REHABILITATING THE 9-POINT HEDONIC SCALE TO MAKE IT VALID FOR ACROSS GROUP COMPARISONS

By

RACHEL GLINTZ

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2014

To the Glintz family for supporting me whilst earning a master’s degree

ACKNOWLEDGMENTS

I thank my parents for their continued support and encouragement in completing my thesis. I thank my committee members, especially Dr. Sims for all of his help from beginning to end during my Masters education and throughout the writing of my thesis. I thank Dr. Bartoshuk for teaching my how to analyze my results and for spending so much time helping me to make the finished copy. I thank my lab mates for all the laughs. I thank Brittany Hubbard for all her patience and guidance while I was working on my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

REHABILITATING THE 9-POINT HEDONIC SCALE TO MAKE IT VALID FOR ACROSS GROUP COMPARISONS

By

Rachel Glintz

May 2014

Chair: Charles A. Sims
Major: Food Science and Human Nutrition

The Hedonic general Labeled Magnitude Scale (HgLMS) was compared to the 9-point hedonic scale to generate data to discriminate liking between products and compare results across different groups of people. The HgLMS rates foods in context of all liking, ranging from -100 (strongest disliking of any kind ever experienced) to 0 (neutral) to 100 (strongest liking of any kind ever experienced). The 9-point hedonic scale does not provide this context.

Panelists were randomly assigned to the HgLMS, 9-point scale, or modified version of the 9-point hedonic scale with verbal directions like the HgLMS (1=strongest disliking of any kind experienced, 9=strongest liking of any kind experienced). Panelists rated liking of 17 food items from memory then rated liking of orange juice, black coffee, and grapefruit juice samples differing in quality. Panelists were instructed verbally and nonverbally (HgLMS and 9-point modified) or nonverbally (original 9-point). Panelists rated the intensity of sweet, salty, sour, and bitter solutions using the general Labeled Magnitude Scale (gLMS) at a separate session.

All scales showed equal discrimination ability in ratings of sample preference based on ANOVA and Duncan's mean separation. The HgLMS was most effective at

comparing different groups of panelists based on perceived taste intensity (i.e., supertasters and others), followed by the modified 9-point scale; the original 9-point scale was least effective at comparing groups.

A second study was conducted to determine if the modified 9-point scale would improve in making across group comparisons utilizing taster status by adding non-food items. However, the distribution of bitter intensity (for comparing supertasters and others) differed between studies 1 and 2 so the aim of the second study could not be achieved.

CHAPTER 1
INTRODUCTION

The research heretofore was conducted to investigate the importance of using the correct scale and giving proper instructions in sensory laboratory applications. Scales are the most important tools sensory scientists use to collect data from participants on their ratings of how much they like products and how intense they rate certain product characteristics. Thus, it is critical to have the appropriate scale to measure such information. It is often difficult to introduce a new scale into the sensory laboratories of the industries conducting this research because the standard scales are adequate at collecting information for the technicians to interpret. However, there may be room for improvement with a novel scale that can still meet the needs of the industries using them and provide more information than currently offered.

In the food science industry there is ongoing research into how products are perceived by consumers. Whether a new product is being developed that is unique to the products on the market or a competitor is working to create a product that is similar to one currently on the market, sensory panels are held to give the development team feedback. Traditionally, the 9-point hedonic scale is used to measure likeability of food products. However, the 9-point scale does not provide panelists with a standard with which to rate their liking of the product they are testing. Instead, panelists must decide the context of the scale, whether they are rating the product on a spectrum with the best sample of its kind they have ever tasted or if they are to compare the set of samples to each other with that set as the only means of comparison. The data that is produced, therefore, can not be used to compare the results across different groups of people, such as women versus men, or young versus elderly. If the panelists are instructed to

treat the scale differently by giving them a standard independent to the sensation they are rating, the scale may then be able to produce data that can be compared across different groups of people. By combining the ease of use of the 9-point scale with this additional quality, industry will be able to explore different research paths without the discomfort of using an entirely new scale.

Literature Review

Sensory evaluation of foods is a critical part of the food industry; product appeal to consumers' visual, olfactory, and gustatory senses will affect sales and ultimately profitability. However, consumers gauge sensory aspects of foods differently depending on multiple factors. Sensory perception varies widely among panelists due to gender, ethnicity, age, and attitudes in addition to countless other factors. Panelists' interpretation and understanding of the sensory scale contributes to differences in responses (Lawless and Heymann 2010). A panelist's translation of the semantics of categorical scales, including both numerical and verbal labels, plays a large role in the collection of accurate sensory data (Lawless and Heymann 2010).

Scaling methods have evolved over time as information about sensory perception has progressed. The importance of providing accurate sensory information to the food industry is apparent especially in product development (Lawless and Heymann 2010). New products are marketed after going through one or more sensory panels to determine product likeability. Sensory ratings of food products must be simple to assess and easy to interpret by the company indicating product approval or disapproval. Different scales require more or less training of panelists prior to usage. Simpler scales may be more attractive to companies doing sensory research on their

food product, but they may not assess across group comparisons as well as those scaling methods requiring more in depth training.

The ability to make across group comparisons of hedonic data depends on a scale's capacity to provide circumstances for each panelist to produce their own anchors with which they interpret food samples. Some scales allow for a broader interpretation of samples whereas others have a very narrow portrayal of consumer sensitivity. Some scales (e.g. the 9-point scale) may be faulted for their lack of sensitivity due to the limiting of statistical analysis to nonparametric means (Lim and others 2009). The ceiling effect occurs when panelists show an aversion to using the uppermost and/or bottommost category when rating a product's likeability. Scales that experience a ceiling effect, such as the 9-point hedonic scale, will likely have inexact data in comparison to scales that do not exhibit this effect (Lim and others 2009).

Gustation and Olfaction

Taste is one of the senses that humans employ on a daily basis. The involvement of taste in our lives relates to the need for nutrition to sustain life. Taste perception has evolved to allow for the recognition of four basic tastes, salty, bitter, sour, and sweet, among which are those that signal danger and others that indicate pleasure. The four basic tastes are hardwired to produce certain responses from birth (Steiner 1973). Bitter taste receptors recognize bitter molecules and send signals to the brain to avoid consuming the food containing them (Mueller and others 2005). This is an evolutionary design that works to prevent ingestion of poison, such as that produced by plants, often given a bitter taste to ward off potential predators (e.g. herbivores). The other three tastes warrant receptors that induce a need based response, be it for sustenance or survival. Often a craving for one of the tastes signifies a deficiency (e.g. a

sodium deficiency) and functions to increase consumption of the necessary nutrient to achieve homeostasis (Stein and others 1996).

Research has brought to light differences in ability among humans to identify the presence of certain tastes relating to the degree of sensitivity to taste stimuli, i.e. tastants, that focuses on the genetic predisposition to perceive bitter taste, as evidenced by the ability to taste the compound phenylthiocarbamide (PTC) (Fox 1931). Those who can taste the bitterness of PTC are considered to be tasters, while those who can not are nontasters. The taster designation is related to the genetics of the individual, with tasters being dominant for the trait and nontasters being recessive (Blakeslee 1932). A few decades later, with the safety of PTC under scrutiny, the test for taste status was shifted to the compound 6-n-propylthiouracil (PROP), which also contains the bitter-tasting chemical group N-C=S (Wheatcroft and Thornburn 1972). Since PROP is a medication, its toxicity is known and can be minimized. PROP functions by attaching to a bitter taste receptor, T2R38 in humans (Kim and others 2003). Tasters' aversion to the bitterness induced by PROP was extended to the consumption of vegetables, dissuading them from consuming the necessary nutrients.

Quinine is another bitter chemical that can be used to differentiate between tasters and nontasters. Nontasters are less sensitive to quinine than tasters (Blakeslee 1932). Women have lower thresholds to quinine, as with PTC/PROP, than men (Falconer 1946). A genetic predisposition for a low threshold of detection of PROP and quinine leads a certain group of people to taste these bitter substances at lower concentrations. Tasters have a higher threshold, needing a greater concentration to stimulate their taste buds for bitter foods (Fischer and Griffin 1961).

In order to understand how taste intensity is perceived across individuals, the concept of cross-modality matching, in which taste stimuli are related to nontaste stimuli (e.g. light intensity) in order to measure the intensity of both on the same scale, was utilized on tasters and nontasters to illuminate a range of abilities (Fast 2005). At the top end of the intensity scale were those who related the taste of PROP to stimuli that was very intense, dubbed supertasters, those who related PROP to a lesser intensity than the former were called medium tasters and so on for nontasters. These designations can be further broken down depending on the concentration of fungiform papillae on the tongue, those papillae that reside on the anterior region of the tongue. A PROP supertaster can taste PROP and also has a large number of fungiform papillae (Hayes and others 2008). A PROP medium taster can taste PROP but has fewer fungiform papillae than the supertaster. A PROP nontaster can not taste PROP; however the number of fungiform papillae is variable (Bartoshuk and others 2005).

Gustation and olfaction go hand in hand to generate flavor; flavor is the integration of retronasal olfaction and taste (Rozin 1982). Retronasal olfaction occurs as odor molecules from food undergoing mastication and swallowing are sent up through the palate into the nose. Orthonasal olfaction occurs when we smell odors through our nostrils. Olfaction, the sense of smell, is the result of a process involving the uptake of odorants by odor receptors of the olfactory epithelium that reside on the olfactory sensory neurons, which send signals to the brain to distinguish among odors we gather in our environment (Schild and Restrepo 1998).

There is an important distinction between sensation and perception of smell. Sensation is the recognition of odors by the primary olfactory cortex, or piriform cortex,

of the brain (Herz 2003). Our awareness of an odor passing into our nostrils is the perceived olfaction. An additional aspect of smell relates to how certain odors stimulate the trigeminal nerve to create a feeling when they reach the brain (e.g. cool, heat, burning). Activation of trigeminal nerves along with odor receptors generates a complete picture for certain odors, such as peppermint (Herz 2012). The combined efforts of olfaction and gustation along with trigeminal sensations result in human perceptions of foods that can be rated on scales to detect affect to be compared among individuals.

History and Development of Scaling Methods

Several scaling methods were designed to measure differences among food preferences. The 9-point hedonic scale was created in 1949 for the rating of foods served in the military canteen (Jones and others 1955). In 1957 Stevens developed magnitude estimation, in which subjects were to assign numbers to stimuli in a manner proportional to the magnitude of sensation prescribed by the stimulus. The 1960s produced the visual analogue scale (VAS) in its attempt to rate sensory intensities along a line anchored by extreme variables equating to minimum and maximum sensory intensities (Aitken 1969). The category-ratio scale was first created by Borg using labels that described increasing intensities along a line, each in measured proportions to the others in perceived intensity (as cited in Borg 1982). The Labeled Magnitude Scale (LMS) was developed using principles based on category-ratio scales to spatially separate the verbal categories used, in this case, to rate oral sensations based on magnitude estimation (Green and others 1993). The LMS used the anchor words “strongest imaginable” to produce data with a fixed end-point. The labeled affective magnitude (LAM) scale was generated for the ratings of affect, or likability, of foods but structurally similar to the LMS with categories from the 9-point scale arranged along a

line with magnitude estimation properties and anchored with the panelist's greatest imaginable liking/disliking. (Schutz and Cardello 2001).

The best-worst scale is a choice-based approach that was developed to have panelists signify their preference for the best and worst stimuli out of a set of three or more and improves upon the paired preference test (Marley and Louviere 2005). The general Labeled Magnitude Scale (gLMS) evolved from the LMS to permit comparisons across groups (Bartoshuk and others 2005). The modeling of the hedonic general Labeled Magnitude Scale (HgLMS) after the gLMS produced a scale with the same ability to create standards for group comparisons with the focus on likability instead of sensory intensity (Bartoshuk and others 2005). In an effort to develop a ratio-level scale that measures panelists' degree of liking of food samples, Lim created the labeled hedonic scale (LHS) to demonstrate the importance of semantics when rating likability of foods on a category-ratio level comparable to methods using magnitude estimation (Lim and others 2009).

Since its inception, the 9-point hedonic scale has been the most commonly used scale in the food industry for rating likability and acceptability of food and drink (Jones and others 1955). The hedonic rating of the panelist is directly related to one of the nine categories on the scale ranging from 1, "dislike extremely," to 9, "like extremely" (Jones and others 1955). A panelist's assessment of the product, therefore, categorically represents the user's sensory perception of the product over a narrow range of choice verbal categories. The 9-point hedonic scale includes "neither like nor dislike" at the center of the scale. Many proponents value its ease of use and simplicity when interpreting data. There are, however, many aspects of the 9-point hedonic scale that

have led to the development of novel scales based on improved technology of sensory perception. The 9-point hedonic scale has a narrow range of descriptors and is susceptible to ceiling effects that may distort the sensory data.

The VAS was a line scale originally created to measure emotions and moods (Aitkens 1969), but was adapted to measure pain. On the VAS, a sensation that was twice as intense as another is rated at a value twice as large as the first. The LMS, a category-ratio scale with the anchors “strongest sensation imaginable” and “no sensation,” allows panelists to score their sensory intensities on a vertical scale, with panelists denoting increased sensations as the ratings increase toward the top of the scale (Green, Shaffer, & Gilmore, 1993). The spacing of the categories on the LMS is based on magnitude estimation so that the verbal descriptors relate to panelists’ perceptible sensation of food likability in a category-ratio manner.

The labeled affective magnitude (LAM) scale was created to help resolve some of the limitations of the 9-point hedonic scale by anchoring the panelists’ hedonic rating with their greatest imaginable liking and disliking to increase the scale’s sensitivity to food preferences (Jaeger and Cardello 2009). This scale allows for a greater degree of difference when rating foods with increasing intensities of likability as compared to the 9-point scale (Schutz and Cardello 2001). On the other hand, studies by Cardello (2008) have shown that the LAM scale has a tendency to narrow the range of categories when using verbal labels such as the “greatest imaginable like for any experience” as an anchor. This compression effect may be due to both consumer interpretation and perception of the terminology. Semantics are very important, especially for reproducibility and comparability to other studies.

The best-worst approach is an indirect scaling method that measures panelists' responses to extreme opposites of stimuli presented in a sensory test (Marley and Louviere 2005). This scaling method involves the panelist's preferential selection for best and worst out of a list of stimuli numbering three or more. The information generated is similar to that of the paired-comparison test; however, it is easier to obtain and can offer further evidence as to the preferences of the panelist. This scale will require minimal training and is more efficient in accurately discriminating between products than monadic studies, which test one product at a time. The ease of use combined with potentially stronger data for sensitivity of food-related testing makes the best-worst scale a worthy approach.

Another category-ratio scale that was developed to better estimate semantic meanings via magnitude estimation is the labeled hedonic scale (LHS) (Lim and others 2009). Panelists are trained in magnitude estimation prior to rating food samples. The semantics of specifically selected hedonic descriptors are chosen to determine the validity of the scale's semantic labels. The LHS can generate more accurate ratio-level statistics than the 9-point hedonic scale and is more sensitive when discriminating between categorical labels, supporting the spatial distance between semantic descriptors on the LHS. Due to the bipolar character of the LHS, it offers statistically symmetrical positioning of the anchor markers around the central neutral descriptor.

The general labeled magnitude scale (gLMS) is used to rate panelists' sensitivity to sensations on a 100-point scale with anchors of "no sensation" at 0 and the "strongest imaginable sensation of any kind" at 100 (Bartoshuk and others 2005). A panelist rates the food samples on the gLMS, with the potentially nonfood anchors, to

determine the relative intensities that foods generate. The gLMS relates sensory intensities across groups to offer comparisons among them. This scale is often utilized to determine the panelist's ability to sense certain flavors and to differentiate between supertasters and non-tasters. Supertasters will have on average a higher rating of sensations on the gLMS for these selected flavors because of their larger quantity of taste buds on their tongue as compared to non-tasters.

The Hedonic general Labeled Magnitude Scale (HgLMS) is currently being researched for its proficiency in rating food acceptability on a continuous line ranging from -100 to +100 with 0 labeled neutral. The anchors were originally labeled the strongest disliking and strongest liking, respectively, of any kind that the panelist had ever imagined (Bartoshuk and others 2005). The panelist will rate food samples based on the scale with the contextual boundaries they have created. A scenario such as reading a favorite book may be rated along the HgLMS as a 50. If a panelist were then to taste a food that they liked half as much as how they liked reading their favorite book, they would rate it at a 25. This scale will allow researchers to compare the data for likability of foods across groups of people. Because foods are rarely liked or disliked using the most extreme anchors, the top and bottom anchors are independent of food preferences.

Scaling Effects

According to Romano (2008), there are three general scaling effects that distort data analysis due to panelist misunderstanding of hedonic scale rating systems. The level effect occurs when panelists differ in their valuation of the average numeral on the scale; an average rating at the lower end of the scale versus the higher end will affect the way a panelist rates their samples (Romano and others 2008). A scaling effect may

occur if participants use only a portion, e.g., the top half, of the scale when rating food preferences. A panelist who gives repeatedly different scores to the same samples on one scale generates the variability affect. Researchers must keep in mind these scaling effects when choosing a hedonic scale to avoid skewed data.

Gaps in the Knowledge

With an understanding of super tasting came the realization that not all scales are alike in their ability to rate food acceptability (Hayes and others 2008). The 9-point scale is unable to differentiate among supertasters and nontasters, but the HgLMS can. When rating likability, supertasters will base their preferences on their sensory intensities. A food that has a hint of bitterness to a nontaster may be unpalatable to a supertaster and yield a much lower rating of affect on the HgLMS due to its range. Research is currently underway on the parallel relationship of hedonic ratings among the versions of the hedonic general labeled magnitude scale.

Preliminary research at the University of Florida was conducted using the three versions of the HgLMS; the first a line scale with specific adjective labels (weak, moderate, strong, very strong) spaced along it, the second a line scale with only the anchors and the midpoint, and the third a number scale without a line that asked participants to provide a number between -100 and 100 (Royuela 2011). The results were statistically significant among the foods designated as having high acceptance ratings. The results for the foods selected as having low acceptance ratings were statistically similar across the three versions of the HgLMS. Kalva et al. compared the HgLMS with the 9-point hedonic scale and found that due to the non-food related experiences used as the endpoints of the HgLMS, it can be used to compare food liking tendencies between different groups of people. The 9-point hedonic scale can not make

these comparisons due to subjects' implicit assumption that the scale is used for food only.

Research is currently being performed using the Hedonic general Labeled Magnitude Scale (HgLMS) in the sensory laboratory at the University of Florida. Due to the novel nature of this scale, panelists must reach a degree of familiarity with it prior to experimental use. The researcher must train the panelist to think of the strongest liking of any kind they can recall feeling in their lifetime; this will anchor the top end of the scale at 100. The researcher will instruct the panelists to do the same for the bottom anchor at -100 with the experience they relate to their strongest disliking. Once the anchors have been established, the researcher will explain that although these experiences are most likely non-food related, the panelists will be rating foods on this scale with the anchors as their basis of comparison. This is a difficult concept for panelists to grasp; however, the understanding of how the HgLMS works is crucial to producing accurate data. It is therefore up to the researcher to thoroughly explain this concept. This misunderstanding by panelists is a shortcoming of the hedonic gLMS.

Another area of study is the further understanding of what standards are required for comparing sensory ratings across individuals (Bartoshuk and others 2005). As noted above, panelists have differences in genetics and in life experiences, and accordingly, will rate food likability on the same hedonic scale differently depending on variation among these factors. Based on this knowledge, the semantic differences in intensity labels on a scale are not similar in absolute perceived intensity as once thought (Bartoshuk and others 2002). To better illustrate similarities in individual responses on taste preferences, it is important to set a useful standard that will have less variability

among respondents. Moreover, using the word “imaginable” in the scale’s anchors that label the most extreme sensations on an LMS scale may broaden the domain of the scale too excessively and its use should be reconsidered (Bartoshuk 2005). Instead, the scale would then contain end points denoting the most intense sensation the panelist has experienced.

The 9-point hedonic scale has many limitations in its ability to accurately measure panelists’ perceived sensations of likability toward foods. This scale was designed to start at 1, with 5 labeled as the neutral category, and equally spaced intervals between the 9 categories. The 9-point scale falls short when attempting to compare data across different groups (Kalva and others 2009). Due to scaling effects, the ceiling effect, and the variability effect, the 9-point scale can produce inaccurate data for across group comparisons. However, if the panelists can be given a context for how to rate foods on the 9-point hedonic scale, there may be potential to make across group comparisons with it. There has yet to be a study on the effects of giving panelists instructions to rate foods using the 9-point hedonic scale with 1 as the strongest disliking they have ever experienced and 9 as the strongest liking they have ever experienced. A study such as this will provide validation for the benefit of the context provided by giving panelists instructions on how to use the hedonic scale and may be a better alternative to using a novel scale in industry.

Alternatively, the Hedonic general Labeled Magnitude Scale gives panelists the context that the 9-point hedonic scale lacks. Hedonic acceptability ratings measured via this methodology may be able to infer relationships between food and non-food stimuli to present a grander scheme of comparisons outside of food science. Testing the

validity of comparing affective experiences across different groups with the HgLMS could lead to many future revelations in the food science industry.

Hypothesis and Specific Aims

Kalva et al (2014) compared the 9-point scale with the hedonic gLMS and showed that both scales are able to show differences across samples (within-subject comparisons); however, the 9-point scale as traditionally used fails to show across-group differences that can be demonstrated with the hedonic gLMS. One purpose of the present research is to provide the same context for the hedonic 9-point scale that is provided for the hedonic gLMS. That is, “9” will be given the label given to “100” (the strongest liking of any kind ever experienced) on the hedonic gLMS. The aim is to determine whether or not this will permit the hedonic 9-point scale to provide valid comparisons across groups.

There are many applications to which the results of this research can be of value. The food science industry uses scaling methods to obtain results on product likability prior to being sent to market. The research and development sector of a company can benefit from the understandings concluded from this research study, such as the novel capacity of the HgLMS to compare product likability across genders, racial categories, and age groups, among others.

Materials and Methods

Study 1

Sensory panel recruitment

In the first study, three hundred and fifteen panelists aged 18 and over were recruited from classes at the University of Florida to the Food Science and Human Nutrition Sensory Lab on the University of Florida campus to participate in the

evaluation. A survey link was provided to determine participant eligibility and availability for one of three days during a thirty-minute time block. Eligibility was based on whether or not the respondent had previously attended a taste panel at the University of Florida Sensory Laboratory. Any person who had participated in a previous sensory test was ineligible for this study. Those eligible to participate were then assigned to the day they were available to attend at a time they chose to come in. Each scale was randomly assigned to a day prior to conducting the survey. Out of the 315 who signed up for the panel, 290 participated in both the first and second sessions in experiment one. Compensation was provided for panelists who participated in both the hedonic session and the intensity session.

Of the 290 panelists who came to both sessions, 95 were assigned to the Hedonic general Labeled Magnitude Scale (HgLMS), the next 103 participants comprised the second group who used the original 9-point hedonic scale, and the third set of 92 panelists encompassed the group assigned to the 9-point hedonic scale with instructions given similarly to that of the HgLMS (referred to as the modified 9-point hedonic scale). The HgLMS is a continuous line scale with markers at -100 (the strongest disliking of any kind ever experienced) 0 (neutral) and 100 (the strongest liking of any kind ever experienced). Therefore, a panelist may choose any number on the scale between -100 and +100 that represents their liking of an item on the questionnaire. The 9-point hedonic scale is a category scale that allows panelists to select only one of the 9 number choices ranging from 1 (dislike extremely) to 9 (like extremely). In the modified 9-point scale, the anchor words at 1 and 9 that are present on the original scale are replaced with those from the HgLMS at -100 and +100,

respectively. All other labels provided on the original 9-point scale were removed from the modified version.

Panelist training

The panelists were trained at the start of each 30-minute session during the panel to understand how to utilize their particular scale. Participants of Group 3 were instructed to use the bottom anchor of the 9-point scale, 1, as the strongest disliking of any kind ever experienced and the top anchor, 9, as the strongest liking of any kind ever experienced. A hard copy of the select scale was provided for each panelist to refer to while participating in the panel. Panelists recorded their anchors on the given copy of the scale for their reference throughout the test. There was no training provided for participants who were assigned to the standard 9-point hedonic scale.

Sensory questionnaire

Each panelist was separated from the person next to him or her in a private booth, with its own computer, keyboard, and mouse. The Compusense program was on the computers to record the panelists' responses. Panelists were assigned a number at the beginning of the hedonic panel that they would also use to register for the intensity panel. All panelists were asked a series of demographic questions at the start of the session including their gender, age, height, and weight as well as their ethnicity and frequency of otitis media (middle ear infection). Height, in feet and inches, and weight, in pounds, were converted to kg and meters, respectively, to calculate BMI using kg/(m^2) (Centers for Disease Control and Prevention 2011). This demographic information was collected to provide data for analysis of the relationship of gender and BMI to the hedonic and intensity data collected.

The panelists assigned to groups 1 and 3 received a set of warm-up questions to familiarize them with the style of question they were to be asked during the tasting portion of the panel. These questions included rating the following life experiences: listening to your favorite music, spending times with your loved ones, the most intense anger you've experienced, the most intense annoyance you've experienced, eating your favorite food, eating your least favorite food, proudest you've ever been of accomplishing a specific goal, the most enthusiastic you've even been about a hobby, the shyest you've ever been, the most amused you have ever been by an anecdote, the most inspired you have ever been by a lecture, and the most disgusted you have ever been by a specific food. Six additional questions, two for each food item they were to taste during the panel, asked the panelists to rate their liking of the best and worst of any orange juice, grapefruit juice, and coffee they had previously experienced. Those panelists assigned to group 2 were asked to rate how much they liked each sample of orange juice, grapefruit juice, and coffee using the 9-point scale on a separate ballot with no further instructions.

Experimental design

The experimental design randomized the order of the samples given to each panelist with 24 different possible combinations for the order of the orange juice samples and two different possible combinations for both samples of grapefruit juice and coffee. Each group of panelists was given three consecutive sets of Styrofoam trays containing one of the following beverage samples: orange juice, grapefruit juice, and coffee. The orange juice samples consisted of four brands that ranged in quality. Tropicana Pure Premium™ and Simply Orange™ should be of higher quality than either Sweetbay™ or Minute Maid™. The grapefruit juice samples consisted of 2 brands,

Simply Grapefruit™ and Publix™ Grapefruit juice, which had a more intense bitter flavor than the former. The coffee samples were served black and consisted of 2 different varieties, Starbucks™ Espresso, an intensely bitter blend, and Starbucks™ Blonde coffee, which was not as strong in flavor and much less bitter than the espresso. The panelists rated the sensory acceptability of each sample with the scale they were assigned to. Panelists were instructed to take a bite of cracker and sip of water before and in-between each sample they tasted.

Panelists from all three groups came to the Sensory Lab the following week during a 20-minute session on one of two separate panels to rate sensory intensities using the general Labeled Magnitude Scale (gLMS). The gLMS is an intensity scale that ranges from 0 = no sensation to 100 = the strongest sensation ever experienced on a continuous line. The panelists had to undergo training for this panel as well. They were asked to provide their strongest sensation and write it on a copy of the scale at 100 for their reference during the session. The questionnaire included warm-up questions that asked panelists to rate the intensities of the following sensory experiences: the loudest sound ever hear, the loudness of a conversation, the brightness of a well-lit room, the brightest light ever seen, the loudness of a whisper, and the brightness of a dimly-lit restaurant. The panelists then rated the sensory intensities of four solutions on the gLMS. The solutions were 1.0 molar sucrose (sweet), 1.0 molar NaCl (salty), 0.03 molar citric acid (sour), and 0.001 molar quinine hydrochloride (bitter). Each of these were prepared the day before and stored in glass jars. The results were used to identify the range of panelists' perceived sensory intensities. This way, supertasters' tendencies toward hedonic ratings of foods were assessed.

Statistical analysis

Data was analyzed using Statistical Analysis Software (SAS) as well as SPSS Statistics. An Analysis of Variance (ANOVA) was run on the data for the orange juice, grapefruit juice, and coffee to determine whether the HgLMS, 9-point hedonic scale, and modified 9-point hedonic scale were able to differentiate equally among the samples. Duncan's multiple range test for mean separation was run on the data. These results were used to evaluate the consistency in the scales' ratings of sample acceptability. Correlation regression analysis was used to evaluate the quantity of significant relationships between the variables on the ballot. These results were used to determine the effectiveness of providing panelists with a context for rating foods. The intensity scale ratings of quinine bitterness were used to separate out the panelists that were supertasters. The range of perceived taste intensity, with supertasters rating quinine most bitter and others rating quinine least bitter, was used to identify how supertasters feel about foods. The extent by which each scale identified both supertasters and their affective leanings toward foods was compared based on the frequency of occurrence of statistically significant results in correlation regression analyses between bitter intensity ratings and hedonic ratings of foods.

Study 2

A second study was conducted to determine if there was a difference in the way the panelists used the modified 9-point hedonic scale if non-food related items were included on the ballot. Two hundred panelists were recruited through university classes as in the first study. Of those recruited, 83 participated in the sessions assigned to the HgLMS and 66 assigned to the modified 9-point scale participated. The panelists were trained to use either scale at the beginning of their prospective panel session as in

study one. Panelists were instructed to record their “anchors” – their strongest disliking and liking of any kind they had ever experienced – on a hard copy of the scale for their reference throughout the test.

The panelists in this study received the same set of warm-up questions as those panelists in the first study to introduce them to the scaling technique. Panelists used the scale to rate the likability of foods and non-food affective items from memory only. The additional affective items on the ballot were chosen based off of previous research indicating that they were rated on average along the length of the hedonic general Labeled Magnitude Scale and included Spending time with loved ones, Accomplishing an important goal, The most enthusiastic you’ve ever been about a hobby, Getting a good grade, A full, sound night’s sleep, Listening to your favorite music, Going to a fun party with friends, Successfully solving a very difficult problem, Meeting a major deadline on time, Getting a great deal on something, Watching your favorite TV show, The most inspired you’ve ever been by a lecture, Riding a roller coaster, Smelling your favorite flower, Working at your job, Speaking in front of an audience, Going to the doctor, Taking medication daily, Getting cut off in traffic, The shyest you’ve ever been, Not getting something you really wanted, Being in a minor car accident, Getting caught doing something you’re not supposed to, Having a deadline which seems impossible to meet, Being made fun of by others, Your most embarrassing moment, The most annoyed you’ve ever been, Getting bad news from your doctor, The most ashamed you’ve ever been of yourself, The end of an important, special relationship, and The angriest you’ve ever been. Panelists were asked to return another day to learn how to

use the gLMS and rate the intensity of the same 4 solutions use in the first study. The results were used to identify panelists as supertasters.

The same conditions for panelist anonymity and assignment hold for study two as in study one as do the collection of demographic information. There were ten different randomized orders of presentation for the questions on the ballot. Each panelist received the same affective and food related items on their ballot. As in study number one, Compusense was used to record panelist responses. Panelist compensation occurred on the second session for each session they attended.

SPSS Statistics was used to analyze the results for correlation regression relationships among the variables for each scale. Each scale was analyzed to determine the ability of the scale to separate out the supertasters among the panelists. The importance of context was analyzed by comparing the results from the HgLMS to the modified 9-point scale. In addition, the use of non-food related items on the ballot to increase understanding of how to utilize the scale as a life experience scale was compared to the first study, which did not include those items.

CHAPTER 2
RESULTS AND DISCUSSION

Study 1

As listed in Table 2-1, the demographics for the first study show that 68% of the participants were female and 32% were male. The majority of panelists fell within the age ranging from 17-22 with only 3% of panelists having an age 23 and above. The ethnic background of the panelists was 77% non-Hispanic and 23% Hispanic. There were five categories for race with 68% white or Caucasian, 10% black or African American, 1% Native American, Alaska Native, or Aleutian, 10% Asian or Pacific Islander, and 11% Other. There were 73% of panelists who had no incidence of middle ear infection, 17% with ear infections during their youth that were not considered serious, 7% who had ear infections that required antibiotics more than once, and 2% who required tubes in their ears.

Within Subject Comparisons

The primary intent of the first research study was to clearly show the discriminability of each of the three scales used. The ANOVA for each of the samples for all three scales verifies this. The separation of the panelists' mean hedonic ratings of the orange juice, black coffee, and grapefruit juice samples were consistent among the HgLMS, the modified 9-point hedonic scale, and the original 9-point scale, as seen in Table 2-2. Of the four brands of orange juice, both Simply Orange™ and Tropicana™ were of better quality than Sweetbay™ brand and Minute Maid™ orange juice. Simply Grapefruit™ is generally better in quality than Publix™ brand grapefruit juice. Espresso coffee is much stronger in bitter flavor than blonde coffee. These brands were chosen for their stark comparison in quality.

The mean likability for Simply Orange™ and Tropicana™ were 19.69 and 20.13, respectively, which is considerably higher on the HgLMS than the ratings of 8.59 for Sweetbay™ and 8.11 for Minute Maid™. The results of the ANOVA for the modified 9-point scale were 5.61 for Simply Orange™ and 5.58 for Tropicana™, which were liked more than both Sweetbay™ with a mean of 4.52 and Minute Maid™ with a mean of 4.22. The mean separation for the original 9-point scale was more spread out but still statistically the same as the other two scales with an average hedonic rating of 7.49 for Simply Orange™, 6.94 for Tropicana™, 5.49 for Sweetbay™, and 5.02 for Minute Maid™. The HgLMS participant responses for their liking of Simply Grapefruit™ had a mean of -11.71 as compared to -19.21 for Publix™ brand. For participants using the modified 9-point hedonic scale the rating for Simply Grapefruit™ was 3.46 while the rating for Publix™ was 2.73. There were similar ratings for the original 9-point scale with Simply Grapefruit™ at 3.75 and Publix™ at 2.95. Although the means for blonde and espresso coffee were within 0.2 of each other, they were still statistically significant. Blonde coffee was liked slightly less with a mean of -21.39 and Espresso a little more with a mean of -21.2. The modified 9-point scale also showed the separation but Blonde coffee was liked more on average with a mean of 3.42; Espresso had a mean just below it at 3.2. The mean separation for the coffees with the original 9-point scale had the same ranking as the modified version with a mean of 2.83 for blonde coffee and 2.51 for espresso.

In sum, all three scales provided the same comparisons across the orange juice, grapefruit juice and coffee samples. Each subject tasted all of the samples. This confirms that all three scales can provide equivalent comparisons.

Across-Group Comparisons

The next test of the three scales concerns across-group comparisons. The subjects in these studies can be divided into various groups. For example, one can compare groups with different body mass indexes (BMIs). Another comparison between supertasters (those who experience the most intense taste sensations) and others can be made using bitter intensity ratings. Since supertasting has been shown to be positively correlated with liking and disliking food, one can also use food liking as the variable to group subjects. Note that although the ratings of different groups can be compared with analysis of variance (ANOVA), when any groupings have underlying continuous variation (e.g., BMI), one can simply correlate that measure with each of the food items. For example, a significant positive correlation between BMI and favorite food would show that those who weigh the most experience the most pleasure from their favorite foods. Overall, the HgLMS had the most significant correlations among the variables. The original 9-point hedonic scale had the fewest significant correlations and the modified 9-point scale fell in between.

Body mass index

For the HgLMS, Body Mass Index (BMI) correlated significantly with five other variables. For example, BMI and Ice Cream were positively correlated indicating that as a participant's BMI increased (i.e. they became more overweight) the participant enjoyed eating ice cream more. The other variables that correlated significantly with BMI had a negative R-value and could be categorized as bitter foods, except for pecan pie. Dark chocolate, blonde coffee, and espresso coffee all correlated negatively with BMI suggesting that with increasing BMI, there is a decreasing liking for these bitter foods. Pecan pie, however, does not fit this mold due to its sweetness. It could be

considered a dichotomous variable with some participants strongly liking it and others strongly disliking it.

There were fewer significant correlations found from either the modified 9-point scale or the original 9-point hedonic scale than the HgLMS. On the modified 9-point scale, BMI correlated positively with cheddar cheese; there were no other significant correlations with BMI. The original 9-point scale had no significant correlations among BMI and liking of foods. Thus the HgLMS was able to provide some meaningful comparisons across BMI. The original 9-point scale failed to do so. The modified 9-point scale performed better than did the original 9-point scale but not as well as the HgLMS. Figure 2-2 shows the correlation between BMI and Ice Cream and BMI and Dark Chocolate for each of the three scales for comparison.

Favorite and least favorite food

The results for favorite food for participants using the HgLMS show that for each significant correlation it had all but those with the three coffee items and least favorite food were positive. Those items positively correlated with favorite food were orange juice, sausage, fresh, ripe strawberries, steak (beef), whole milk, butter, cheddar cheese, sweets, candy, peanut butter, black coffee, ice cream, and all four orange juice samples the panelists tasted. The opposite can be said for coffee. As favorite food ratings increased, the taste for coffee decreased. Least favorite food had fewer significant correlations than favorite food. Those items positively correlated with least favorite food were mayonnaise, a fatty food, and the bitter foods dark chocolate, black coffee, including the two samples the participants tasted and the one they rated from memory, and the two grapefruit juice samples tested. Three sweet foods, sweets, candy, peanut butter, and ice cream, were negatively correlated with least favorite food.

As participants' liking of their least favorite food decreased, so did their liking of bitter foods; however their liking of sweet foods increased.

Favorite food on the modified 9-point scale correlated negatively with least favorite food. Favorite food correlated positively with orange juice, sausage, fresh, ripe strawberries, steak (beef), pecan pie, peanut butter, ice cream, and Simply Orange™. These foods range over a wide variety of flavors. Some, like orange juice and fresh, ripe strawberries, are sweet and fruity. Sausage and steak are salty and savory. Pecan pie, peanut butter, and ice cream are all sweet and fatty foods. Participants who used this scale liked these foods similar in rating to their favorite foods. Favorite food and least favorite food were negatively correlated as with the other two scales. Least favorite food was significantly correlated with mayonnaise, butter, cheddar cheese, Minute Maid™ orange juice, Simply Grapefruit™, Publix™ grapefruit juice, and blonde coffee in a positive direction. Panelists who gave increasingly low ratings for the disliking of least favorite food disliked all the aforementioned items more. The positively correlated items that panelists liked more the more they disliked their least favorite food were sweets, candy, and ice cream.

On the original 9-point scale, favorite food was correlated with butter, grapefruit juice, and blonde coffee all in the positive direction and Minute Maid™ orange juice in the negative direction. The participants who rated their liking of the favorite food high on the 9-point scale also rated their liking of butter, grapefruit juice, and blonde coffee high and Minute Maid™ orange juice low on the scale. Least favorite food had a significant positive correlation with mayonnaise and a significant negative correlation with sweets, candy. These two results were identical to the other two scales.

In sum, using favorite food to classify subjects the HgLMS was most successful and the original hedonic 9-point scale was least successful. The modified hedonic 9-point scale was intermediate at making across group comparison with favorite and least favorite food.

Supertasting

Supertasters have previously been shown to experience greater pleasure from their favorite food as well as greater displeasure from their least favorite food (Kalva and others 2014). Using the perception of the bitterness of quinine as a measure of supertasting, correlation of bitter with hedonic food ratings can reveal increased liking by supertasters. Table 2-3 charts the significant Pearson correlations for each of the three scales between bitter and each of the seventeen food related items on the ballot. A significant correlation is any with a p value of <.05. Those items with significant correlations to bitter show the distribution of liking by supertasters and others. This table shows that there are more significant correlations with those participants who used the HgLMS than either the modified or the original 9-point hedonic scale indicating that the HgLMS can more accurately identify supertasters' food preferences.

Figure 2-2 shows the relationship graphically for bitter correlated with Favorite Food and Least Favorite Food for the HgLMS, modified 9-point scale, and original 9-point scale. Bitter intensity had a significant correlation with favorite food for participants of the HgLMS corroborating a higher affinity of supertasters for their favorite food. The correlations that resulted from the participants using both the modified and the original 9-point scale were not significant for liking of favorite food and bitter intensity. Therefore, the HgLMS is able to differentiate among the supertasters who derive a greater pleasure in consuming their favorite food and others who derive less pleasure from

foods than the other scaling methods used. Both the HgLMS and the modified 9-point scale exhibited significant negative correlations for the rating of bitter intensity and liking of least favorite food in; in effect, supertasters dislike their least favorite food more than others. Oddly, the significant correlation between bitter intensity and liking of least favorite food for participants of the original 9-point scale was positive. This correlation can be interpreted as bitter intensity ratings increase so too does the liking of least favorite food.

In figure 2-3, the ratings of bitter intensity are correlated with black coffee and orange juice among the three hedonic scales. These two foods are representative of bitter and sweet foods, respectively, that humans are hard-wired to recognize. Humans react negatively toward bitter due to its potential for indicating consumption of poison while they react positively to sweet for its signal that glucose has been consumed, which has a role in providing energy for normal brain and body functioning. Thus, supertasters should have a stronger disliking of black coffee. This is evidenced by the HgLMS only where bitter and black coffee have a significant negative correlation. Bitter was not significantly correlated with black coffee for the modified 9-point scale or the original 9-point scale. Orange juice likeability correlated with bitter intensity for the HgLMS but not for either version of the 9-point hedonic scale. The modified 9-point scale was not able to be used for across group comparisons for either of these foods.

The original 9-point hedonic scale was unable to correctly identify supertasters due to its lack of context as compared to the HgLMS. Without the wider range for panelists to accurately rate the spread of their liking of the food related items, the 9-point scale can’t pick up the differing degrees of tasters. For example, in Figure 2-2 the

significantly positive correlation of bitter with the ratings of least favorite food is in the opposite direction than is statistically expected. If the original 9-point scale had been effective at identifying supertasters, it would have been a negative correlation.

Study 2

The results of study one showed that changing the original 9-point scale labels did not totally permit it to make valid across groups comparisons. The purpose of study 2 was to add nonfood items to see if that would improve the performance of the modified 9-point scale.

Demographics

There were 157 total participants in the second study. These participants were assigned to either the HgLMS or the modified 9-point hedonic scale. The total number of participants who showed to both the hedonic and intensity scale sessions was 84 for the HgLMS and 66 for the modified 9-point scale. These are the participants whose data was analyzed for statistical significance. The gender, age, and race distribution for the combined study data can be seen in table 3.

Correlation Regression

The main focus of the second research study was to determine if the inclusion of non food related items would increase the ability of the modified 9-point scale to identify the correlation regression relationships between the grouping variables used in the first study and the food related items on the ballot similarly to the HgLMS. The study focused on the supertaster analysis.

Across-Group Comparisons

In order to compare groups involving supertasting and hedonic ratings of favorite and least favorite foods in experiment 2, the proportion of supertasters, based on

quinine bitterness intensity ratings, should be similar in studies 1 and 2. Unfortunately, this was not the case as seen in Figure 2-4. Although the subjects who volunteered for the two studies appear to have been drawn from the same sample (students and staff of UF), there were markedly more supertasters in experiment two. Comparing the distributions of ratings of bitter in the two studies showed that the distribution for experiment 2 was skewed toward higher bitter ratings (chi square = 21/19; $p<.0001$). This was true for both males (chi square = 14.45; $p=.00014$) and females (chi square = 7.94; $p= .0048$). Therefore, the data were inappropriate to be used to make across group comparisons for supertasters.

Further data analyses revealed, however, that the relationships between the nonfood data among the sexes can be used to make across group comparisons. Because of the simplicity and ubiquitous use of the 9-point hedonic scale in many industry settings, the possibility that data from research using the scale can be used to make comparisons across groups of interest in those fields is very promising. In essence, very little time and effort may be required to educate these fields on how to make the 9-point scale a more useful research tool.

Nonfood items were added in order to provide a larger context for the 9-point scale. The intent was never to analyze the nonfood items. However, once the data was collected and some trial analyses we run some interesting associations were found for differences of BMI and sex with the nonfood items. Although it was not part of the original intent of the thesis, these associations are worth analyzing in the future.

Table 2-1. Study 1 participant demographic information.

Demographic	Participant Quantity
Gender	
Male	92
Female	196
Age	
17-22	278
23-29	8
30+	2
Ethnic Background	
Hispanic	66
Non-Hispanic	222
Race	
White or Caucasian	197
Black or African-American	28
Native American, Alaska Native, Aleutian	2
Asian/Pacific Islander	28
Other	33
Incidence of Middle Ear Infection	
No	210
Yes, but not serious	50
Yes, required antibiotics more than once	21
Yes, required tubes in ears	7

Table 2-2. Mean separation (Duncan's Multiple Range Test) of participant likability ratings of orange juice, grapefruit juice, and black coffee samples with the HgLMS, modified 9-point scale, and original 9-point scale.

		HgLMS		Modified 9-point		Original 9-point	
Orange Juice	Simply Orange™	19.69	a	5.61	a	7.49	a
	Tropicana™	20.13	a	5.58	a	6.94	a
	Sweetbay™	8.59	b	4.52	b	5.49	b
	Minute Maid™	8.11	b	4.22	b	5.02	b
Grapefruit Juice	Simply Grapefruit™	-11.71	a	3.46	a	3.75	a
	Publix™ Grapefruit Juice	-19.21	b	2.73	b	2.95	b
Black Coffee	Blonde	-21.39	a	3.42	a	2.83	a
	Espresso	-21.2	b	3.2	b	2.51	b

Table 2-3. Correlations of bitter intensity of quinine with food variables for the HgLMS, Modified 9-point scale, and Original 9-point scale.

	HgLMS		Modified 9-point scale		Original 9-point scale	
Variable	Pearson Correlation	p-value	Pearson Correlation	p-value	Pearson Correlation	p-value
Favorite Food	.509	.000				
Least Favorite Food	-.257	.012	-.215	.043	.200	.044
Mayonnaise						
Orange Juice	.386	.000				
Sausage						
Fresh, Ripe Strawberries	.353	.000				
Steak (beef)			.245	.021		
Dark Chocolate						
Whole Milk						
Butter						
Cheddar Cheese	.258	.012				
Sweets, Candy	.306	.003				
Pecan Pie						
Peanut Butter						
Black Coffee	-.461	.000				
Ice Cream	.442	.000				
Grapefruit Juice			-.235	.027		

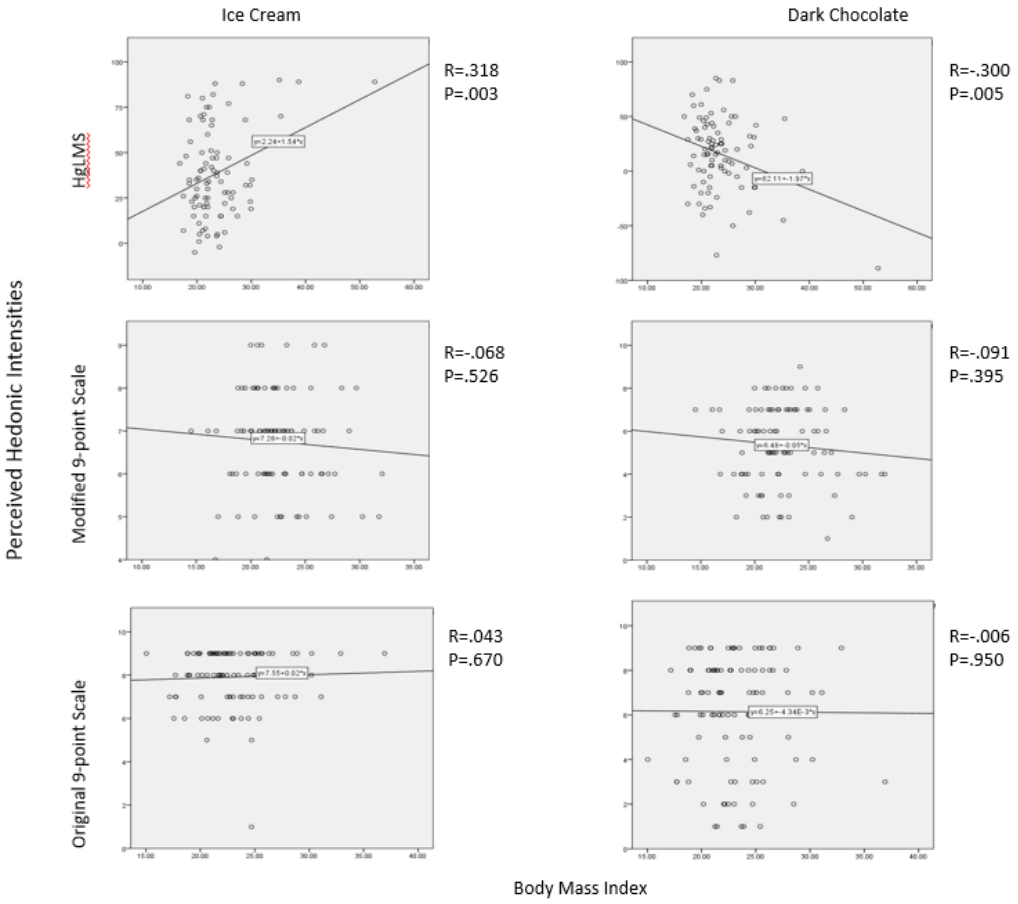

Figure 2-1. BMI correlated with liking of Ice Cream and Dark Chocolate for the HgLMS, the Modified 9-point hedonic scale, and the original 9-point hedonic scale.

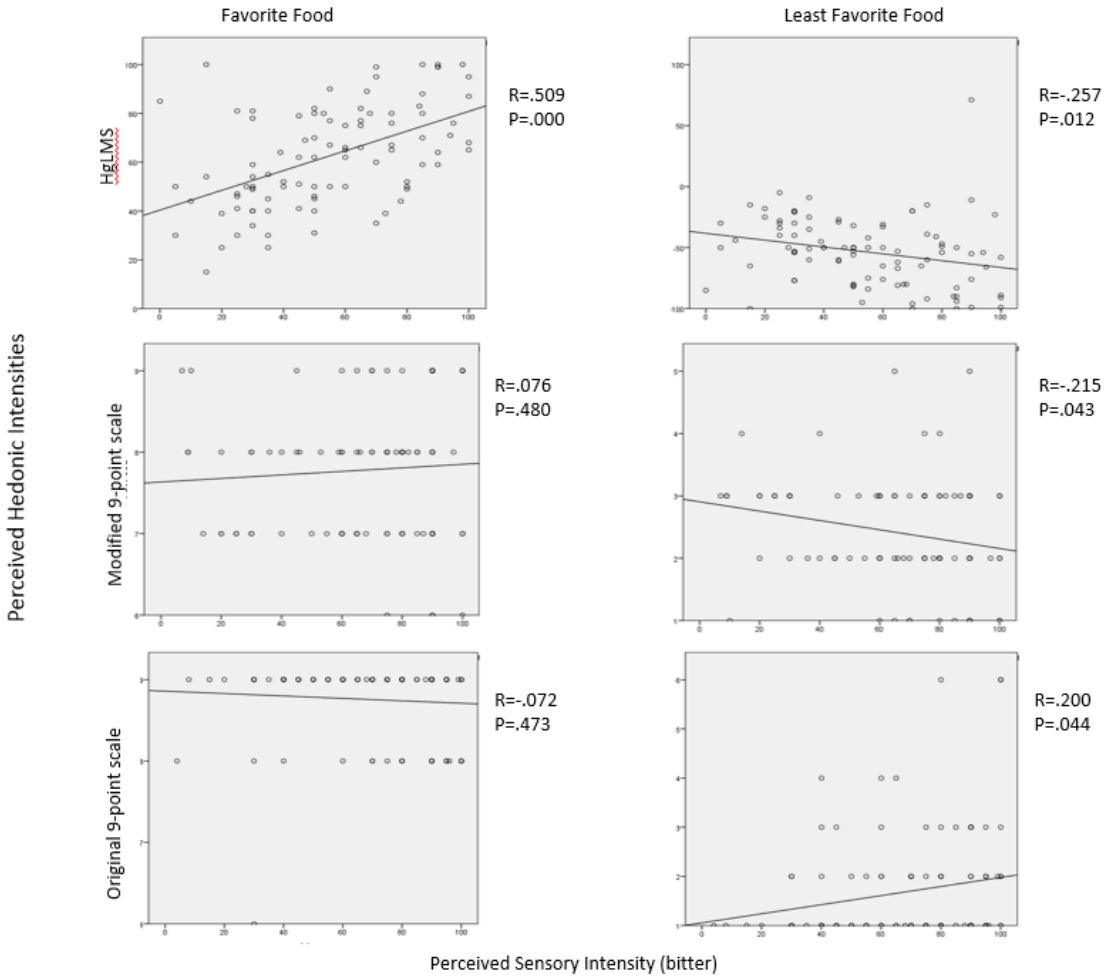

Figure 2-2. Hedonic ratings of participants of the HgLMS, Modified 9-point scale, and original 9-point scale correlated with ratings of perceived sensory intensity of quinine bitterness.

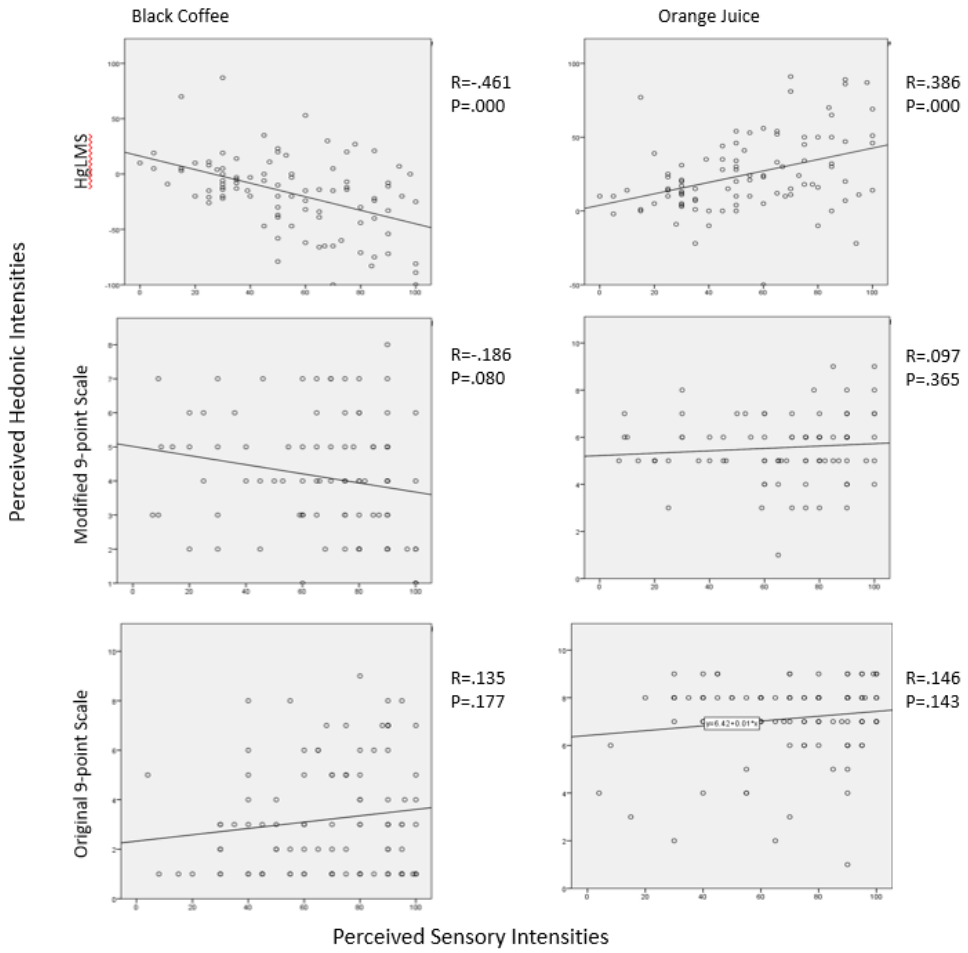

Figure 2-3. Hedonic intensities of black coffee (representing bitter foods) and orange juice (representing sweet foods) correlated with quinine bitter intensity ratings for the HgLMS, Modified 9-point scale and Original 9-point scale.

Table 2-4. Study 2 participant demographic information.

Demographic	Participant Quantity
Gender	
Male	52
Female	98
Age	
<17	1
17-22	138
23-29	9
30+	2
Race	
White or Caucasian	100
Black or African-American	7
Native American, Alaska Native, Aleutian	0
Asian/Pacific Islander	26
Other	17
Incidence of Middle Ear Infection	
No	
Yes, but not serious	
Yes, required antibiotics more than once	
Yes, required tubes in ears	

Table 2-5. HgLMS participant demographic information.

Demographic	Participant Quantity
Gender	
Male	34
Female	50
Age	
<17	1
17-22	76
23-29	5
30+	2
Race	
White or Caucasian	53
Black or African-American	4
Native American, Alaska Native, Aleutian	0
Asian/Pacific Islander	18
Other	9

Table 2-6. Modified 9-point hedonic scale participant demographic information.

Demographic	Participant Quantity
Gender	
Male	18
Female	48
Age	
<17	0
17-22	62
23-29	4
30+	0
Race	
White or Caucasian	47
Black or African-American	3
Native American, Alaska Native, Aleutian	0
Asian/Pacific Islander	8
Other	8
Incidence of Middle Ear Infection	
No	

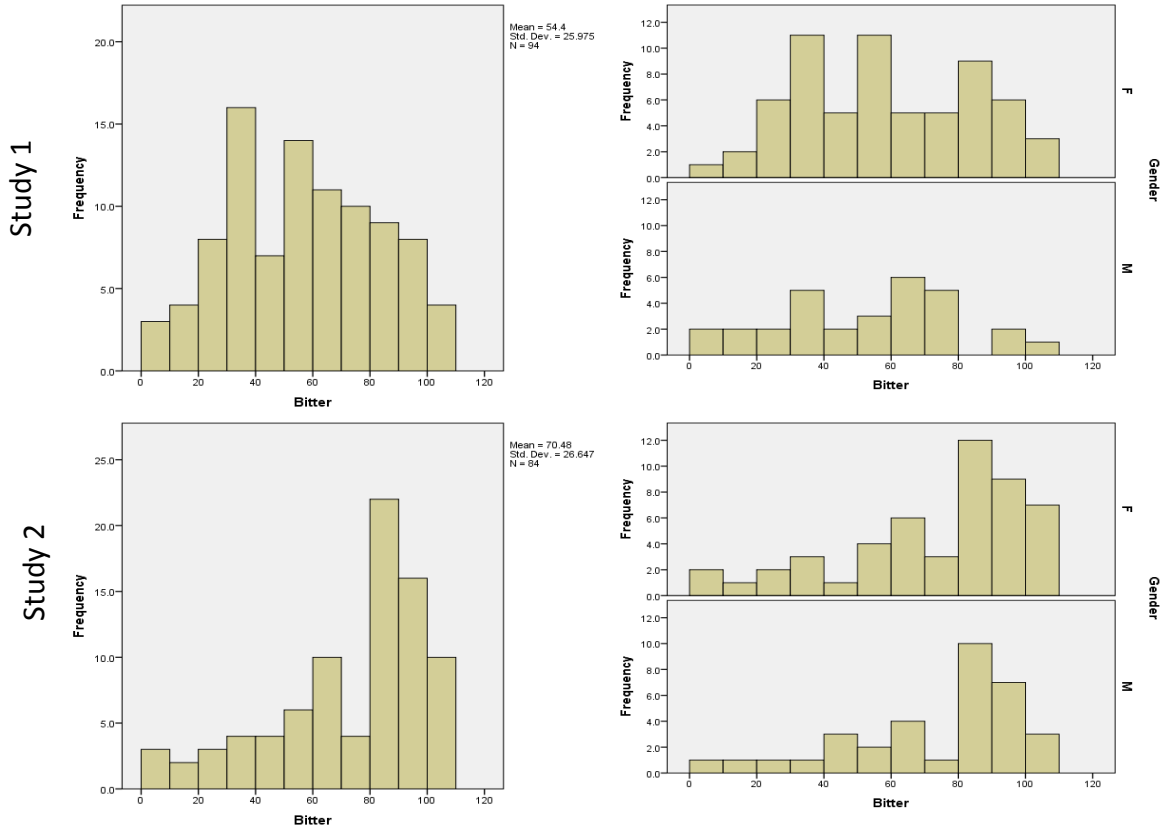

Figure 2-4. Perceived bitterness of quinine; distribution for Study 1 and Study 2 showing that the data was not distributed the same between studies.

CHAPTER 3
CONCLUSION

Study 1

The results from the first study indicated that the HgLMS, modified 9-point scale, and original 9-point scale can equally discriminate the differences in preference among the samples of orange juice, grapefruit juice, and black coffee the participants tested. Each of the hedonic scales showed the same mean separation of the samples' ratings exhibiting the differing qualities of the four samples of orange juice, two samples of grapefruit juice, and two samples of black coffee.

Across Group Comparisons

For all of the groups tested, the HgLMS found more significant correlations between the grouping variable and the items on the ballot. Whether using BMI or favorite food, the HgLMS was more effective at showing how people view foods and can be used to compare these results across different groups. Although having attempted to rehabilitate the 9-point hedonic scale to do the same for these food items, there was limited success in finding grouping variables that could be used to compare food preferences across them. The original 9-point scale was still perceived to be a food scale and without an independent standard by which to compare the results of food preferences across different groups, this scale can not be used for this purpose because the participants will always view the top of the scale as something food related.

Supertasting

Comparing the intensity of bitter with the hedonic ratings of favorite food and least favorite food was used as a way to make comparisons across groups of tasters. Supertasters, who rate bitter intensity the highest, also rate their liking of favorite food

the highest and their disliking of favorite food the lowest on the HgLMS. The original 9-point hedonic scale can not make these comparisons and the modified 9-point hedonic scale is only marginally better than the original.

Study 2

The second study was intended to show that by including nonfood items on the ballot the modified 9-point hedonic scale would be realized as more than just a food scale providing participants with a standard by which to anchor the scale. This would make the modified 9-point scale more effective at making across group comparisons than in the first study. However, due to the uneven distribution between study one and study two with regard to supertasters evidenced by the intensity ratings of bitterness, the original aim was not achieved. The potential value of a modified 9-point scale is still to be determined.

APPENDIX A
EXAMPLE SCALES USED IN STUDY ONE AND TWO

Hedonic general Labeled Magnitude Scale

Strongest Dislike of any kind ever experienced	**Neutral**	**Strongest Like of any kind ever experienced**
-100	0	100

Modified 9-point Hedonic Scale

Strongest Dislike of any kind ever experienced				**Neutral**				**Strongest Like of any kind ever experienced**
1	2	3	4	5	6	7	8	9

Original 9-point Hedonic Scale

Dislike extremely	Dislike very much	Dislike moderately	Dislike slightly	Neither like nor dislike	Like slightly	Like moderately	Like very much	Like extremely
1	2	3	4	5	6	7	8	9

APPENDIX B
STUDY 1 HEDONIC QUESTIONNAIRE

Orange Juice, Grapefruit Juice, & Black Coffee

Question # 1.

Please indicate your gender.

- ❍ Male
- ❍ Female

Question # 2.

Please enter your age.

Mage _____

Question # 3.

Please enter your age.

Fage _____

Question # 4.

Please enter your height (For example: If you are 5 feet and 3 inches in height, enter 503).

Height _____

Question # 5.

Please enter your weight in pounds.

Weight _____

Question # 6.

What is your ethnic background?

- ❍ Hispanic
- ❍ Non-Hispanic

Question # 7.

Which of the following best describes you?

- ❍ Asian/Pacific Islander
- ❍ Black or African-American
- ❍ White or Caucasian
- ❍ Native American, Alaska Native, Aleutian
- ❍ Other

Question # 8.

Have you ever suffered from middle ear infections?

- ❍ No
- ❍ Yes, but not serious
- ❍ Yes, required antibiotics more than once
- ❍ Yes, required tubes in ears

Question # 9.

How often do you drink orange juice?

- ❍ More than once a day
- ❍ Once a day
- ❍ 2-3 times a week
- ❍ Once a week
- ❍ 2-3 times a month
- ❍ Once a month
- ❍ Twice a year
- ❍ Once a year
- ❍ Less than once a year
- ❍ Never

Question # 10.

How often do you drink grapefruit juice?

- ❍ More than once a day
- ❍ Once a day
- ❍ 2-3 times a week
- ❍ Once a week
- ❍ 2-3 times a month
- ❍ Once a month
- ❍ Twice a year
- ❍ Once a year
- ❍ Less than once a year
- ❍ Never

Question # 11.

How often do you drink coffee?

- ❍ More than once a day
- ❍ Once a day
- ❍ 2-3 times a week
- ❍ Once a week

- ❍ 2-3 times a month
- ❍ Once a month
- ❍ Twice a year
- ❍ Once a year
- ❍ Less than once a year
- ❍ Never

S C A L E 1

Please click on the 'Continue' button below.

Question # 12.

Please type the strongest **LIKING OF ANY KIND YOU'VE EXPERIENCED** in the space below and remember that this experience will be 100 on your scale (**S C A L E 1**).

_____

_____

_____

Please click on the 'Continue' button below.

Question # 13.

Please type the strongest **DISLIKING OF ANY KIND YOU'VE EXPERIENCED** in the space below and remember that this experience will be -100 on your scale (**S C A L E 1**).

_____

_____

_____

Question # 14 - Sample <<Sample1>>

Please rate the following foods (from memory) using your **S C A L E 1**.

Hedonic general Labeled Magnitude Scale

Strongest Dislike of any kind ever experienced	**Neutral**	**Strongest Like of any kind ever experienced**
-100	0	100

Your favorite food

Your least favorite food

Mayonnaise

Orange Juice

Sausage

Fresh, ripe strawberries

Steak (beef)

Dark chocolate

Whole milk

Butter

Cheddar cheese

Sweets, candy

Pecan Pie

Peanut butter

Black coffee

Ice Cream

Grapefruit juice

Take a bite of cracker and a sip of water to rinse your mouth.
Make sure to do this before tasting and between each sample.

Make sure the number on the cup you are evaluating matches up with the number on the screen!

Your first set of samples will be ORANGE JUICE.

Question # 1.

Please use your **S C A L E 1** to rate how much you like each sample **OVERALL**

Simply Orange™

Tropicana Pure Premium™ Orange Juice

Sweetbay™ Orange Juice

Publix™ Orange Juice

Please lift your window to receive your GRAPEFRUIT JUICE samples.

Be sure to take a bite of cracker and a sip of water before tasting and between each sample.

WHEN ANSWERING ANY QUESTION, MAKE SURE THE NUMBER ON THE CUP MATCHES THE NUMBER ON THE MONITOR.

Please click on the 'Continue' button below.

Question # 1.

Please use your **S C A L E 1** to rate how much you like each sample **OVERALL**.

Simply Grapefruit™

Publix™ Grapefruit Juice

Please lift your window to receive your COFFEE samples.

*****CAUTION: coffee samples may be HOT!*****

Be sure to take a bite of cracker and a sip of water before tasting and between each sample.

WHEN ANSWERING ANY QUESTION, MAKE SURE THE NUMBER ON THE CUP MATCHES THE NUMBER ON THE MONITOR.

Please click on the 'Continue' button below.

Question # 1.

Please use your **S C A L E 1** to rate how much you like each sample **OVERALL**.

Blonde Coffee
Espresso Coffee

APPENDIX C
STUDY 2 HEDONIC QUESTIONNAIRE

Question # 1.

Please indicate your gender.

- ❍ Male
- ❍ Female

Question # 2.

Please enter your age.

Mage _____

Question # 3.

Please enter your age.

Fage _____

Question # 4.

Please enter your height (For example: If you are 5 feet and 3 inches in height, enter 503).

Height _____

Question # 5.

Please enter your weight in pounds.

Weight _____

Question # 6.

Which of the following best describes you?

- ❍ Asian/Pacific Islander
- ❍ Black or African-American
- ❍ White or Caucasian
- ❍ Native American, Alaska Native, Aleutian
- ❍ Other

S C A L E 1

1. Now, please take a few minutes to identify the strongest **LIKING** (i.e., pleasure) of any kind that you have ever experienced.

2. Once you have identified your strongest **LIKING** experienced,
please type it in on the next screen.

**3. Please remember to use the strongest liking that you've identified,
and typed in as the top of your scale (100).**

Please click on the 'Continue' button below.

Question # 7.

Please type the strongest **LIKING OF ANY KIND YOU'VE EXPERIENCED** in the space below and remember that this sensation will be 100 on your scale (**SCALE 1**).

_____

_____

_____

S C A L E 1

1. Now, please take a few minutes to identify the strongest **DISLIKING** (i.e., displeasure) of any kind that you have ever experienced.

2. Once you have identified your strongest **DISLIKING** experienced,
please type it in on the next screen.

3. Please remember to use the strongest disliking that you've identified,
and typed in as the bottom of your scale (-100).

Please click on the 'Continue' button below.

Question # 8.

Please type the strongest **DISLIKING OF ANY KIND YOU'VE EXPERIENCED** in the space below and remember that this sensation will be -100 on your scale (**SCALE 1**).

_____

_____

_____

Question # 1 - Sample <<Sample1>>

Please use your **SCALE 1** to rate the following experiences and food items from memory**.**

Hedonic general Labeled Magnitude Scale

Strongest Dislike of any kind ever experienced	**Neutral**	**Strongest Like of any kind ever experienced**
-100	0	100

Dark Chocolate

Eating your favorite food

Butter

Watching your favorite tv show

Whole Milk

Your most embarrassing moment

Orange Juice

Steak (beef)

Mayonnaise

A full, sound night's sleep

Going to a fun party with friends

Getting bad news from your doctor

Black Coffee

Successfully solving a very difficult problem

Riding a roller coaster

Being in a minor car accident

Peanut butter

Sweets, candy

The most inspired you've ever been by a lecture

Cheddar cheese

The most annoyed you've ever been

Ice cream

The most ashamed you've ever been of yourself

Getting a great deal on something

Smelling your favorite flower

Speaking in front of an audience

Meeting a major deadline on time

The most enthusiastic you've ever been about a hobby

Getting cut off in traffic

Grapefruit juice

Going to the doctor

Getting a good grade

Getting caught doing something you're not supposed to

Being made fun of by others

Sausage

Taking medication daily

The most nervous you've ever been

Fresh, ripe strawberries

The end of an important, special relationship

The shyest you've ever been

Pecan pie

The angriest you've ever been

Accomplishing an important goal

The death of a loved one

Listening to your favorite music

Having a deadline which seems impossible to meet

Spending time with loved ones

Not getting something you really wanted

Eating your least favorite food

The most amused you have ever been by an anecdote

APPENDIX D
SENSORY QUESTIONNAIRE

Question # 1.

Please indicate your gender.

- ❍ Male
- ❍ Female

Question # 2.

Please enter your age.

Mage _____

Question # 3.

Please enter your age.

Fage _____

Question # 4.

Please enter your height (For example: If you are 5 feet and 3 inches in height, enter 503).

Height _____

Question # 5.

Please enter your weight in pounds.

Weight _____

Question # 6.

Which of the following best describes you?

- ❍ Asian/Pacific Islander
- ❍ Black or African-American
- ❍ White or Caucasian
- ❍ Native American, Alaska Native, Aleutian
- ❍ Other

S C A L E 2

Now, we would like you to rate sensory intensities rather than liking/disliking. Rate the following sensations from no sensation (0) to the strongest sensation of any kind that you have ever experienced (100). *For example*, for some individuals, the brightest light ever seen (usually the sun) is the most intense sensation they have ever experienced. For others, the loudest sound ever heard (e.g., like a jet plane taking off nearby) might be the most intense. For still others, a particular pain might be the most intense. Whatever the most intense sensation is for you, that is the sensation that goes at the top of the scale.

Keep in mind that the scale is like a sensory ruler. If the sweetness of the sample is 1/10th of the way from zero to maximum (100), then enter it at 10. If it is twice as intense as that, it should be entered at 20, etc.

Please type your most intense sensation experienced (100 on your scale) on the next screen.

Question # 7.

Please type your most intense sensation experienced (100 on your **S C A L E 2**) in the space provided below.

_____

_____

_____

Question # 8.

Please, using your **S C A L E 2**, enter a number from zero (no sensation) to 100 (strongest sensation of any kind you've had) that best describes the experiences listed below.

Loudest sound ever heard	_____
Loudness of a conversation	_____
Brightness of a well-lit room	_____
Brightest light ever seen (usually the sun)	_____
Loudness of a whisper	_____
Brightness of a dimly-lit restaurant	_____

Take a bite of cracker and a sip of water to rinse your mouth.

Remember to do this <u>before you taste each sample</u>.

Please click on the 'Continue' button below.

Question # 1.

You will now be asked to rate the intensity of 4 solutions: **Sweet**, **Salty**, **Sour**, & **Bitter**.

Be sure to rate these using the SCALE 2 you just created.

Intensity

Sweet

Salty

Sour

Bitter

LIST OF REFERENCES

Aitken RC. 1969. Measurement of feelings using visual analogue scales. *Proc R Soc Med*. 62:989-993.

Bartoshuk LM, Duffy VB, Fast K, Green BG, Prutkin J, Snyder DJ. 2002. Labeled scales (e.g., category, Likert, VAS) and invalid across-group comparisons: what we have learned from genetic variation in taste. *Food Qual Pref*. 14:125-138.

Bartoshuk LM, Duffy VB, Chapo AK, Fast K, Yiee JH, Hoffman HJ, Ko C, Snyder DJ. 2004. From psychophysics to the clinic: missteps and advances. *Food Qual Pref*. 15:617-632.

Bartoshuk LM, Duffy VB, Green BG, Hoffman HJ, Ko C-W, Lucchina LA, Marks LE, Snyder DJ, Weiffenbach JM. 2004. Valid across-group comparisons with labeled scales: the gLMS versus magnitude matching. *Physiol Behav*. 82:109-114.

Bartoshuk L, Fast K, Snyder DJ. 2005. Differences in our sensory worlds: Invalid comparison with labeled scales. *Curr Direct Psychol Sci*. 14:122-125.

Blakeslee AF. 1932. Genetics of sensory thresholds: Taste for phenyl thio carbamide. *Proc Natl Acad Sci*. 18:120-130.

Borg GAV. 1982. Psychophysical bases of perceived exertion. *Med Sci Sports Exer*. 14:377-381.

Cardello A, Lawless HT, Schutz HG. 2008. Effects of extreme anchors and interior label spacing on labeled affective magnitude scales. *Food Qual Pref*. 19:473-480.

Centers for disease control and prevention. 2011. *Body Mass Index*. Centers For Disease Control and Prevention.

Chandrashekar J, Mueller KL, Hoon MA, Adler E, Feng L, Guo W, Zuker CS, Ryba NJP. 2000. T2Rs function as bitter taste receptors. *Cell*. 100:703-711.

Falconer DS. 1946. Sensory thresholds for solutions of phenyl-thio-carbaminde: results of tests on a large sample, made by R. A. Fisher. *Ann Hum Gen*. 13:211-222.

Fast K. 2004. *Developing a scale to measure just about anything: Comparisons across groups and individuals*. New Haven, CT: Yale University School of Medicine.

Fischer R, Griffin F. 1961. Quinine dimorphism among nontasters of 6-n-propylthiouracil. *Experientia*. 17:36-38.

Fischer R, Griffin F. 1964. Pharmacogenetic aspects of gustation. *Drug Res*. 14:673-686

Fox AL. 1931. Six in ten “tasteblind” to bitter chemical. *Sci News Lett*. 9:249.

Green BG, Shaffer GS, Gilmore MM. 1993. A semantically-labeled magnitude scale of oral sensation with apparent ratio properties. *Chem Senses*. 18, 683-702.

Green BG, Shaffer GS, Gilmore MM. 1993. Derivation and evaluation of a semantic scale of oral sensation magnitude with apparent ratio properties. *Chem Senses*. 18:683-702.

Guo SW, Reed DR. 2001. The genetics of phenylthiocarbamide perception. *Ann Hum Bio*. 28:111-142.

Hayes JE, Bartoshuk LM, Kid JR, Duffy VB. 2008. Supertasting and PROP bitterness depends on more than the Tas2r38 gene. *Chem Senses*. 33:255-265.

Herz RS. 2003. The effect of verbal context on olfactory perception. *J Exp Psychol*. 132:595-606.

Herz RS. 2012. *Olfaction*. In: Wolfe JM. Sensation and Perception. 3rd ed. Massachusetts: Sinauer Associates, Inc. 399-430p.

Jaeger SR, Cardello AV. 2009. Direct and indirect hedonic scaling methods: A comparison of the labeled affective magnitude (LAM) scale and best-worst scaling. *Food Qual Pref*. 20:249-258.

Jones LV, Peryam DR, Thurstone LL. 1955. Development of a scale for measuring soldiers' food preferences. *J Food Sci*. 20:512-520.

Kalva JJ, Sims CA, Puentes LA, Snyder DJ, Bartoshuk LM. 2009. Comparison of the hedonic general labeled magnitude scale with the hedonic 9-point scale. *J Food Sci*. 101

Kim UK, Jorgenson E, Coon H, Leppert M, Risch N, Drayna D. 2003. Positonal cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science.* 229:1221-1225.

Keller A, Zhan H, Chi Q, Vosshall LB, Matsumani H. 2007. Genetic variation in a human odorant receptor alters odour perception. *Nature*. 449:468-472.

Lawless HT, Heymann H. 2010. *Sensory evaluation of food*. 2nd ed. New York: Springer Science+Business Media, LLC. 596 p.

Lea P, Rodbotten M, Naes T. 1995. Measuring validity in sensory analysis. *Food Qual Pref*. 6:321-326.

Lim J, Wood A, Green BG. 2009. Derivation and evaluation of a labeled hedonic scale. *Chem Senses*. 34:739-751.

Lim J, Fujimaru T. 2010. Evaluation of the labeled hedonic scale under different experimental conditions. *Food Qual Pref*. 21:521-530.

Marley AAJ, Louviere JJ. 2005. Some probabilistic models of best, worst, and best-worst choices. *Journ Math Psych*. 49:464-480.

Mueller KL, Hoon MA, Erlenbach I, Chandrashekar J, Zuker CS, Ryba NJ. 2005. The receptors and coding logic for bitter taste. *Nature*. 434:225-229.

Nicolas L, Marquilly C, O'Mahony M. 2010. The 9-point hedonic scale: Are words and numbers compatible? *Food Qual Pref*. 21:1008-1015.

Romano R, Brockhoff PB, Hersleth M, Tomic O, Naes T. 2008. Correcting for different use of the scale and the need for further analysis of individual differences in sensory analysis. *Food Qual Pref*. 19:197-209.

Rozin P. 1982. Taste-smell confusions and the duality of the olfactory sense. *Percept Psychophys*. 31:397-401.

Schild D, Restrepo D. 1998. Transduction mechanism in vertebrate olfactory receptor cells. *Physiol Rev*. 37:369-375.

Schutz HG, Cardello AV. 2001. A labeled affective magnitude (LAM) scale for assessing food liking/disliking. *J Sens Stud*. 16:117-159.

Stein LJ, Cowart BJ, Epstein AN, Pilot LJ, Laskin CR, Beauchamp GK. 1996. Increased liking for salty foods in adolescents exposed during infancy to a chloride-deficient feeding formula. *Appetite*. 27:65–77.

Steiner JE. 1973. *The gustofacial response: Observation on normal and anencephalic newborn infants.* In: Bosma JF, author. Oral sensation and perception: Development in the fetus and infant: Fourth symposium. Oxford, England: US Government Printing Office. 419pp.

Stevens SS. 1957. On the psychophysical law. *Psychol Rev*. 64:153-181.

Wheatcroft PEJ, Thornburn CC. 1972. Toxicity of the taste testing compound phenylthiocarbamide. *Nature New Bio*. 235:93-94.

Wolfe JM, Kleunder KR, Levi DM, Bartoshuk LM, Herz RS, Klatzky R, Lederman SJ, Merfeld DM. 2012. *Sensation and Perception*. 3rd ed. Massachusetts: Sinauer Associates, Inc. 399-455.

BIOGRAPHICAL SKETCH

Rachel Glintz was born and raised in Wellington, Florida. She has one brother, Adam, a first lieutenant in the Marine Corps. Her parents excitedly anticipate her graduation and future career. Rachel graduated with her Bachelor of Science in food science and human nutrition in 2012 from the University of Florida. Upon completion of her Master of Science, Rachel looks forward to a prosperous career in the food science industry in the area of product development.