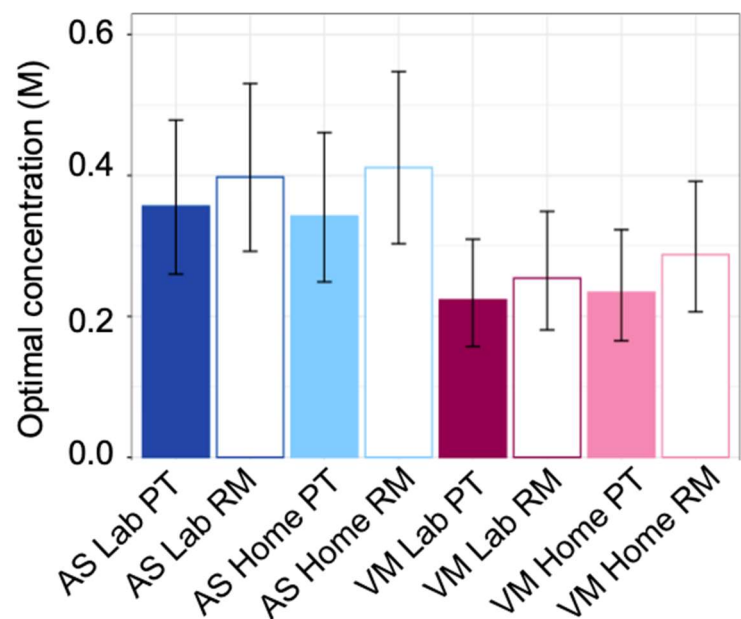


Figure 1. Optimal sucrose concentration (Molar; 95% confidence interval). Dark blue bars: aqueous solutions tested in laboratory. Light blue bars: data for aqueous solutions tested at home. Red bars: vanilla milk tested in laboratory. Pink bars: vanilla milk tested at home. Filled bars: most preferred concentration estimated from preference tracking. Open bars: most liked concentration estimated from the rating method.

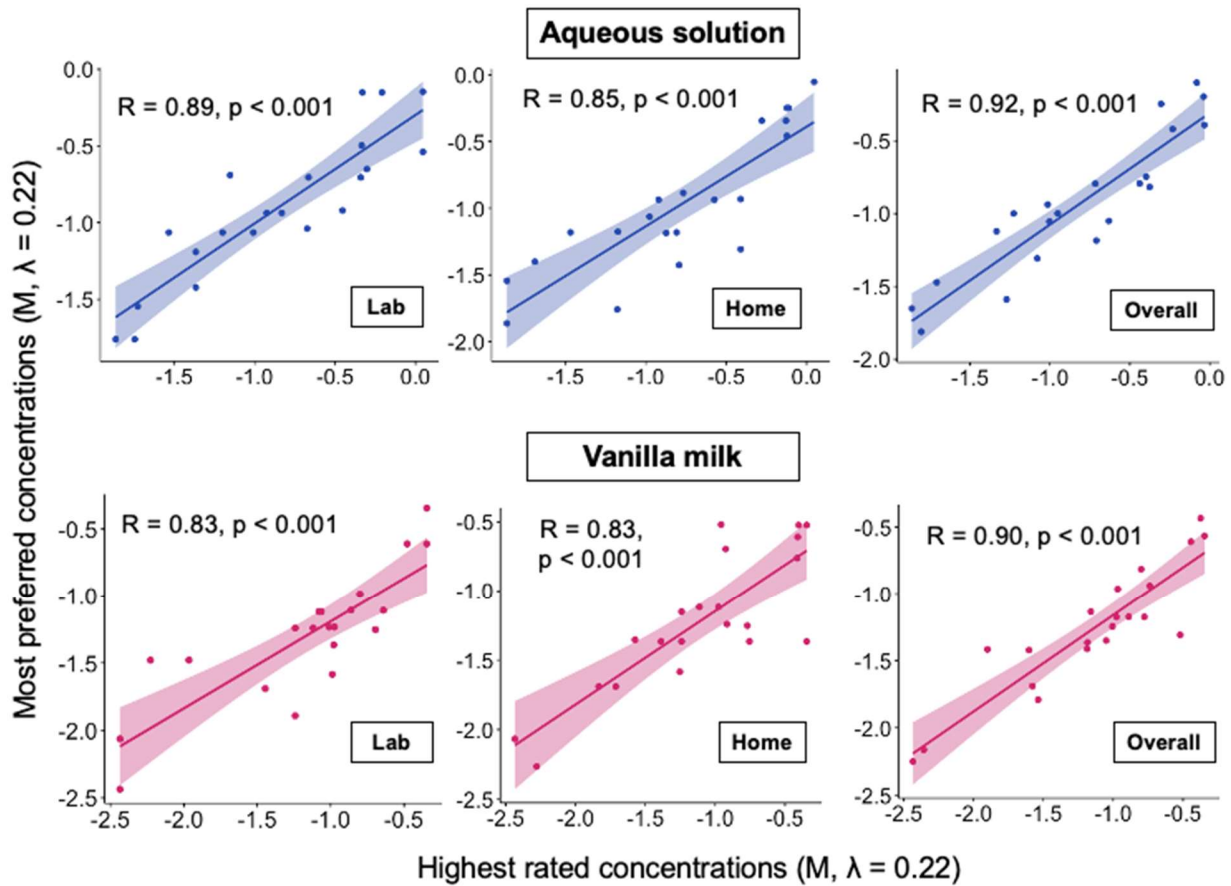


Figure 2. Scatter plots for most preferred sucrose concentration (averaged across repeated sessions) measured via paired-comparison preference tracking (M, Box-Cox transformation with $\lambda = 0.22$) vs. most liked (highest rated) concentration via ratings of liking. Top (blue points): data for aqueous solutions. Bottom (red points): vanilla milk. Left to right: Measured at home, measured in the lab, and averaged across the two test locations. Functions: linear fits via least-squared regression, with correlation coefficients and significance indicated in each panel.

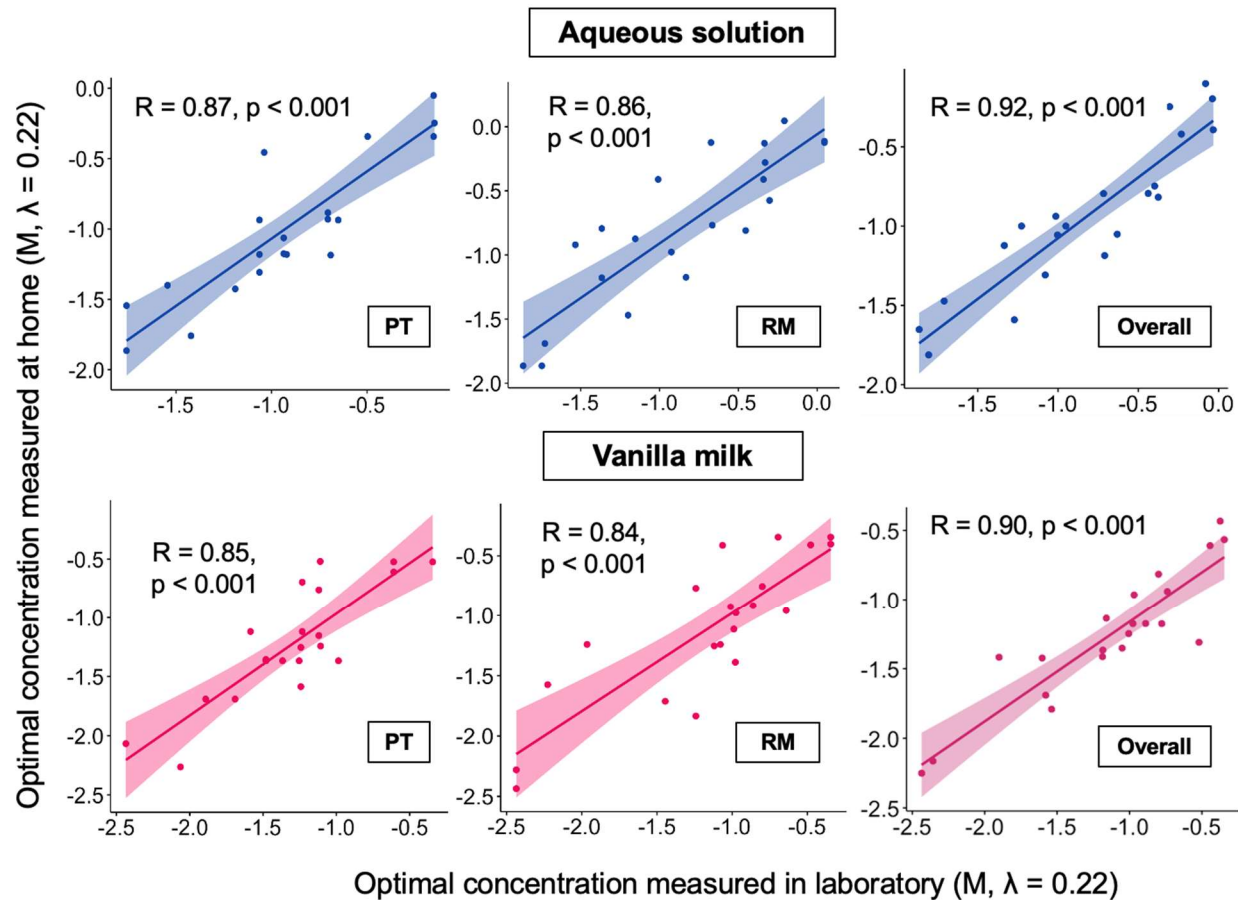
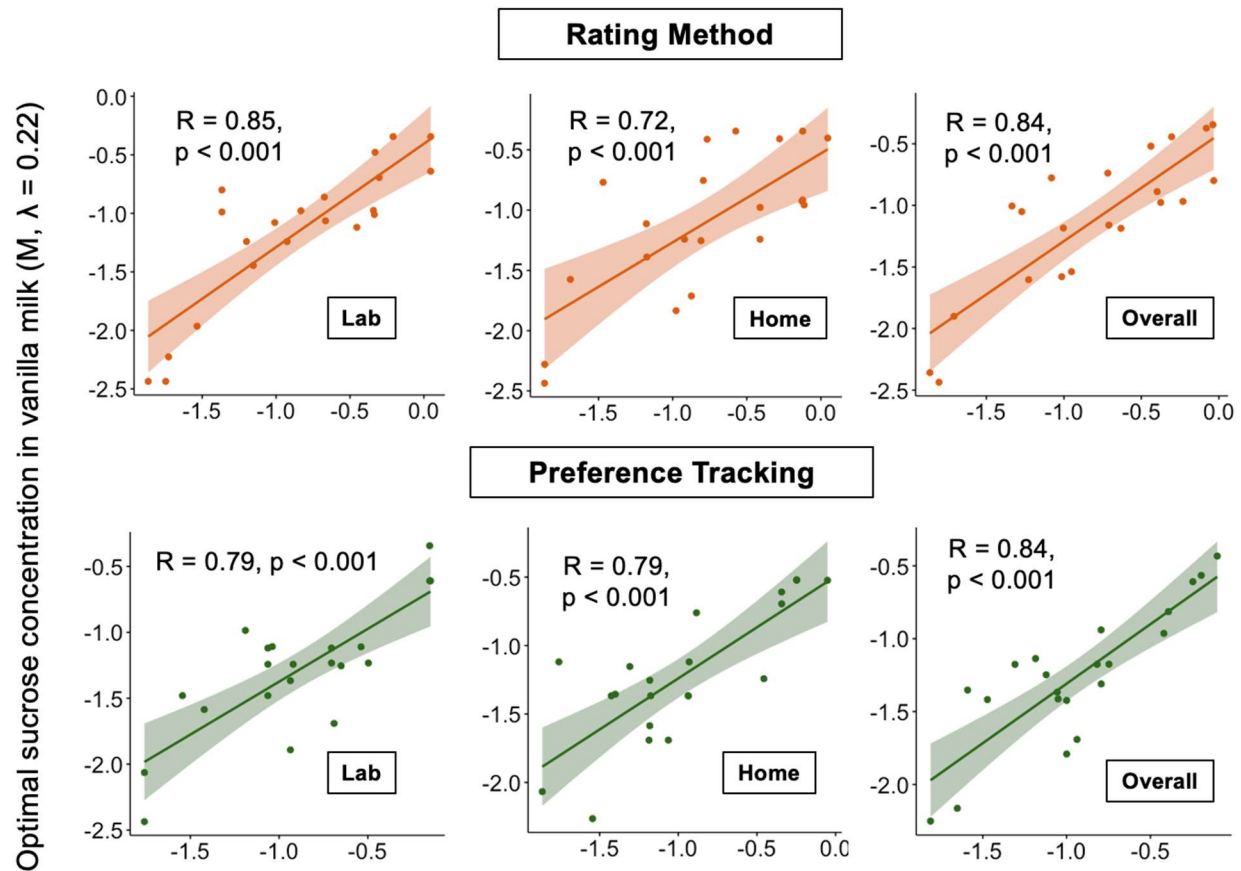


Figure 3. Scatter plots for optimal concentration measured in the laboratory (averaged across repeated sessions, M , Box-Cox transformation with $\lambda = 0.22$) vs. at home. Top (blue points): data for aqueous solutions. Bottom (red points): vanilla milk. Most preferred concentration estimated using preference tracking. Left to right: Measured using preference tracking, measured using rating method, and averaged across the two test methods. Functions: linear fits via least-squared regression, with correlation coefficients and significance indicated in each panel.



Optimal sucrose concentration in aqueous solution (M , $\lambda = 0.22$)

Figure 4. Scatter plots for optimal concentration measured using vanilla milk (averaged across repeated sessions, M , Box-Cox transformation with $\lambda = 0.22$) vs. hedonically optimal concentration measured using aqueous solutions. Top (orange points): Most liked concentration estimated from ratings of liking (rating method). Bottom (green points): most preferred concentration estimated using preference tracking. Left to right: Measured at home, measured in the lab, and averaged across the two test locations. Functions: linear fits via least-squared regression, with correlation coefficients and significance indicated in each panel.

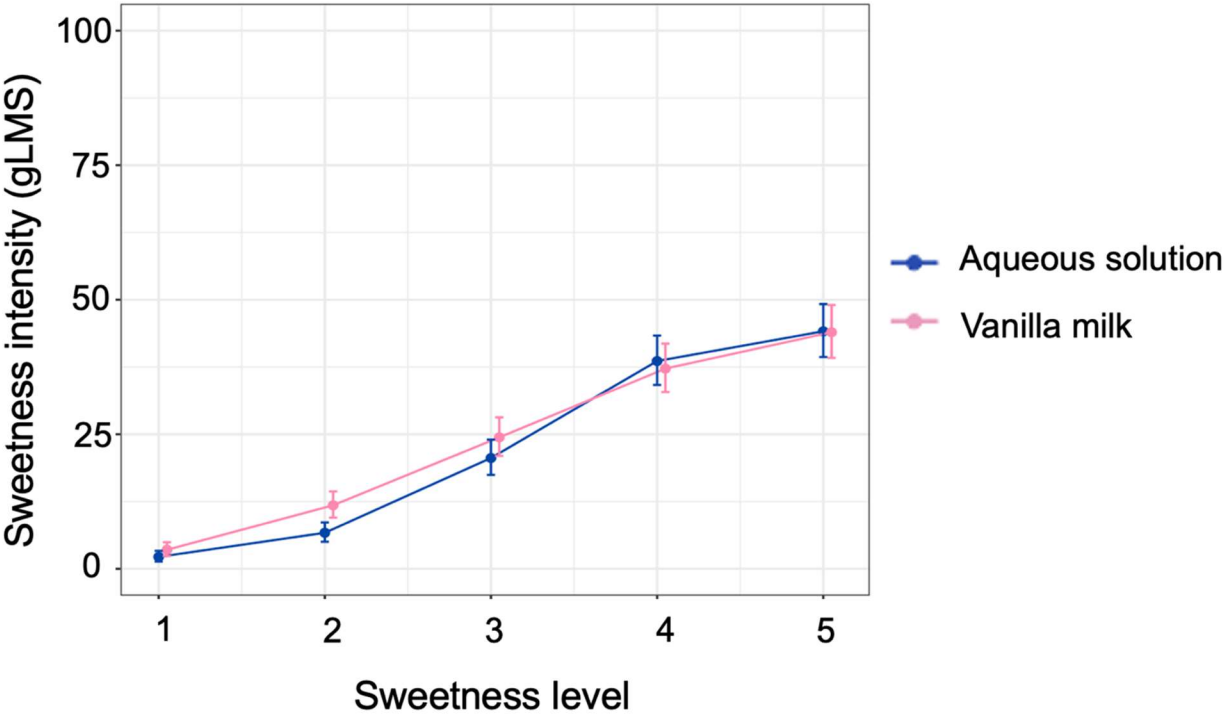
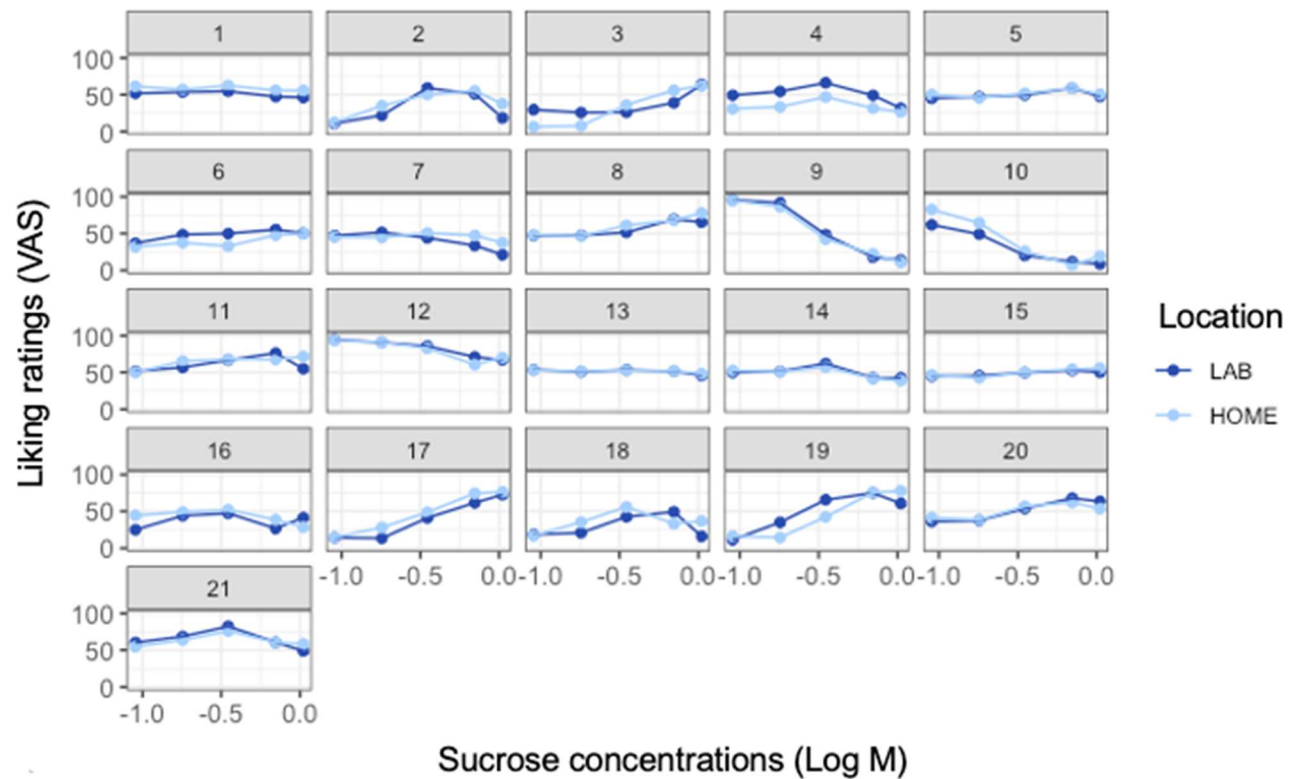
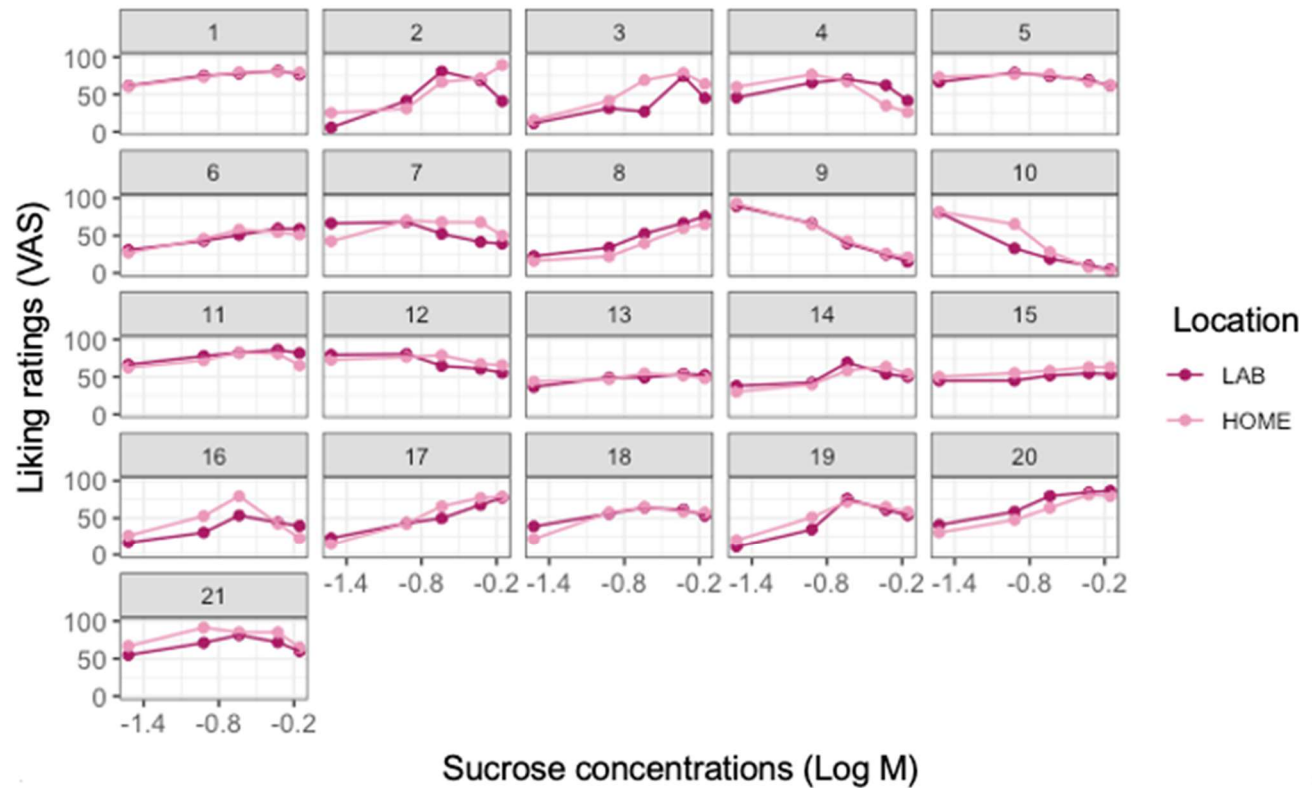


Figure 5. Average back-transformed ratings of sweetness intensity made using the general Labeled Magnitude Scale (gLMS). Error bars represents 95% confidence intervals on means. Data are separated by stimulus, averaged across other conditions (filled symbols for aqueous solutions, pink symbols for vanilla milk).

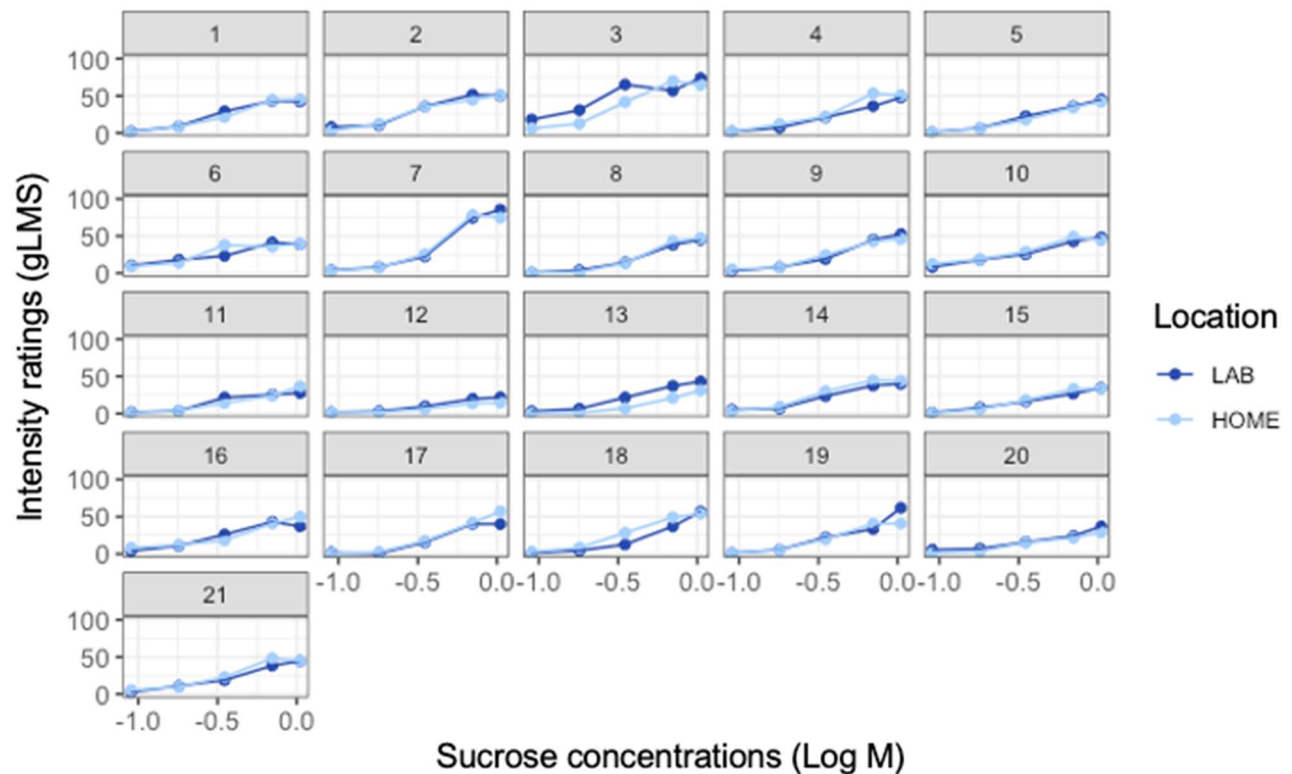
Supplementary figures



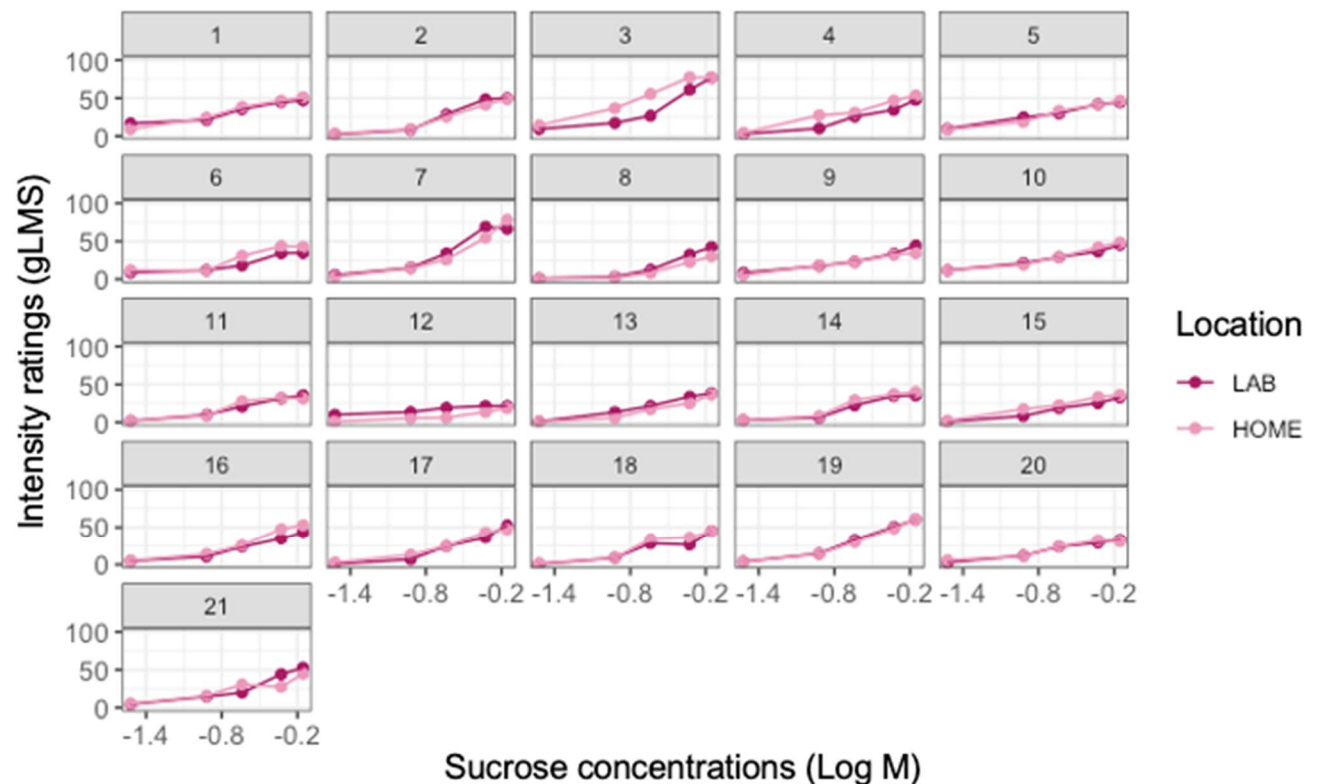
Supplemental figure 1. Liking vs. concentration functions for individual participants, AS. Rated liking for aqueous solutions for individual participants, averaged across repeated test sessions. Numbers at top of each panel are participant ID number. X-axis: sucrose concentration, log M. Dark blue points and curves: measured in the laboratory. Light blue points and curves: measured at home.



Supplemental figure 2. Liking vs. concentration functions for individual participants, VM. Rated liking for vanilla milk for individual participants, averaged across repeated test sessions. Numbers at top of each panel are participant ID number. X-axis: sucrose concentration, log M. Red blue points and curves: measured in the laboratory. Pink points and curves: measured at home.



Supplemental figure 3. Intensity vs. concentration functions for individual participants, AS. Rated intensity of aqueous solutions for individual participants, averaged across repeated test sessions. Numbers at top of each panel are participant ID number. X-axis: sucrose concentration, log M. Dark blue points and curves: measured in the laboratory. Light blue points and curves: measured at home.



Supplement figure 4. Intensity vs. concentration functions for individual participants, VM. Rated intensity of aqueous solutions for individual participants, averaged across repeated test sessions. Numbers at top of each panel are participant ID number. X-axis: sucrose concentration, log M. Red points and curves: measured in the laboratory. Pink points and curves: measured at home.

Supplementary Table 1. Parameters for fits of rated liking vs. sucrose concentration to determine hedonically optimal (most liked) concentration.

Participant	Stimulus	Location	Session	Intercept	Linear	Quadratic	Cubic	r ²	Optimal
1	AS	HOME	1	59.72	0.00	0.00	0.00	0.00	0.18
1	AS	HOME	2	57.35	0.00	0.00	0.00	0.00	0.86
1	AS	LAB	1	44.63	-6.64	0.00	0.00	0.30	0.09
1	AS	LAB	2	59.20	0.00	-2.70	0.00	0.21	0.38
1	VM	HOME	1	82.71	8.52	-6.56	0.00	0.87	0.43
1	VM	HOME	2	75.23	4.06	-2.98	0.00	0.57	0.44
1	VM	LAB	1	83.74	0.00	-8.16	0.00	0.36	0.23
1	VM	LAB	2	75.37	7.57	-3.05	0.00	0.75	0.69
2	AS	HOME	1	51.94	-28.44	-20.73	18.74	0.67	0.25
2	AS	HOME	2	53.95	60.27	-12.09	-24.87	0.87	0.70
2	AS	LAB	1	62.67	48.38	-21.76	-24.22	0.82	0.61
2	AS	LAB	2	46.07	0.00	-25.30	0.00	0.72	0.38
2	VM	HOME	1	46.83	31.48	0.00	0.00	0.58	0.70
2	VM	HOME	2	65.32	33.54	0.00	-10.04	0.80	0.60
2	VM	LAB	1	71.76	60.04	-38.96	-32.92	0.92	0.37
2	VM	LAB	2	73.37	0.00	-20.14	8.84	0.71	0.23
3	AS	HOME	1	32.89	27.82	0.00	0.00	0.65	1.05
3	AS	HOME	2	36.54	56.23	0.00	-25.47	0.62	0.76
3	AS	LAB	1	17.73	0.00	23.04	0.00	0.42	1.05
3	AS	LAB	2	35.03	23.52	0.00	0.00	0.75	1.05
3	VM	HOME	1	71.53	0.00	-18.51	0.00	0.27	0.23
3	VM	HOME	2	63.72	56.59	-15.05	-18.48	0.88	0.48
3	VM	LAB	1	59.23	33.03	-27.73	-19.68	0.59	0.35
3	VM	LAB	2	29.26	49.96	9.07	-12.97	0.91	0.70
4	AS	HOME	1	22.57	0.00	0.00	0.00	0.00	0.49
4	AS	HOME	2	51.88	-8.77	-7.55	0.00	0.62	0.21
4	AS	LAB	1	54.08	-8.41	-12.15	0.00	0.82	0.27
4	AS	LAB	2	69.88	6.40	-13.82	-6.14	0.91	0.45
4	VM	HOME	1	67.92	-32.58	-10.78	9.71	0.82	0.09
4	VM	HOME	2	60.08	-36.81	-10.47	12.73	0.92	0.10
4	VM	LAB	1	71.43	0.00	-13.07	-6.58	0.65	0.23
4	VM	LAB	2	70.79	0.00	-19.09	0.00	0.55	0.23
5	AS	HOME	1	56.32	23.39	0.00	-12.34	0.71	0.72
5	AS	HOME	2	47.56	0.00	0.00	0.00	0.00	0.18
5	AS	LAB	1	63.39	14.42	-8.65	-3.61	0.94	0.63
5	AS	LAB	2	44.01	0.00	0.00	-3.00	0.29	0.09
5	VM	HOME	1	75.16	-4.93	-5.98	0.00	0.39	0.15
5	VM	HOME	2	75.89	-5.63	-3.61	0.00	0.83	0.09
5	VM	LAB	1	77.59	0.00	-5.68	0.00	0.27	0.23
5	VM	LAB	2	74.23	-9.69	-6.67	3.22	0.92	0.13
6	AS	HOME	1	38.88	0.00	8.83	0.00	0.30	1.05
6	AS	HOME	2	36.99	25.27	-3.47	-9.65	0.90	0.74
6	AS	LAB	1	50.16	0.00	0.00	0.00	0.00	0.35
6	AS	LAB	2	53.50	10.09	-7.98	0.00	0.74	0.64
6	VM	HOME	1	53.03	7.13	-7.38	0.00	0.79	0.37
6	VM	HOME	2	57.76	8.23	-10.28	0.00	0.68	0.35
6	VM	LAB	1	60.73	0.00	-6.56	3.91	0.80	0.23
6	VM	LAB	2	42.02	27.74	0.00	-8.18	0.93	0.60

Supplementary Table 1 (continued). Parameters for fits of rated liking vs. sucrose concentration to determine optimal (most liked) concentration.

Participant	Stimulus	Location	Session	Intercept	Linear	Quadratic	Cubic	r ²	Optimal
7	AS	HOME	1	44.22	0.00	0.00	0.00	0.00	0.49
7	AS	HOME	2	45.87	0.00	0.00	0.00	0.00	0.25
7	AS	LAB	1	31.50	-13.95	0.00	0.00	0.80	0.09
7	AS	LAB	2	54.13	-6.57	-7.40	0.00	0.53	0.25
7	VM	HOME	1	66.37	0.00	-11.71	0.00	0.40	0.23
7	VM	HOME	2	78.54	0.00	-16.60	0.00	0.56	0.23
7	VM	LAB	1	36.54	-32.11	14.55	13.11	0.78	0.03
7	VM	LAB	2	72.70	-12.35	-14.99	0.00	0.85	0.15
8	AS	HOME	1	59.88	11.76	0.00	0.00	0.71	1.05
8	AS	HOME	2	58.03	12.27	3.25	0.00	0.90	1.05
8	AS	LAB	1	54.25	18.44	0.00	-7.76	0.83	0.78
8	AS	LAB	2	59.23	20.05	0.00	-6.60	0.91	0.84
8	VM	HOME	1	44.07	32.99	0.00	-8.05	0.89	0.65
8	VM	HOME	2	35.52	31.79	0.00	-7.56	0.96	0.66
8	VM	LAB	1	47.59	23.27	0.00	0.00	0.86	0.70
8	VM	LAB	2	53.05	18.06	0.00	0.00	0.86	0.70
9	AS	HOME	1	42.35	-55.69	5.56	13.19	0.97	0.09
9	AS	HOME	2	54.38	-32.87	0.00	0.00	0.71	0.09
9	AS	LAB	1	51.20	-61.50	0.00	17.48	0.88	0.12
9	AS	LAB	2	53.41	-53.39	0.00	13.16	0.92	0.10
9	VM	HOME	1	54.79	-25.07	0.00	0.00	0.88	0.03
9	VM	HOME	2	37.01	-28.95	7.86	0.00	0.97	0.03
9	VM	LAB	1	48.38	-29.94	0.00	0.00	0.87	0.03
9	VM	LAB	2	45.66	-27.15	0.00	0.00	0.74	0.03
10	AS	HOME	1	23.16	-26.38	20.91	0.00	0.82	0.09
10	AS	HOME	2	27.25	-47.81	10.69	11.77	0.94	0.09
10	AS	LAB	1	31.26	-16.20	0.00	0.00	0.42	0.09
10	AS	LAB	2	13.63	-27.83	17.51	0.00	0.98	0.09
10	VM	HOME	1	46.16	-63.14	0.00	18.07	0.97	0.06
10	VM	HOME	2	20.94	-36.47	11.32	5.59	0.98	0.03
10	VM	LAB	1	23.66	-26.61	10.72	0.00	0.93	0.03
10	VM	LAB	2	13.44	-7.83	12.30	-10.12	0.99	0.03
11	AS	HOME	1	65.88	0.00	0.00	6.22	0.76	1.05
11	AS	HOME	2	69.68	0.00	-8.14	3.72	0.52	0.38
11	AS	LAB	1	79.85	20.55	-19.32	-15.73	0.79	0.52
11	AS	LAB	2	61.68	17.78	0.00	-4.48	0.95	0.93
11	VM	HOME	1	82.93	3.98	-6.84	0.00	0.67	0.31
11	VM	HOME	2	81.05	21.20	-15.84	-14.77	0.55	0.35
11	VM	LAB	1	93.16	0.00	-6.03	0.00	0.41	0.23
11	VM	LAB	2	74.54	19.27	-5.71	-6.75	0.79	0.46
12	AS	HOME	1	75.91	-37.64	0.00	17.63	0.58	0.16
12	AS	HOME	2	82.06	-11.05	0.00	0.00	0.65	0.09
12	AS	LAB	1	88.32	-20.06	-2.95	5.53	0.96	0.14
12	AS	LAB	2	77.65	-10.26	0.00	0.00	0.95	0.09
12	VM	HOME	1	62.03	0.00	0.00	0.00	0.00	0.40
12	VM	HOME	2	86.68	-11.82	-4.54	0.00	0.76	0.04
12	VM	LAB	1	65.79	-9.01	0.00	0.00	0.63	0.03
12	VM	LAB	2	71.82	-25.31	0.00	9.19	0.79	0.07

Supplementary Table 1 (continued). Parameters for fits of rated liking vs. sucrose concentration to determine optimal (most liked) concentration.

Participant	Stimulus	Location	Session	Intercept	Linear	Quadratic	Cubic	r ²	Optimal
13	AS	HOME	1	57.55	6.76	-3.34	-3.70	0.56	0.59
13	AS	HOME	2	48.60	-3.99	0.00	0.00	0.60	0.09
13	AS	LAB	1	56.55	0.00	-4.25	0.00	0.34	0.38
13	AS	LAB	2	49.53	0.00	0.00	-3.01	0.81	0.09
13	VM	HOME	1	48.46	0.00	0.00	0.00	0.00	0.33
13	VM	HOME	2	54.10	0.00	-5.02	0.00	0.55	0.23
13	VM	LAB	1	59.79	-10.75	-5.45	7.10	0.90	0.13
13	VM	LAB	2	42.50	10.44	0.00	0.00	0.64	0.70
14	AS	HOME	1	46.00	-6.91	0.00	0.00	0.76	0.09
14	AS	HOME	2	56.68	-3.89	-6.92	0.00	0.56	0.29
14	AS	LAB	1	56.71	-11.59	-7.82	6.26	0.58	0.24
14	AS	LAB	2	53.56	-3.95	-4.92	0.00	0.70	0.26
14	VM	HOME	1	54.28	28.28	-8.10	-10.08	0.83	0.47
14	VM	HOME	2	58.69	19.52	-10.77	-8.33	0.78	0.40
14	VM	LAB	1	62.57	0.00	-12.68	0.00	0.37	0.23
14	VM	LAB	2	55.44	0.00	-5.88	0.00	0.26	0.23
15	AS	HOME	1	52.35	13.09	0.00	-3.65	0.89	0.89
15	AS	HOME	2	47.79	0.00	0.00	0.00	0.00	0.31
15	AS	LAB	1	51.87	4.05	0.00	0.00	0.65	1.05
15	AS	LAB	2	45.13	0.00	0.00	0.00	0.00	0.49
15	VM	HOME	1	56.03	6.42	0.00	0.00	0.67	0.70
15	VM	HOME	2	59.71	3.71	0.00	0.00	0.26	0.70
15	VM	LAB	1	52.64	15.27	0.00	-5.76	0.64	0.55
15	VM	LAB	2	46.73	0.00	0.00	0.00	0.00	0.40
16	AS	HOME	1	45.28	22.58	-12.22	-10.21	0.90	0.60
16	AS	HOME	2	51.37	-20.32	0.00	0.00	0.65	0.09
16	AS	LAB	1	35.18	0.00	0.00	0.00	0.00	0.61
16	AS	LAB	2	38.44	0.00	0.00	0.00	0.00	0.25
16	VM	HOME	1	60.50	-11.80	-15.60	0.00	0.51	0.15
16	VM	HOME	2	74.07	0.00	-34.14	0.00	0.84	0.23
16	VM	LAB	1	57.95	0.00	-17.15	0.00	0.37	0.23
16	VM	LAB	2	36.58	29.73	-7.38	-12.44	0.62	0.46
17	AS	HOME	1	62.17	34.65	-11.18	-7.84	0.96	0.74
17	AS	HOME	2	45.32	28.47	0.00	0.00	0.96	1.05
17	AS	LAB	1	45.53	23.89	0.00	0.00	0.79	1.05
17	AS	LAB	2	29.68	34.66	7.13	-6.68	0.99	1.05
17	VM	HOME	1	57.15	22.32	-6.93	0.00	0.85	0.70
17	VM	HOME	2	69.16	24.46	-9.40	0.00	0.78	0.70
17	VM	LAB	1	53.07	16.07	0.00	0.00	0.54	0.70
17	VM	LAB	2	51.40	23.77	0.00	0.00	0.87	0.70
18	AS	HOME	1	63.25	-21.01	-20.35	17.07	0.84	0.27
18	AS	HOME	2	32.16	8.84	-8.09	0.00	0.61	0.60
18	AS	LAB	1	51.96	31.97	-24.57	-18.84	0.80	0.55
18	AS	LAB	2	31.64	25.21	0.00	-14.98	0.35	0.70
18	VM	HOME	1	60.95	-24.38	-6.75	20.24	0.89	0.13
18	VM	HOME	2	66.03	24.83	-16.87	-8.09	0.96	0.39
18	VM	LAB	1	65.88	28.41	-12.57	-9.24	0.82	0.44
18	VM	LAB	2	53.68	0.00	0.00	0.00	0.00	0.17

Supplementary Table 1 (continued). Parameters for fits of rated liking vs. sucrose concentration to determine optimal (most liked) concentration.

Participant	Stimulus ¹	Location ²	Session ³	Intercept ⁴	Linear ⁵	Quadratic ⁶	Cubic ⁷	r ² ⁸	Optimal ⁹
19	AS	HOME	1	45.74	49.97	0.00	-11.52	0.94	0.96
19	AS	HOME	2	47.21	48.41	0.00	-18.45	0.87	0.80
19	AS	LAB	1	85.23	14.88	-34.26	0.00	0.94	0.46
19	AS	LAB	2	44.68	28.78	0.00	0.00	0.86	1.05
19	VM	HOME	1	46.48	0.00	0.00	0.00	0.00	0.33
19	VM	HOME	2	76.06	18.21	-17.81	0.00	0.89	0.38
19	VM	LAB	1	60.83	0.00	-20.68	0.00	0.34	0.23
19	VM	LAB	2	68.36	45.58	-20.08	-17.20	0.77	0.42
20	AS	HOME	1	48.48	24.60	0.00	-15.15	0.37	0.69
20	AS	HOME	2	59.12	11.89	-6.20	0.00	0.74	0.81
20	AS	LAB	1	48.83	41.24	0.00	-14.95	0.98	0.81
20	AS	LAB	2	60.51	6.23	-5.35	0.00	0.48	0.62
20	VM	HOME	1	64.43	28.63	0.00	-8.32	0.92	0.61
20	VM	HOME	2	55.71	23.90	0.00	0.00	0.66	0.70
20	VM	LAB	1	74.13	39.30	-7.29	-11.82	0.89	0.52
20	VM	LAB	2	77.69	5.88	-6.86	4.21	0.97	0.70
21	AS	HOME	1	70.73	0.00	-3.74	4.01	0.78	1.05
21	AS	HOME	2	71.51	0.00	-15.37	0.00	0.45	0.38
21	AS	LAB	1	76.70	0.00	-10.81	0.00	0.32	0.38
21	AS	LAB	2	75.34	-11.69	-15.49	0.00	0.69	0.27
21	VM	HOME	1	93.62	0.00	-10.01	0.00	0.59	0.23
21	VM	HOME	2	86.65	0.00	-15.55	0.00	0.32	0.23
21	VM	LAB	1	77.46	17.91	-13.02	-11.28	0.44	0.36
21	VM	LAB	2	81.34	0.00	-13.98	0.00	0.35	0.23

¹ Model beverage. AS = aqueous solutions; VM = vanilla milk.

² Test location. HOME = test conducted at home; Lab = test conducted in the laboratory.

³ Replicate sessions (first and second).

⁴ Model fit parameter, intercept..

⁵ Model fit parameter, linear term. A value of 0 indicates that the best fit model did not include a linear term.

⁶ Model fit parameter, quadratic term. A value of 0 indicates that the best fit model did not include a quadratic term.

⁷ Model fit parameter, cubic term. A value of 0 indicates that the best fit model did not include a cubic term.

⁸ Proportion of variance for which the model fit accounts. A value of 0 indicates an intercept-only model.

⁹ Most liked concentration of sucrose (M). Local maximum of fitted function within range of presented concentrations. For intercept-only fits, this value is the geometric mean of the two highest rated concentrations.