AN ABSTRACT OF THE THESIS OF

Samuel J. Hammond for the degree of Honors Baccalaureate of Science in Food Science and Technology presented on March 24th, 2010. Title: Evaluation of the Labeled Hedonic Scale (LHS): Obtaining higher level data in consumer testing for product development research

Abstract Approved: ______________________________

Juyun Lim

Among sensory and food scientists, the 9-point hedonic scale has been the most commonly used scale for measuring liking and disliking of products. Recently, the Labeled Hedonic Scale (LHS) was developed to overcome some recognized limitations of the 9-point hedonic scale. One of the claimed advantages of the LHS is the ability to provide ratio-level data, which can be extremely useful when making product comparisons. The current study was aimed to confirm that the LHS can produce ratio-level data in a product development setting, where samples only differ slightly, by comparing it with magnitude estimation (ME), which yields ratio-level data. Subjects (N= 40, 12 M, 28 F, Age range: 19-32) attended two separate testing sessions. During each session, they used one of the two scaling methods (i.e., LHS, ME) to rate their liking and disliking of two product systems (cherry flavored Kool-Aid and vanilla custard) with varying sucrose concentrations (0.14, 0.20, 0.28, 0.40, 0.56 M). The results indicated that the LHS yielded data that were not significantly different to that obtained using ME for both product systems (repeated measures ANOVA, $p>0.05$), implying that the LHS, in fact, produced ratio-level data. The results also indicated that the LHS offered slightly better discrimination power than ME. The present study demonstrates the potential utility of the LHS as a tool in food development that enables sensory scientists to make statements about proportional (ratio) differences in liking and disliking among samples.

Keywords: labeled hedonic scale (LHS), consumer testing, ratio-level data, discrimination power, product development.

Corresponding e-mail address: hammonds@onid.orst.edu

Evaluation of the Labeled Hedonic Scale (LHS): Obtaining higher level data in consumer testing for product development research

By

Samuel J. Hammond

A PROJECT

submitted to

Oregon State University

University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Food Science and Technology
(Honors Associate)

Presented (March 24th, 2010)
Commencement June 2010

Honors Baccalaureate of Science in Food Science and Technology project of Samuel J. Hammond presented on (March 24th, 2010)

APPROVED:

____

Mentor, department of Food Science and Technology

____

Committee Member, department of Food Science and Technology

____

Committee Member, department of Food Science and Technology

____

Chair, Department of Food Science and Technology

____

Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

____

Samuel J. Hammond, Author

ACKNOWLEDGEMENT

First of all, I would like to thank my mentor, Dr. Juyun Lim, for assisting me throughout my entire thesis process. She pushed me to do my best, and I appreciate all her valuable time and effort she put into me succeeding. I would also like to thank my committee members, Dr. Tom Shellhammer and Dr. Robert McGorrin for their contributions to my thesis.

I owe a great deal of thanks to my advisor, Dr. J. Antonio Torres, who welcomed me from day one into the Food Science program and helped me every step of the way, even when I had challenges outside of school.

Everyone, faculty and students alike, in the Food Science department here at Oregon State University has helped me at some point throughout my career here, and I thank you all for your contributions to my success.

Lastly, but surely not least, I would like to thank my parents, Sam and Marla Hammond. They are a constant source of encouragement and have always believed in me. I owe everything in my life to them; thank you.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

INTRODUCTION

During product development, consumer testing of newly finished products is often completed before the product can go into commercial production (Resurreccion, 1998; Lawless & Heymann, 1998). Product development can also consist of optimizing already existing food formulations based on changes in taste, flavor and/or ingredients for a variety of different reasons (e.g., ingredient change to decrease cost). Newly optimized products are often compared to their previous formulations in consumer sensory tests. Depending on objectives, sensory scientists conduct consumer testing using one of two testing methods: preference or acceptance test (Jellinek, 1964; Lawless & Heymann, 1998). Preference measurement involves the panelist choosing a single product over one or more products. Acceptance measurement involves the panelist using a scale to measure the degree of hedonic value (i.e. liking/disliking) of a product. Both types of measurement are commonly used in various situations due to its own advantages. Preference testing can be extremely useful when testing new modifications of a known, widely accepted product (e.g., Coke®) against the original formulation. Often these situations do not require hedonic values from acceptance testing because the company already knows consumers accept their product. However, results from preference test alone can be misleading in situations where a company does not have previous consumer acceptance information. Data indicating that one product is preferred over another does not necessarily indicate acceptance of the product. If both products are disliked then the preferred product will be the least disliked of the two products, but unaccepted none the less. For this reason, it is beneficial for company to consider using acceptance measurement during consumer testing of newly formulated products.

Acceptance testing involves the use of scaling to determine the degree of hedonic value of a product. The method of scaling is based on applying numbers to quantify, or measure, sensory experiences (Lawless & Heymann, 1998). Measurement theory (Stevens, 1951) explains that applied numbers can be assigned to the sensory experience in different ways. The four common levels of measurement are nominal, ordinal, interval, and ratio scaling (Stevens, 1951). Nominal scaling refers to the method of assigning numbers as labels (e.g., 1 for like, 2 for dislike, 3 for neutral). Ordinal scaling refers to the method of assigning numbers as a rank order (e.g., 1, 2, 3 in an order of hedonic reaction, 1 being most liked 3 being least liked). Interval scaling refers to the method of having responses with equal subjective spacing and an arbitrary zero point so that assigned numbers indicate degrees of difference (e.g., a 1 to 9 hedonic scale, 1 being least liked and 9 being most liked, where a difference between 2 and 3 is equal to the difference between 8 and 9). Ratio scaling refers to the method of assigning numbers reflecting relative proportions with a true zero point. (e.g., a product rated 20 is two times more liked than a product rated 10). These four levels of measurement are useful in different situations and will yield nominal, ordinal, interval and ratio-level data respectively. As a scale advances from nominal to ratio level measurement, the data becomes more powerful. It also has been proposed that the types of statistical analysis permissible with each data level are different (Stevens, 1951; Townsend and Ashby, 1980).

Interval-level data has been the most common type of data obtained during consumer acceptance testing (e.g., category scaling). It offers the advantage of numbers taking on significance. This significance offers preference rank (ordinal data) as well as a numerical difference between products, allowing researchers to know degrees of difference. Interval scaling also allows the use of parametric statistics, such as t-tests and analysis of variance (Stevens, 1951; Townsend and Ashby, 1980). However, scales often used for interval measurement are *only assumed* to yield interval-level data on the basis of equal spacing between ratings (Lawless & Heymann, 1998). Without the equal spacing of an interval scale, the assumptions for the use of parametric statistics are invalid. Another disadvantage is the arbitrary zero point of interval scaling. This prevents ratio comparisons of rated degree of liking/disliking for products. Ratio scaling, however, offers some advantages over interval scaling that are potentially useful in product development research. In addition to the ability to determine preference rank and degree of difference, ratio scaling can obtain the relative magnitudes of products which can be directly compared using proportions (e.g., a product rated 40 is 4 times more liked than a product rated 10). This capability is attributed to ratio scaling providing a true zero point. The ability to determine how much more a product is liked or disliked over another product allows product developers to fine tune, or optimize, their formulations (e.g., amount of sugar for added sweetness, amount of salt for added saltiness) to enhance acceptance level of products for their potential consumers. However, it is difficult to establish that a scale is providing ratio-level data. Ratio scaling also tends to be more complicated for the panelist to use properly.

Among sensory and food scientists, the 9-point hedonic scale (Peryam & Giradot, 1952; Peryam & Pilgrm, 1957) has been the most commonly used scale for measuring liking and disliking of products. This category scale was first invented in the 1940s at the Food Research Division of the Quartermaster Food and Container Institute in Chicago

(Peryam & Giradot, 1952). It consists of nine categories: four positive, four negative, and one neutral. The descriptors on the scale used to indicate degree of liking or disliking are slightly, moderately, very much, and extremely (e.g., like slightly, dislike extremely). The 9-point hedonic scale has many positive attributes pertaining to consumer testing. The scale is very simple and therefore easy to use for both consumers and researchers due to its categorical nature and limited choices. Secondly, it also provides semantic information of a rated product. While the 9-point hedonic scale has positive aspects, it also has many limitations due to its simplicity. All category scales are prone to ceiling effects; panelists' tendency to avoid using the extreme categories (Hollingworth 1910; Moskowitz 1982). This reduces the discrimination power of the scale for highly liked or disliked samples and effectively shrinks the scale down to seven categories (Lim et al., 2009; Peryam et. al., 1960). Also, the categories used are shown equally spaced but could be psychologically unequal (Moskowitz 1980). Without equally spaced categories, the scale would provide ordinal instead of interval-level data which should not be subjected to parametric statistics even though most researchers will still perform the analysis with the data obtained (Gay and Mead 1992; Villanueva et al. 2000; Lim et al. 2009; Peryam et al., 1960). The arbitrary zero point of the 9-point hedonic scale also prevents ratio comparisons between ratings of products (Lawless & Heymann, 1998). Sensory scientists aware of the limitations of the 9-point hedonic scale have continued to adapt strategies to design new scales that overcome some of its limitations.

The method of Magnitude Estimation (ME) was originally developed to measure proportional magnitudes of perceived sensations but has since been adapted to rate the hedonic values of food samples (Stevens, 1956; Stevens & Galanter, 1957; Moskowitz, 1971, 1982; Moskowitz & Sidel, 1971; Engen & McBurney 1964). This scale allows consumers to freely assign numbers that reflect the ratio of hedonic perception (Moskowitz 1977). For example, if product A is given a value of 50 for liking and product B is liked twice as much then product B would be given a magnitude of 100. The ME method offers advantages over traditional category scales. First of all, ME has a true zero point which enables ratio comparisons between ratings. Secondly, the open ended nature of ME reduces ceiling effects. However, ME also has disadvantages that make it less likely to be used in consumer acceptance testing. First, it requires more training time for consumers due to its complicated nature of number usage. Secondly, when panelists are instructed to freely assign numbers to samples, they tend to pick a few favorite numbers (e.g., 1, 10, 50, 100) and restrict their number ranges (e.g., 1 to 10, 1 to 100), both of which can affect the results obtained (Moskowitz 1977). Third, data obtained needs to be normalized to a single scale before statistical analysis due to individual differences in using different ranges of numbers (Moskowitz 1977). This increases the amount of work that must be done by the researcher even before they analyze the data. Lastly, the ratings obtained by ME do not provide semantic information, which is often desired from consumer acceptance testing. The same numerical rating given by two separate panelists to the same product could mean two different degrees of liking or disliking (e.g., the first panelist may moderately like the product and the second panelist may extremely like the product but both gave the same numerical rating). These disadvantages are often seen to out weigh the benefits of ME and therefore not used often in consumer acceptance testing.

Recently, the Labeled Hedonic Scale (LHS) was developed (Lim et al., 2009) to overcome some recognized limitations of the 9-point hedonic scale while providing the advantages of ME (i.e. ratio-level data and reduced ceiling effect). The LHS is a semantically-labeled, continuous hedonic line scale that yields the magnitude of liking and disliking of samples (see Fig. 1 & Table 1) and offers advantages over other scales used in consumer acceptance testing. First, it has been shown to provide ratio-level data in a well controlled psychophysical setting with a wide range of samples (Lim et al., 2009). Secondly, it was also shown to provide slightly better discrimination power over the 9-point hedonic scale (Lim et al., 2009; Lim et al., 2010). Third, the scale is much more resistant to ceiling effects than the 9-point hedonic scale due to being bound by "most liked sensation imaginable" and "most disliked sensation imaginable," which encompasses the entire range of sensations (Lim et al., 2009). Fourth, it retains the advantages of simplicity (ease of use) and ability to provide semantic information as the 9-point hedonic scale does. These advantages of the LHS can provide data that is very useful during product development research. For example, a company may discover that consumers like their product 25% more when 2% more vanilla extract is added to the formula. In addition to the ratio data, the company will also receive semantic information similar to the 9-point hedonic scale. This provides a verbal descriptor to go along with a numerical value and increases information obtained about a specific product.

The current study was aimed to confirm that the LHS can produce ratio-level data by comparing it to ME, which has been shown to provide ratio-level data (Stevens & Galanter, 1957; Moskowitz, 1971; Moskowitz & Sidel, 1971), in a product development setting where tested samples are closely related variations of the same formulation.

Previous experiments have shown that the LHS can provide ratio-level data; however, they were in a well controlled psychophysical test setting or in a consumer test setting, with a wide range of samples (Lim et al., 2009; Lim & Fujimaru, 2010). In the current study, stimuli will only slightly vary by the level of sucrose to simulate a product development research setting in which a company is trying to optimize the sweetness of their product. This will also test the scales discrimination power for closely related products. If the LHS achieves ratio-level data in this type of setting, then it could become a valuable tool for product development research.

MATERIALS AND METHODS

Subjects

A total of 41 subjects (12 males and 29 females), most of whom were undergraduate students, between the ages of 19 and 32 were recruited from the Oregon State University campus to participate in the experiment. A single subject's data was dropped due to inability to follow the directions and procedures, bringing the total subjects to 40 (12 males and 28 females). Subjects gave written informed consent and were compensated for their participation. The subjects were non-smokers, in good general health, and had no deficit in taste or smell. Subjects were asked to refrain from eating or drinking for a minimum of one hour prior to their individual testing sessions. Some subjects had prior experience with ME and the LHS.

Samples

Two product systems, consisting of six samples each, were used in the experiment. The product systems used were a cherry flavored beverage (Kool-Aid®, Kraft Foods Global Inc., Northfield, IL) and vanilla custard (Birds Custard Powder, Premier Ambient Products (UK) Ltd., Spalding, Lincs). Each product system used five different sucrose concentrations with molarities equaling 0.14M, 0.20M, 0.28M, 0.40M, and 0.56M. Both systems also used a duplicate concentration to ensure that panelists were correctly using the hedonic scales. Duplicate concentrations were 0.2 M sucrose and 0.4 M sucrose for the cherry flavored beverage and vanilla custard respectively. The product formulas are listed in Table 2. Samples were made fresh weekly prior to testing and anything left over was disposed of one week after preparation.

Kool-Aid® Preparation Directions

One packet of cherry flavored Kool-Aid® was combined with 1.9 liters of miliQ water in a large pitcher and stirred with a spoon until it was completely dissolved. Then the sucrose (Table 2) was added to the stock solution. A stir bar was used to mix the sugar into solution until it was completely dissolved, and then promptly refrigerated at about 39 degrees Fahrenheit.

Vanilla Custard Preparation Directions

Twenty grams of Birds Vanilla Custard powder (or about 2 tablespoons) was combined with 500 mL of whole milk (HY-TOP) and whisked until blended. Then the sucrose (Table 2) was whisked into the milk and powder solution, but not completely dissolved. The bowl containing the mixture was then microwaved on high for a total of 6 minutes. The solution was stirred after 3 minutes of microwaving, and then about every 45-60 seconds until the whole 6 minutes was complete. The hot custard was removed from the microwave, covered with saran wrap touching the surface of the hot custard (to prevent a skin from forming), and then promptly refrigerated for at least 24 hours.

Procedure

Each subject attended two sessions on two separate days, one session for ME and the other session for the LHS. At the beginning of each session, verbal instructions were given on how to use either the LHS or ME scale. Following verbal instructions, a short practice session was performed to ensure panelists understood how to use the scale. For the LHS session, subjects were asked to use the LHS to rate 15 real and remembered sensations read aloud (see Table 3 for a list of sensations used). For the ME session, subjects were asked to use ME to rate the length of four different lines and the area of four different squares (Moskowitz 1977). The lines and squares were held up one at a time for rating. Subjects then used one of the two scale types to rate their liking and disliking for six cherry flavored beverages and six vanilla custards, given as two separate testing blocks of six samples, one at a time. The panelist consumed (i.e. did not spit) either 10 mL of cherry flavored Kool-Aid® or one small plastic spoonful of vanilla custard. After tasting the sample, the panelists rated how much they liked or disliked the sample. In between each sample, panelists were asked to rinse their mouth at least twice with water during a one minute break. A 5 minute break was given between testing blocks to prevent fatiguing. The order panelists tested the stimuli was randomized and counterbalanced across all subjects. All ratings were made on paper ballots, with each stimulus rated on a separate ballot. Booklets of ballots were made prior to each testing session period. All testing was conducted one-on-one in a sensory testing booth.

Data Analysis

Before the collected data could be statistically analyzed, a data transformation was made to compare the two scales directly. The LHS data was translated to a range of -

100 (most disliked sensation imaginable) to 100 (most liked sensation imaginable). The ME data was normalized across subjects by dividing the grand mean of the absolute value of all subjects by the mean of the absolute values of the ratings of each subject, then multiplying individual subject ratings by each subjects individual factor. The normalized data was then put through a standardizing procedure (Moskowitz 1977) to reflect the LHS range of -100 (most disliked sensation imaginable) to 100 (most liked sensation imaginable). The collected data was analyzed using repeated-measures analyses of variance (ANOVAs) followed by the post-hoc Tukey's test. All statistical analyses were performed using Statistica 8 (StatSoft, Inc.).

RESULTS AND DISCUSSION

Figure 2 and 3 compares the data obtained using the LHS with the data obtained using ME for the cherry flavored Kool-Aid® (Figure 2) and vanilla custard (Figure 3) sample sets respectively. The most liked cherry flavored Kool-Aid® sample was the 0.40 M sucrose concentration (26.90) followed by the 0.56 M (23.63), 0.28 M (20.97), 0.20(1) M (16.87), 0.20(2) M (14.20), and 0.14 M (-2.73). The 0.40 M sucrose concentration was liked 10.66 times more than the least liked sample (0.14 M) but was liked only 1.14 times more than the second highest rated sample (0.56 M). The most liked vanilla custard sample was the 0.56 M sucrose concentration (28.33) followed by the 0.40(1) M (27.03), 0.40(2) M (24.53), 0.28 M (24.27), 0.20 M (13.70), and 0.14 M (7.10). The 0.56 M sucrose concentration was liked 3.99 times more than the least liked sample (0.14 M) but was liked only 1.05 times more than the second highest rated sample (0.40(1) M).

A repeated-measures ANOVA performed on the data indicated that there was no significant effect between the two scales for either sample set (Kool-Aid®: $F(1, 39) =$

2.07, $p = 0.16$; Custard: $F(1, 39) = 1.98$, $p = 0.17$). The scale by stimulus interaction also had no significant effect for either sample set (Kool-Aid®: $F(5, 195) = 0.19$, $p = 0.97$; Custard: $F(5, 195) = 0.43$, $p = 0.83$). This indicates that there is no significant difference between the ratings of samples between scales. This conclusion was strengthened with the post-hoc Tukey test indicating the ratings of the samples were not significantly different between the two scales as well. These results indicate that the data obtained from the LHS is not statistically different than ME, confirming that the LHS produces data that is equivalent to ratio-level data produced by ME in this product development research setting where products only differ slightly. However, there are trends on the graph that come into question. In Figure 3, it can also be seen that the 0.28 M custard's 95% confidence interval does not overlap the mean hedonic rating obtained using the LHS. This infers that for this sample, ratings between scales were significantly different. This could be due to the order samples were given to subjects. Eleven of the 40 subjects received the vanilla custard at 0.28 M sucrose concentration as their last sample for both the LHS and ME. In ME, subjects are required to remember their standard of reference in order to properly rate the product. This is difficult for consumers to do across six samples, and therefore could have affected the final outcome of the ME data for this sample. In contrast, the LHS does not require the consumer to remember a reference sample and can be more consistent throughout the entire sample block.

Table 4 and 5 show the mean hedonic ratings for cherry flavored Kool-Aid® samples and for vanilla custard samples respectively. The mean ratings for each sample set are separated by two scales used. The repeated measure ANOVA results for each scale method indicated that there was a significant effect of stimulus for both sample sets

(Kool-Aid® $F(5, 195) = 24.80$, $p < 0.00001$; Custard: $F(5, 195) = 18.25$, $p < 0.00001$). This indicates that both the LHS and ME were able to discriminate different samples as expected. To compare discrimination power, significant difference for all possible pairs was compared via the Tukey test. Each scale, per sample set, has 15 possible comparisons. In Table 4, the LHS discriminated seven of the possible 15 pairs while the ME discriminated only six of the possible 15 pairs. In Table 5, the LHS discriminated eight of the possible 15 pairs while the ME discriminated only six of the possible 15 pairs. For both sample sets, the LHS had a higher number of significantly different pairs. This data indicates that the LHS has slightly better discrimination power over the ME method because of its ability to determine more statistically different sample pairs.

Tables 4 and 5 also show that in all cases for both scales, duplicates were not rated significantly different from each other (Tukey test, $p > 0.90$). The duplicates were used to ensure that subjects were properly using the scales. They also served to determine if the subjects were able to discriminate between samples. Because there is no significant difference between any of the duplicates on both scales, it can be assumed that subjects were properly using the scales as instructed. It can also be assumed that the subjects were able to discriminate between samples properly. These assumptions are important in order to validate a subject's ability to discriminate between products and ensure subjects are not guessing or randomly assigning hedonic ratings.

SUMMARY

The results of the present study indicate that the LHS can be implemented into a large scale consumer testing for product development research and obtain a higher degree of information than the more commonly used 9-point hedonic scale. The results indicate the LHS provides ratio-level data equivalent to data obtained from ME in a setting similar to product development research. The LHS was also shown to have slightly better discrimination power over ME. The ability to see how much more a product is liked or disliked compared to another can be a great advantage when testing new products. The slightly better discrimination power that the LHS provides also allows the comparison of closely related products. This can often happen in product development to allow optimization of a specific ingredient. For example, table 5 shows the custard data obtained from the LHS method; it can be seen that the highest rated custard contained 0.56 M sucrose (mean = 28.33). This rating is just about 4 times larger than that of the custard with the lowest rating (0.14M, mean = 7.10) but only 1.05 times (or 5%) larger than the second highest rated custard (0.4M (1), mean = 27.03), which is insignificantly different by the Tukey test. A company may decide, based on this data, to launch the product at 0.4M sucrose concentration instead of the highest rated 0.56M in order to keep ingredient cost low because they would be using 28.5% less sucrose per batch of custard made. If the company had decided to use the 9-point hedonic scale during their consumer testing, they most likely would have gone with the highest rated product without knowing how much more it was liked than the others, and end up spending more on ingredient cost than may be needed. In conclusion, with the LHS's ability to provide ratio-level data with a higher discrimination power on top of semantic information gives it the potential to become a powerful tool in product development research as well as the general food and sensory science community.

BIBLIOGRAPHY

Engen T, McBurney DH. 1964. Magnitude and category scales of the pleasantness of odors. J Exp Psychol. 68:435-440.

Gay C, Mead R. 1992. A statistical appraisal of the problem of sensory measurement. J Sen Stud. 7:205-228.

Hollingworth HL. 1910. The central tendency of judgement. J Philos Psych Sci Meth. 7:461-469.

Jellinek G. 1964. Introduction to and critical review of modern methods of sensory analysis (odour, taste and flavour evaluation) with special emphasis on descriptive sensory analysis (flavour profile method). J Nutri & Diet. 1:219-260.

Lawless HT, Heymann H. 1998. Sensory evaluation of food: principles and practices. New York: Chapman & Hall.

Lim J, Wood A, Green BG. 2009. Deviation and Evaluation of a Labeled Hedonic Scale. Chem Sense. 34:739-751.

Lim J, Fujimaru T. 2010. Evaluation of the labeled hedonic scale under different experimental conditions. Food Qual Pref. 5:521-530.

Moskowitz HR, Sidel, JL. 1971. Magnitude and hedonic scales of food acceptability. J Food Sci. 36:677-680.

Moskowitz HR. 1971. The sweetness and pleasantness of sugars. Am J Psychol. 84:387-405.

Moskowitz HR. 1977. Magnitude estimations: notes on what, how, when, and why to use it. J Food Qual. 3:195-227.

Moskowitz HR. 1980. Psychometric evaluation of food preferences. J Foodserv Sys. 1: 149-167.

Moskowitz HR. 1982. Utilitarian benefits of magnitude estimation scaling for testing product acceptability. Philadelphia (PA): American society for testing and materials.

Peryam DR, Girardot NF. 1952. Advanced taste-test method. Food Eng. 24:58-61.

Peryam DR, Pilgrim FJ. 1957. Hedonic scale method of measuring food preference. Food Tech. 11:9-14.

Peryam DR, Polemis BW, Kamen JM, Eindhoven J, Pilgrim FJ. 1960. Food preferences of men in the U.S. armed forces. Chicago (IL): Quartermaster Food and Container Institute for the Armed Forces. p. 1-160.

Resurreccion AVA. 1998. Consumer sensory testing for product development. Maryland: Aspen Publishers, Inc.

Stevens SS. 1951. Mathematics, measurement, and psychophysics. Handbook of experimental psychology. New York: John Wiley and Sons.

Stevens SS, Galanter EH. 1957. Ratio scales and category scales for a dozen perceptual continua. J Exp Psychol. 54:337-411.

Townsend JT, Ashby FG. 1984. Measurement scales and statistics: the misconception misconceived. Pysch Bulletin. 96: 394-401.

Villanueva NDM, Petenate AJ, Da Silve MAAP. 2000. Performance of three affective methods and diagnosis of the ANOVA model. Food Qual Pref. 11:363-370.

Figure 1 The labeled hedonic scale (LHS) used in the ballots for the experiment. See table 1 for numerical ratings for the descriptors

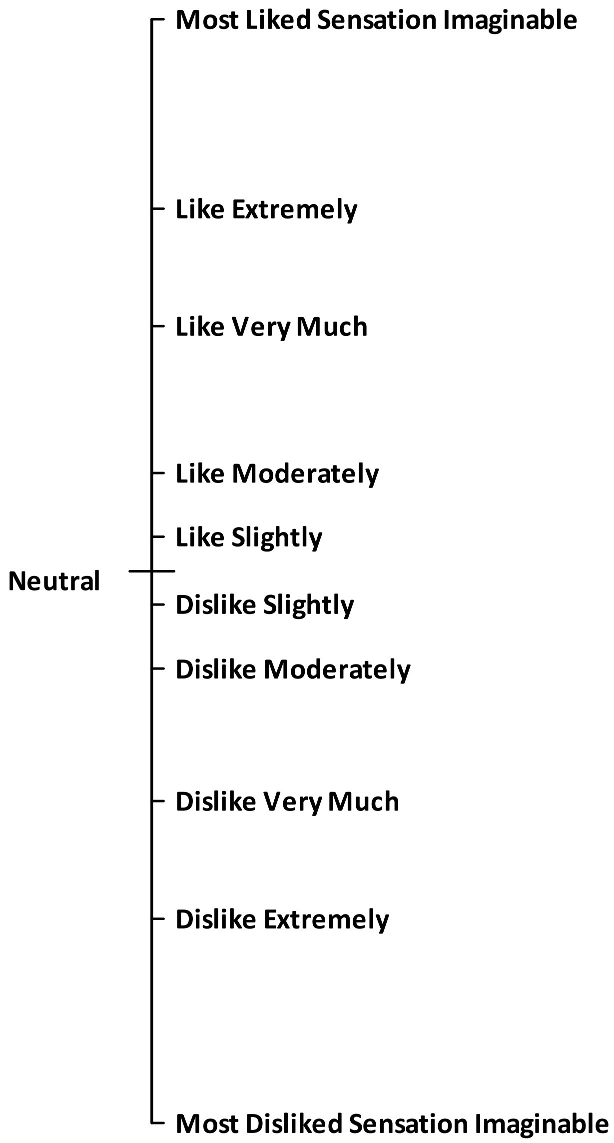
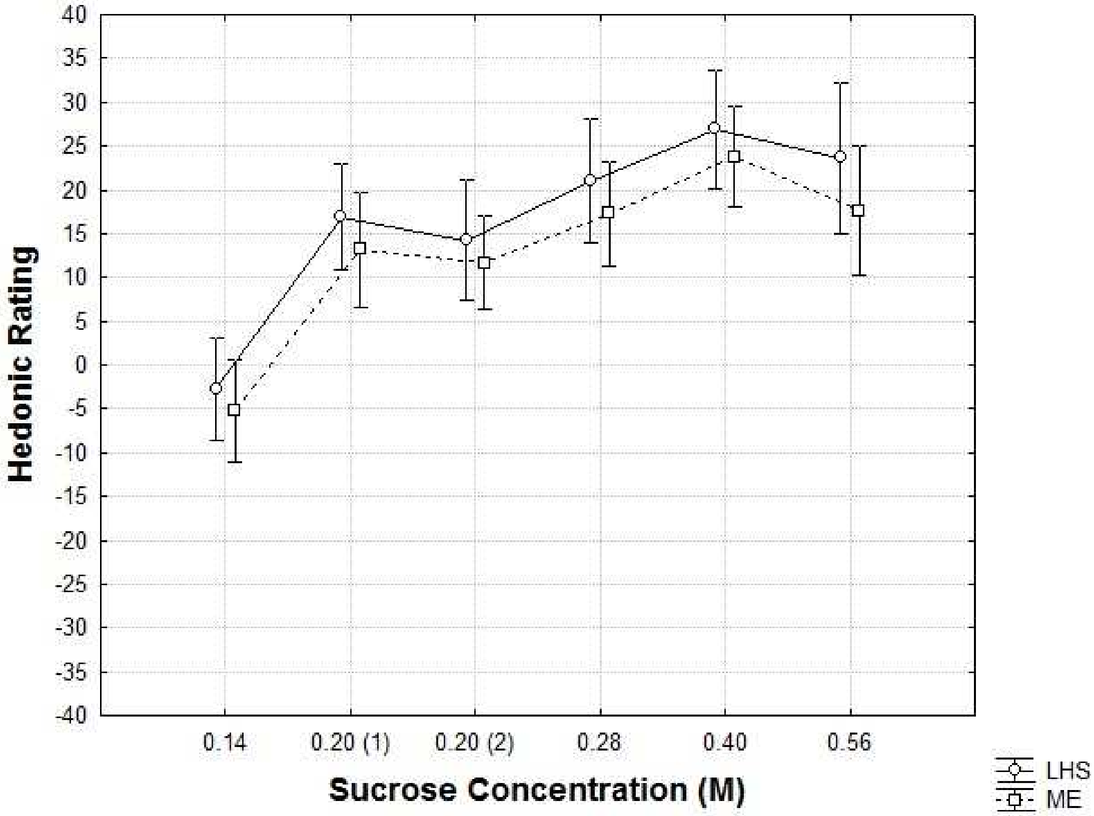
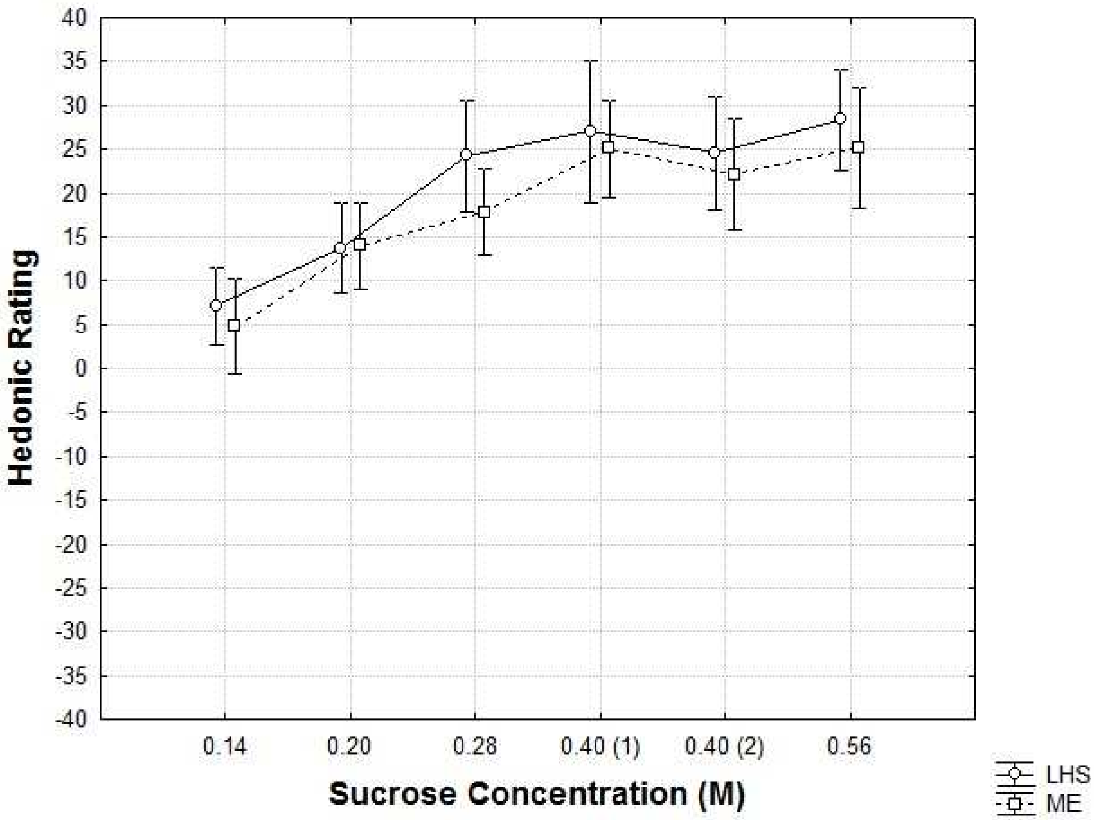


Figure 2 Mean Hedonic Ratings of cherry flavored Kool-Aid® Samples with varying sucrose concentrations



95% confidence intervals are denoted by the bars. The duplicate for this sample set was 0.20 M and are distinguished by numbers in parenthesis on the chart. The means for ME were standardized to the LHS data.

Figure 3 Mean Hedonic Ratings of Vanilla Custard Samples with varying sucrose concentrations



95% confidence intervals are denoted by the bars. The duplicate for this sample set was 0.40 M and are distinguished by numbers in parenthesis on the chart. The means for ME were standardized to the LHS data.

Table 1 The labeled hedonic scale's descriptors shown as numerical values

Descriptors	Scale Value
Most like sensation imaginable	100
Like extremely	65.72
Like very much	44.43
Like moderately	17.82
Like slightly	6.25
Neutral	0
Dislike slightly	-5.92
Dislike moderately	-17.59
Dislike very much	-41.58
Dislike extremely	-62.89
Most disliked sensation imaginable	-100

Table 2 Formulations of samples used in the experiment

Cherry Flavored Beverage	
Cherry Kool-Aid® Powder	1 packet
DI water	1.9 L
Sucrose	X*

Vanilla Custard	
Whole Milk	500 mL
Birds Custard Powder	20 g
Sucrose	X**

*X = 91.1g (0.14 M or 4.6%), 130.1g (0.20 M or 6.4%), 182.1g (0.28 M or 8.7%), 260.1g (0.40 M or 12.0%), or 364.2g (0.56 M or 16.1%)

**X = 24.9g (0.14 M or 4.6%), 35.6g (0.20 M or 6.4%), 49.8g (0.28 M or 8.7%), 71.2g (0.40 M or 12.0%), or 99.7g (0.56 M or 16.1%)

Table 3 Imagined sensations used for LHS practice session

Real and Remembered Sensations
The taste of plain bread
The taste of a soggy potato chip
The feel of pure silk
The sound of fingernails dragging across a blackboard
The taste of your favorite chocolate
The feel of a massage
The smell of clean laundry
The smell of vomit
The taste of water
The smell of a rose
Stinging eyes from cutting an onion
The feel of course sandpaper
The smell of bad body odor
The feel of a minor scratch
The taste of room temperature soda

Table 4 The mean hedonic ratings of cherry flavored Kool-Aid® samples with varying sucrose concentrations

	Sucrose Concentration (M)					
	0.14	0.20(1)	0.20(2)	0.28	0.4	0.56
LHS	-2.73[a]	16.87[b]	14.20[b]	20.97[bc]	26.90[c]	23.63[bc]
ME	-5.19[a]	13.15[bc]	11.64[b]	17.30[bc]	23.80[c]	17.62[bc]

Different super scripts indicate significant differences across ratings of products by Tukey HSD test ($p<0.05$).

Table 5 The mean hedonic ratings of the vanilla custard samples with varying sucrose concentrations

	Sucrose Concentration (M)					
	0.14	0.20	0.28	0.4(1)	0.4(2)	0.56
LHS	7.10^{a}	13.70^{a}	24.27^{b}	27.03^{b}	24.53^{b}	28.33^{b}
ME	4.78^{a}	13.99^{ac}	17.83^{bc}	25.05^{b}	22.12^{bc}	25.18^{b}

Different super scripts indicate significant differences across ratings of products by Tukey HSD test ($p<0.05$).