ABSTRACT

ROVAI, DOMINIC. Oat Milk, Protein Beverages, Creamer, Gin: Understanding Consumer Preferences, Novel Evaluation Techniques, and the Impact of Context on the Sensory Evaluation of Beverages. (Under the direction of Dr. MaryAnne Drake).

Beverages have increased in sales and popularity in recent years with consumers choosing them for a variety of occasions. Sensory science provides a set of tools to help understand consumer motivations and preferences, and to evaluate and compare beverages to one another to inform the decision-making process.

Sensory evaluation is a critical component of a quality control (QC) program. Sensory encompasses many techniques, each with a unique objective and application. Chapter 1 reviewed sensory methodologies and their application and offers guidance for implementing a sensory QC program in dairy processing operations. Descriptive analysis (DA) is a robust tool that is versatile and can be adapted to aid in monitoring the quality control of dairy products. Degree of difference scaling and DA with limited attributes are viable modifications of DA that can be applied to a QC setting. Implementing sensory tools like these will help ensure consistent product sensory quality.

Chapter 2 evaluated consumer perception of oat milk through an online survey (n=415) and qualitative focus groups (n=24). Trial motivations for oat milk are that it is new, trendy, and fits within dietary restrictions. Repeat purchase motivations are motivated by the unique flavor and texture that oat milk provides compared to dairy milk and other milk alternatives. The ideal oat milk for consumers is original flavor with claims of excellent source of calcium, vitamins A & D, and free from artificial flavors and colors ($p<0.05$).

Consumers drink coffee creamers added to a cup of coffee. When evaluating consumer liking, the context and use of a carrier may influence liking. Chapter 3 determined the impact of serving method (ratio of creamer:coffee) and type of coffee, on consumer liking of coffee creamers. In phase I, consumers evaluated creamers (n=4) using three methods: fixed amount (n=127), free pour (n=120), and warm-up (n=122). No significant creamer*method interaction was observed for overall liking ($p>0.05$). In phase II, consumers (n=134) evaluated creamers in two different coffees: light roast and dark roast. No significant creamer*coffee interaction was observed for overall liking ($p>0.05$), but significant interactions ($p<0.05$) were observed for many diagnostic attributes (appearance, aroma, flavor, mouthfeel, creaminess). Any of the serving methods were suitable for creamer evaluation, but researchers should consider the impact of coffee on creamer evaluation.

Temporal methods provide objective documentation of product attribute changes over the course of evaluation. In chapter 4, gin was used as a case study to compare three temporal methods: temporal dominance of sensations (TDS), temporal check-all-that-apply (TCATA) and temporal ranking (TR). TDS differentiated gins based on dominant flavor attributes but was ineffective at distinguishing many attributes. TCATA captured differences in subtle flavor attributes but did not discriminate samples by juniper flavor. TR differentiated gins based on differences in juniper flavor intensity and documents additional differences in complex gins, but did not discriminate subtle flavor attributes.

Ingredients are added to protein beverages during processing to preserve the quality and prevent undesirable protein aggregation. In chapter 5, an online survey (n=405) and qualitative focus groups (n=25) were used to understand consumer perception of the ingredients in protein beverages. The amount of protein and type of protein were key factors in protein beverages. The most appealing sources of protein were plant protein, whey protein, and milk protein ($p<0.05$). Natural sweeteners (agave, monk fruit, cane sugar) were the most appealing sweeteners ($p<0.05$).

Fibers and starches were more appealing than gums (carrageenan, gellan gum) ($p<0.05$). Stabilizers were the least desirable class of beverage ingredients, with sodium and potassium phosphates the least desirable ($p<0.05$). Consumers preferred a label with minimal ingredients and the omission of ingredients they did not recognize.

Oat Milk, Protein Beverages, Creamer, Gin: Understanding Consumer Preferences, Novel Evaluation Techniques, and the Impact of Context on the Sensory Evaluation of Beverages

by
Dominic Rovai

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Food Science

Raleigh, North Carolina
2024

APPROVED BY:

Dr. MaryAnne Drake Committee Chair	Dr. Haotian Zheng
Dr. Clinton Stevenson	Dr. Daniel Poole

BIOGRAPHY

Dominic Rovai was born on November 24th, 1997. Dominic grew up in Aptos, California and attended Aptos High School. After high school, Dominic went to California Polytechnic State University (Cal Poly), San Luis Obispo where he received his B.S. and M.S. in Food Science in 2021 through a blended Master's program. At Cal Poly, Dominic worked under the direction of Dr. Amy Lammert where he found and interest in sensory science running consumer tests for clients and conducting researched on insects for human consumption. Following graduation, Dominic moved to Raleigh, North Carolina to pursue his PhD in Food Science at North Carolina State University under the direction of Dr. MaryAnne Drake.

ACKNOWLEDGMENTS

Thank you Tenaya for moving across the country and supporting me every step of the way. Thank you to the current and previous MAD lab members I have been able to work with for your friendship, support, and encouragement. Getting things done in this lab requires collaboration, teamwork, and resilience, and I am grateful for the great group of peers I have been able to work with. I am also grateful for the guidance provided by my committee members: Dr. Zheng, Dr. Stevenson, Dr. Poole. Lastly, thank you Dr. Drake for your mentorship and dedication to the lab and my success. You exemplify the benefits of hard work and critical thinking, and I am grateful for the guidance you have provided me.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 5

LIST OF FIGURES

CHAPTER 5

CHAPTER 1:

Sensory Methodologies and the Quality Control of Dairy Products

Sensory Methodologies and the Quality Control of Dairy Products

D. Rovai, M. A. Drake

Southeast Dairy Foods Research Center, North Carolina State University, Raleigh, NC

Contact information for Corresponding author
MaryAnne Drake
Box 7624, Department of Food Science
North Carolina State University
Raleigh, NC 27695-7624
Phone: 919-513-4598
Fax: 919-513-0014
Email: mdrake@ncsu.edu

Choice of journal/topic: Elsevier book chapter

Abstract

Sensory quality is an important aspect of food quality control (QC) as it drives repeat purchase decisions for consumers. Many sensory methodologies exist, and these tools have been adapted to specific objectives. Defect judging is an example of a traditional sensory tool that has been used to document product quality. Descriptive analysis (DA) has begun to replace defect judging due to its robustness, versatility, and adaptability to the QC setting. DOD scaling and DA with limited attributes are powerful QC tools. These methods benefit from a trained panel to objectively document products and their attributes. QC can leverage components from DA to facilitate panelist training and implementation. This chapter summarizes best practices and practical approaches for implementing sensory tools within the QC program to ensure consistent product sensory quality.

1.1 Introduction

Quality control (QC) in food production systems focuses on maintaining standards in all aspects of the product–technological, physical, chemical, microbiological, nutritional, and sensorial. The implementation of a quality control program helps ensure wholesome, safe, consistent, and appealing food products. Often, quality control programs are focused on instrumental or physical analyses (pH, brix, Aw, moisture content, PV, microbial counts, etc.). A target can be set for each of these values with tolerance limits that warrants actions if the values fall out of range. These measures are necessary to ensure product safety and signal potential process variation that may need to be addressed. However, often overlooked in the QC process is the need to evaluate the sensory properties of the product. When it comes to ensuring the quality of a product, safety is crucial, but it is equally important that the product tastes good and is consistent with consumer expectations. Ultimately, consumers make repeat purchase decisions based on how much they like the taste of the product. Thus, understanding the sensory properties of foods and ensuring the product meets these expectations during production is essential when implementing a QC program.

Sensory evaluation at its core is the science of human perception. Food products undergo constant sensory evaluation before and during consumption. The dairy industry yields many different types of products at different production stages (ingredients, additives, final products, etc.), each facing different types of challenges. Depending on the nature of the product, the need for sensory evaluation may differ. For example, the sensory methodologies necessary for ensuring the quality of raw milk are different from the set of tools needed to validate a new supplier of vanilla flavor in a soft serve ice cream mix. Sensory methods have evolved over time to address different types of questions. With many different sensory methods, it is critical to make sure that the right sensory tool is being used. This chapter will explore the evolution of sensory methodologies, provide guidance for choosing the right tool for the job, and list practical steps on how sensory techniques can be adapted and utilized within a QC operation.

1.2 Evolution of Sensory Methodologies

Traditional sensory techniques include product grading and scorecard judging (Drake and Clark, 2023). These tools were developed by the dairy industry to train the incoming workforce and ensure consistent product quality. The first example of collegiate dairy judging took place in 1916 on butter (Clark, 2023), and milk and Cheddar cheese were added a year later (Bodyfelt et al., 2008). Using these methods, products are graded based on the presence and severity of common defects and then assigned an overall quality grade based on the summation of deductions across all defects (Alvarez, 2023). These methods are limited in their practical application beyond gross quality screening for a variety of reasons. The primary concern with these methods is that there is a lack of consistency in how scores are assigned. Ideally, identification of defects should be an objective process where all assessors are standardized in how they are identifying and scoring attributes. In reality, many defects commonly used are dated and not well defined, making it difficult to calibrate panelists or apply statistical analyses (Drake, 2023). In addition, there is inconsistency regarding the weight of point deductions, both across defects and within any single defect (Drake, 2023). Since scores are defect-based, quality scores cannot provide comprehensive product profiles or predict consumer liking. These tools still find applications today in situations where a large number of products or samples may require sensory quality screening in a brief period of time.

While traditional sensory techniques persist today, their limitations facilitated the development of descriptive methods. Descriptive Analysis (DA) techniques are analytical techniques that utilize trained panelists (typically 6-12) operating as a detecting instrument to document the intensity of product attributes. One research study used both industry graders and a trained panel to evaluate Cheddar cheeses manufactured using different starter cultures and found that no consistent differences were identified by graders, but trained panelists differentiated samples by several flavor attributes (Shakeel-ur-Rehman et al., 2008). Because of the discriminatory power of DA and the lack of statistics that can be applied to traditional grading scorecards (Bodyfelt et al., 2008), DA has replaced quality judging techniques in research publications (Schiano et al., 2017). Objective documentation of flavor and/or texture profiles of food products via DA is a powerful tool that allows product developers to understand how their product is similar and/or different from a competitor, can be combined with analytical techniques to identify sources of off-flavors, and can be combined with consumer data to understand drivers of liking. The flavor profile method, considered the foundation of modern DA, was introduced by Arthur D. Little Co. in 1957 (Drake and Clark, 2023). Since then, Tragon Corp introduced quantitative descriptive analysis (QDA) in 1974 (Stone & Sidel, 2004) and Gail Civille introduced the Spectrum DA method in 1979 (Society of Sensory Professionals, 2021). Today, QDA and Spectrum, and hybrid approaches based on these techniques, are the most widely used DA methods. DA is considered one of the most powerful tools when it comes to sensory evaluation because of the versatility of the information it can provide and its application in answering many different sensory questions. As such, derivations of DA are commonly used in QC labs, which will be discussed in the next section.

In the 1940s, the triangle difference test was introduced (Drake and Clark, 2023). Difference tests answer the question: “does a difference exist between two samples?”. These tests are limited to answering this single question but provide an efficient and powerful means of doing so. Indeed, DA and consumer testing also allow leverage of statistics to determine if there are perceived differences in samples, but DA requires extensive time and training, and consumer testing is expensive and necessitates recruiting a lot of panelists. Difference tests, on the other hand, generally require only a few minutes of time from 25-50 untrained panelists. Thus, difference tests are often a precursor to further analysis. If a significant difference does not exist between samples there is no need to follow up with DA or consumer testing. Difference tests are implemented when companies make process changes, ingredient substitutions, or look to approve additional ingredient vendors. In all these instances, companies hope to improve efficiency and flexibility in the supply chain and/or reduce costs, but they want to make sure that these changes do not result in noticeable differences in their product. Today, a variety of difference tests exist such as affective forced choice (AFC), duo-trio, triangle, and tetrad tests. The decision of which difference test to use depends on several factors including the number of samples, the quantity of samples available, testing conditions, and historical use (i.e. your company has been using a particular test for a number of years). More detail about difference tests can be found in Lawless and Heymann (2010) and Meilgaard et al. (2025).

Affective consumer tests are a set of sensory techniques aimed at assessing subjective liking of products. In 1949, the US Army Quartermaster Laboratory developed the 9-point hedonic scale to document liking of food by U.S. soldiers (Peryam and Pilgrim, 1957). This scale, which ranges from 1 = dislike extremely to 9 = like extremely, has become a popular way to measure consumer liking and preferences. Today, understanding consumer liking of samples is critical to the R&D process as products are typically benchmarked against competitors and/or current products to predict consumer acceptability in the market. Consumer testing methods can be either quantitative or qualitative in nature. Quantitative methods utilize a ballot where consumers use scales to indicate degree of liking (such as the 9-point hedonic scale) and preferences for blinded samples. In addition, just-about-right (JAR) scales are commonly used to determine if supplemental texture or flavor attributes are perceived as too much, too little, or JAR (Lawless and Heymann, 2010). These are often conducted as a central location test (CLT) where consumers come to a central location like a testing site but can also be conducted as a home usage test (HUT). CLTs help control extraneous variables by ensuring all samples are prepared the same way and served in a similar environment to maximize discriminatory power, while HUTs embrace the context of how the product would be used at home to understand pain points and how consumers interact with and use these products in the comfort of their own homes. Quantitative tests yield numeric scores that can leverage statistics to determine significant product preferences among consumers. Qualitative methods include tests like focus groups and interviews. These techniques are often used in conjunction with quantitative tests and provide insights into the nuances of what consumers like/dislike, how they anticipate using it, and other emotional responses that are not captured by a simple product questionnaire. Additional guidance on best practices for understanding consumer preferences are provided in a review by Drake et al. (2023).

Threshold tests are another type of sensory methodology that evolved to provide context to analytical data on volatile flavors. In many instances, it is useful to understand sensory detection limits of specific compounds that cause an off flavor in a product. Techniques such as gas chromatography (GC) are often used to identify and quantify flavor-active compounds (Ismail, 2017). Using a GC, scientists can identify and quantify volatile flavor compounds in a sample. However, just because a compound is present in a sample does not mean that it is perceived or detected by humans. Threshold tests determine the lowest concentration of a compound where a sensory difference is detected (detection threshold). The threshold value is determined by a series of triangle difference tests. In each triangle test, three samples are presented, with one of the samples spiked with a known concentration of the compound of interest. Across the series of tests, the concentration of the compound increases by a concentration step factor. Panelists evaluate (either aroma or taste) one set at a time and are asked to indicate the one sample that is different among the three samples. After evaluating the entire series, a threshold value is calculated per panelist based on the lowest concentration they were able to correctly identify. The population threshold is also estimated based on the threshold values for all panelists. In this way, understanding the threshold of a compound (matrix specific) allows you to determine if the concentration of compound in the sample is high enough to be considered flavor active.

1.3 Choosing the Right Tool for the Job – Quality Control

Each of the sensory methods evolved to fill a gap and provide a tool for a specific job. Indeed, each method is effective when used properly, but should not be utilized without thorough understanding of the usefulness and limitations of each method. These considerations are summarized by Table 1.1. While there are many other sensory techniques and variations of those listed, these methods broadly cover the categories of sensory techniques.

The job of sensory quality programs is to understand if products are objectively within quality range. While Quality Assurance (QA) focuses on proactive approaches to prevent defects, QC focuses on identifying any defects that may arise (and correcting them). From a sensory perspective, this means that the goal of the QC program is to identify defects or attributes and determine if they fall within acceptable range. DA is the most robust tool for identifying deviations and quantifying differences between products but the time and cost requirement for a comprehensive panel may be prohibitive. As such, sensory quality approaches that involve trained or experienced assessors using an identified set of attributes or defects and a specific numerical scale adapted to the specific product and plant environment can be a viable option.

Modified DA methodologies include approaches such as degree of difference (DOD) or difference from control (DFC) and cut down DA (using a limited number of attributes) have emerged as a practical tool in QC. DOD scaling considers the total sensory difference between two products (Meilgaard et al., 2025). In QC applications, DOD–also referred to as difference from control (DFC)--has been used due to its flexibility. The DOD score quantifies how different a sample is from the control, which can then determine if the product is within spec or if follow-up action is needed. It is best practice to include a blinded control within the sample set to serve as a baseline for the panel (i.e. the panel should not indicate a difference between the control and the blinded control sample). Most importantly, the control or gold standard or target sensory profile must be defined so there is a "target" or control that the assessor is judging samples relative to. Ideally, common defects should also be identified and clearly described. Different scales can be used to for DOD/DFC, including both unidirectional and bidirectional scales. Unidirectional scales (Figure 1.1) are more commonly used to capture total differences, while a series of bidirectional scales (Figure 1.2) may be used to score DOD from control for different attributes. In this way, a bidirectional DOD scale provides some insight into the nature of differences, if they exist. When unidirectional DOD scales are used in the QC setting, they may be used with cut down DA (using a limited number of attributes). The list of attributes for cut down DA would encompass key product attributes.

Regardless of the methodology used, QC aims at objective documentation of product attributes. Aside from consumer tests–which measure subjective consumer liking–all the tools mentioned benefit from a panel composed of trained panelists. Research by Carey et al. (2020) proposed a training program for fluid milk QC focused on defects in fluid milk and emphasized the importance of initial training of panelists, as well as the benefits of retraining certain attributes to improve discriminatory power. Because DA is the gold standard for product discrimination and product profiling, the next section will walk through some of the components of DA, specifically the approach to panel training. Understanding the theory of how to train and implement a DA panel directly translates to benefit a sensory QC program.

1.4 Guidance and Considerations for Training a DA Panel

The quality of data generated by DA is dictated by the experience of the panel and the amount of training (Drake et al., 2007). The amount of training required is dependent on the complexity of the product and the attributes of interest. Training can be as little as a few hours if the product and attributes are simple but is typically much longer (20-40 hours minimum). The purpose of training is to calibrate the panel so that all panelists are familiar with the characteristic attributes of the product (flavor, aroma, texture) and scale them in a similar way. Panel training is typically composed of three steps: 1) attribute identification, 2) scaling, 3) merging. The goal of attribute identification is to identify the key attributes present in the products and ensure panelists are speaking the same language. To do this, a sensory lexicon is used specific to the product being evaluated (Drake and Civille, 2003). Lexicons provide a standardized vocabulary that serves as a communication tool for understanding the differences between products (Lawless and Civille, 2013). Sources of published lexicons on various dairy products are included in Table 1.2. These lexicons provide standardized terminology for discussing the product, and references (chemical and/or product references) for standardized stimuli to help train panelists and calibrate them on the lexicon terms. Initial training sessions are carried out which focus solely on attribute identification using the lexicon and references.

Once panelists feel comfortable with attribute identification, the next step is scaling attributes. Scaling refers to how attribute intensities are ranked. The philosophy behind which scale to use is a point of contention among different DA methods, which will be discussed later in this section. Scaling training uses standardized solutions spiked with known amounts of stimuli. For example, water may be spiked with standardized concentrations of sugar or salt to exemplify how panelists should scale saltiness or sweetness attributes (ex. Labeled cups with sweet 2, sweet 4, sweet 6 solutions). Panelists may alternate between labeled references to calibrate themselves, and blinded references to test their ability to use the scale.

The final step of training is merging. This requires panelists to merge the skills of attribute identification and scaling. During this step, panelists use the lexicon to identify product attributes in the product matrix, and then use the scale to provide a value for the intensity of the perceived attribute. Simultaneously identifying and scaling attributes can be a cognitively difficult task for new panelists. As such, most of the time spent training will be spent practicing merging. It is common to keep references on hand so that panelists can go back and refresh themselves about the attribute they are looking for in the products. There is no set rule that indicates a panel is performing optimally or individual panelists are performing well, but rather it is based on the judgment of the panel leader (Drake et al., 2007). Is the panel able to function for the job it was applied to accomplish? That function might be very simple such as differentiating a few texture attributes or it may be more complex such as utilizing a lexicon of >10 flavor aromatics. Optimal panel performance may depend on the panelists, the product, the attributes and the ultimate objective. Research by Chambers et al. (2004) evaluated the effect of panel training time on their ability to discriminate tomato sauce attributes and found that texture and some flavor attributes required minimal training (4 hours), while other attributes required extensive training (60+ hours).

While all DA methods have the same end goal in mind, there are two popular DA methods that differ based on their philosophy regarding panel training and implementation. These two methods are the Spectrum method and the QDA method. There are many nuances between these methods (Meilgaard et al., 2025), but we will focus on key differences. The Spectrum method uses a universal scale (0 - 15). This means that a sweet 3 in yogurt is the same intensity as a sweet 3 in ice cream mix. Additionally, the stimuli intensity for a 3 in sweetness should be equally intense as the stimuli intensity for a 3 in cooked flavor. This scale method allows for consistent scaling across product categories. The initial training investment for a Spectrum panel is extensive, but easily adaptable once they are well trained. QDA is distinct from Spectrum in that the panels are product specific, not universal. QDA utilizes line scales that are unique to each product panel. For example, if a panel was evaluating yogurt, the range on the scale would reflect the lowest and highest intensity for each attribute as observed in the yogurt category. Thus, for sweetness, the upper end of the scale would capture the sweetest yogurts tasted by the panel. The upper end of a scale for an ice cream mix panel would also capture the sweetest ice cream mixes, but you would not expect the sweetness intensity for the sweetest yogurts to have the same sweetness intensity as the sweetest ice cream mixes. Another difference between Spectrum and QDA is the use of references. Both methods use references in training during attribute identification, but the Spectrum method uses chemical references (i.e. diacetyl for buttery aroma), while QDA favors product references (i.e. a bag of microwave popcorn for buttery aroma). Finally, these two methods differ in how the panel is led. In the Spectrum method, the panel leader is highly experienced and tastes along with the panel. They help calibrate, train, and guide the panel, and participate in product scoring. In a QDA panel, the panel leader acts as a moderator. They do not taste or participate in the panel, just facilitate the training and discussion to make sure panelists have appropriate references and keep the training progress on track.

While these two DA methods present different views regarding ideology in training a panel, both methods have proved to be effective at discriminating samples based on objective differences in the intensities of product attributes. Spectrum panelists are more adaptable due to the universal scale and the trained panel profiles generated can be compared across panels (i.e. the trained panel profile of a product should not change based on the set of samples it was evaluated with/compared to). However, the Spectrum method requires extensive training and must be led by an individual who is extremely trained and familiar with tasting and Spectrum scaling. In contrast, QDA requires comparatively less training and does not require the level of expertise of the panel leader (though more experience is always better!). The downside of QDA is that results are relative and specific to the product category and the products that were compared by the panel (i.e. we would expect the trained panel profile of a product to change based on the set of samples it was evaluated with/compared to). In practice, some programs adopt hybrid techniques where they utilize some of the characteristics from each methodology to create an adaptive method that fits within their objectives and capabilities.

1.5 Practical Application of DA -- Dairy Products and Quality Control

The robustness and versatility of DA means it is effective in many practical research applications. Figure 1.3 shows a few examples of how DA has been applied to gain insights about various dairy products.

While DA has many uses, in a QC program the sole purpose of DA is defect identification—both the presence of off-flavors and other attributes outside of tolerable limits. With this narrow objective, training and evaluation can be targeted towards this application.

In a perfect world, every member of the production team would be an expert in tasting dairy products, recognizing defects, and have extensive knowledge about the manufacture process and where defects may arise. Food scientists who are experienced in the sensory evaluation of dairy foods have an advantage over those who are only trained in chemical and/or microbial QC methods (Bodyfelt et al., 2008). A team trained in sensory evaluation will be better equipped at identifying defects and their sources. Sensory evaluation should be performed in a clean area, free from all odors (good and bad), and be in an environment with good lighting, no outside odors, and no noises or distractions. By controlling the external environment, panelists are allowed to focus solely on evaluating the product attributes.

While we know this scenario is idealistic, sensory analysis in a QC situation does not have to be an all or nothing approach and can be scaled based on individual facility capabilities. As a bare bones approach, panelists should be identified, and a schedule should be set up for sampling and tasting products in an appropriate environment. Panelists should be aware of what the target product should look and taste like and have an idea of what it means for a product to be out of acceptable range. Records of product evaluations should be maintained by filling out a standardized ballot, and action should be defined for when product does not meet sensory quality expectations. Below is a step-by-step outline for implementing a sensory QC program that utilizes DA principles.

1. *Identify your panelists*

   Depending on the facility, this may be one person or a small group. Ideally, multiple panelists will be used to balance each other in case of time off, job departures or illness. While sensory acuity varies among individuals (Meilgaard et al., 2025), every healthy individual is capable of being proficient at sensory evaluation with reasonable dedication and effort (Bodyfelt et al., 2008). When identifying panelists, desire to participate and dedication are more promising attributes than innate tasting ability.

2. *Set the panel up for success*

   At a minimum, tasting should be set up in an environment with minimal sensory distractions. Evaluation should take place away from the production floor and the loud noises and aromas that come with it. Panelists should not wear perfume, strong deodorant, scented lotions, etc., and there should be adequate overhead lighting. Implementing simple measures like these will help ensure evaluators are sensitive to the sensory characteristics of the products.

3. *Establish a tasting schedule*

   Sampling and tasting should be performed on all incoming raw materials and outgoing finished products. Additional tasting should occur at times that are deemed necessary based on the production process (before/after heating, culturing, etc.). Like any QC measure, more sampling times provides increased confidence about when a problem likely occurred, and helps you narrow down the amount of product that needs to be put on hold/discarded.

4. *Use a ballot for record keeping*

   Having a standardized ballot makes it easier to track changes over time and document any trends or deviations. This does not have to be a complex document and should in fact be easy to understand and implement. The ballot should be specific to the product being produced and reflect relevant product attributes/defects. Paper ballots can (still) be used or an online process can be used within a company or through the use of external commercial software.

5. *Identify a target product*

   The target is product (raw ingredients, finished products) with ideal sensory attributes. If everything is running perfectly, target product should be the output. It is important that panelists know what this product looks like so that they can identify when the product deviates from target and in what ways.

6. *Identify references*

   References provide examples of what evaluators should look for when evaluating the product. These differ from the target because they typically exemplify defects, not ideal attributes. Depending on the product being produced, references could include paint chips that represent a range of acceptable and unacceptable colors, pictures of common defects that have been observed in your product, or flavor/aroma references for off-flavors/aromas. These references should be available during training and evaluations.

7. *Adaptability and continuous improvement*

   As products are regularly tasted and defects arise, it is important to document them and add them as references, attributes, and/or defects on the ballot (also noting their cause, if known). It may seem like extra work, but customizing the ballot to your product and process will maximize its usefulness in your QC program.

8. *Periodic training/calibration*

   Recurring training/calibration using target, references, test samples, and the ballot ensures the panel is aligned on expectations. Research has found that while initial training is critical, the efficacy of retraining sessions can be maximized when only a select number of attributes are targeted (Carey et al., 2020). This may include a group evaluation session where the panel evaluates samples individually and then reviews results together, or a sample with a known defect may be included in the standard set of samples to verify whether panelists are correctly flagging defects when appropriate. If blinded defect samples are not correctly identified by panelists, it may be an indication that further training is needed.

9. *Following up on product failures*

   Inevitably, products will be evaluated which do not meet the sensory criteria. It is important to identify the course of action before this issue arises. Outlining which defects are unacceptable (product must be discarded) and which require other follow-up action will optimize the decision-making process.

10. *Do not extend the panel beyond its scope*

    The goal of the sensory QC panel is to identify attributes/defects and/or deviation from target/gold standard. Trained panelists should not be used to indicate liking or acceptance of the product (Drake, 2007). If additional sensory insights are needed, appropriate training, methodology, and protocol should be implemented (Lawless and Heymann, 2010).

This chapter highlighted the evolution of sensory methodologies. Quality judging identifies defects but has limited application. DA is a robust approach for identifying and quantifying attributes, utilizing a lexicon and standardized references for attribute identification and scaling. These characteristics from DA can be applied to the QC of dairy products by using DOD/DFC scaling and/or a cut down DA approach. The guidance and references provided in this chapter should be used to help establish sensory within a QC program, with room for adaptation based on individual production considerations.

References

Alvarez, V. B. (2023). Fluid Milk Products. In S. Clark, M. Drake, & K. Kaylegian (Eds.), The Sensory Evaluation of Dairy Products (3rd ed., pp. 79–142). Springer.

Bodyfelt, F. W., Drake, M. A., & Rankin, S. A. (2008). Developments in dairy foods sensory science and education: From student contests to impact on product quality. In International Dairy Journal (Vol. 18, Issue 7, pp. 729–734). https://doi.org/10.1016/j.idairyj.2008.03.011

Carey, N. R., Kent, D. J., Murphy, S. C., Boor, K. J., Wiedmann, M., & Martin, N. H. (2020). A practical training program for fluid milk defect judging should focus on initial training of panelists. Journal of Dairy Science, 103(7), 6716-6726. https://doi.org/10.3168/jds.2019-17336

Chambers, D. H., Allison, A. M. A., & Chambers IV, E. (2004). Training effects on performance of descriptive panelists. Journal of Sensory Studies, 19(6), 486–499. https://doi.org/10.1111/j.1745-459X.2004.082402.x

Clark, S. (2023). Dairy Products Evaluation Competitions. In S. Clark, M. Drake, & K. Kaylegian (Eds.), Sensory Evaluation Techniques (3rd ed., pp. 53–78). Springer.

Coggins, P. C., Schilling, M. W., Kumari, S., & Gerrard, P. D. (2008). Development of a sensory lexicon for conventional milk yogurt in the United States. Journal of Sensory Studies, 23(5), 671–687. https://doi.org/10.1111/j.1745-459X.2008.00179.x

Desai, N. T., Shepard, L., & Drake, M. A. (2013). Sensory properties and drivers of liking for Greek yogurts. Journal of Dairy Science, 96(12), 7454–7466. https://doi.org/10.3168/jds.2013-6973

Drake, M. (2023). Modern Sensory Practices. In S. Clark, M. Drake, & K. Kaylegian (Eds.), Sensory Evaluation Techniques (3rd ed., pp. 621–648). Springer.

Drake, M. A. (2007). Invited review: Sensory analysis of dairy foods. In Journal of Dairy Science (Vol. 90, Issue 11, pp. 4925–4937). Elsevier Inc. https://doi.org/10.3168/jds.2007-0332

Drake, M. A., & Civille, G. V. (2003). Flavor lexicons. Comprehensive Reviews in Food Science and Food Safety, 2(1), 33–40. https://doi.org/10.1111/j.1541-4337.2003.tb00013.x

Drake, M. A., Karagul-Yuceer, Y., Cadwallader, K. R., Civille, G. V., & Tong, P. S. (2003). Determination of the sensory attributes of dried milk powders and dairy ingredients. Journal of Sensory Studies, 18(3), 199–216. https://doi.org/10.1111/j.1745-459X.2003.tb00385.x

Drake, M. A., Mcingvale, S. C., Gerard, P. D., Cadwallader, K. R., & Civille, G. V. (2001). Development of a descriptive language for Cheddar cheese. Journal of Food Science, 66(9), 1422–1427. https://doi.org/10.1111/j.1365-2621.2001.tb15225.x

Drake, M. A., Watson, M. E., & Liu, Y. (2024). Downloaded from www.annualreviews.org. Guest (guest) IP: 152.1.118.51 On: Wed. Annual Review of Food Science and Technology. https://doi.org/10.1146/annurev-food-060721

Drake, M., & Clark, S. (2023). History of Sensory Analysis. In S. Clark, M. Drake, & K. Kaylegian (Eds.), The Sensory Evaluation of Dairy Products (3rd ed., pp. 1–8). Springer.

Drake, S. L., Lopetcharat, K., & Drake, M. A. (2009). Comparison of two methods to explore consumer preferences for cottage cheese. Journal of Dairy Science, 92(12), 5883–5897. https://doi.org/10.3168/jds.2009-2389

Drake, S. L., Yates, M. D., & Drake, M. A. (2010). Development of a flavor lexicon for processed and imitation cheeses. Journal of Sensory Studies, 25(5), 720–739. https://doi.org/10.1111/j.1745-459X.2010.00300.x

Ismail, B. P. (2017). Basic Principles of Chromatography. In S. S. Nielsen (Ed.), Food Analysis (5th ed., pp. 185–212). Springer.

Krause, A. J., Lopetcharat, K., & Drake, M. A. (2007). Identification of the characteristics that drive consumer liking of butter. Journal of Dairy Science, 90(5), 2091–2102. https://doi.org/10.3168/jds.2006-823

Lawless, H. T., & Heymann, H. (2010). Sensory Evaluation of Food. In H. T. Lawless & H. Heymann (Eds.), Sensory Evaluation of Food (2nd ed.). Springer.

Lawless, L. J. R., & Civille, G. V. (2013). Developing Lexicons: A review. In Journal of Sensory Studies (Vol. 28, Issue 4, pp. 270–281). https://doi.org/10.1111/joss.12050

Lee, A. P., Barbano, D. M., & Drake, M. A. (2017). The influence of ultra-pasteurization by indirect heating versus direct steam injection on skim and 2% fat milks. Journal of Dairy Science, 100(3), 1688–1701. https://doi.org/10.3168/jds.2016-11899

Liggett, R. E., Drake, M. A., & Delwiche, J. F. (2008). Impact of flavor attributes on consumer liking of Swiss cheese. Journal of Dairy Science, 91(2), 466–476. https://doi.org/10.3168/jds.2007-0527

McCarthy, K. S., Lopetcharat, K., & Drake, M. A. (2017). Milk fat threshold determination and the effect of milk fat content on consumer preference for fluid milk. Journal of Dairy Science, 100(3), 1702–1711. https://doi.org/10.3168/jds.2016-11417

Meilgaard, M. C., Civille, G. V., Carr, T. B., & Osdoba, K. E. (2025). Sensory Evaluation Techniques. In M. C. Meilgaard, G. V. Civille, T. B. Carr, & K. E. Osdoba (Eds.), Sensory Evaluation Techniques (6th ed.). CRC Press.

Peryam, D. R., and Pilgrim F. J. (1957). Hedonic scale method of measuring food preferences. Food Technol. 11:9-14.

Rétiveau, A., Chambers, D. H., & Esteve, E. (2005). Developing a lexicon for the flavor description of French cheeses. Food Quality and Preference, 16(6), 517–527. https://doi.org/10.1016/j.foodqual.2004.11.001

Schiano, A. N., Harwood, W. S., & Drake, M. A. (2017). A 100-Year Review: Sensory analysis of milk. Journal of Dairy Science, 100(12), 9966–9986. https://doi.org/10.3168/jds.2017-13031

Shakeel-ur-Rehman, Drake, M. A., & Farkye, N. Y. (2008). Differences between Cheddar cheese manufactured by the milled-curd and stirred-curd methods using different commercial starters. Journal of Dairy Science, 91(1), 76–84. https://doi.org/10.3168/jds.2007-0003

Shepard, L., Miracle, R. E., Leksrisompong, P., & Drake, M. A. (2013). Relating sensory and chemical properties of sour cream to consumer acceptance. Journal of Dairy Science, 96(9), 5435–5454. https://doi.org/10.3168/jds.2012-6317

Society of Sensory Professionals. (2021). Spectrum descriptive analysis. Accessed Oct. 31, 2024. https://www.sensorysociety.org/knowledge/sspwiki/Pages/Spectrum%20Descriptive%20analysis.aspx

Stone, H., and Sidel, J. L. (2004). Sensory Evaluation Practices. 3rd ed. Elsevier Academic Press, San Diego, CA.

Talavera, M., & Chambers, D. H. (2016). Flavor lexicon and characteristics of artisan goat cheese from the United States. Journal of Sensory Studies, 31(6), 492–506. https://doi.org/10.1111/joss.12239

Thompson, J. L., Drake, M. A., Lopetcharat, K., & Yates, M. D. (2004). Preference mapping of commercial chocolate milks. Journal of Food Science, 69(9). https://doi.org/10.1111/j.1365-2621.2004.tb09958.x

Table 1.1 Usefulness and limitations of common sensory methods

Method	Usefulness	Limitations
Product grading/quality judging	Training students, competitions, rapid quality analysis of multiple samples	Cannot run statistics, lacks scientific rigor, grades provide no indication of how products differ or how well liked they are by consumers
Difference tests	Determine if a difference exists between two or more samples, efficient as a preliminary test to see if follow-up methods are needed (DA, consumer tests)	Does not tell you what the difference is, how big it is, or if consumers like it
Descriptive analysis (DA)	Provides a product attribute profile, determines how samples are different (if at all), combine with analytical tests or consumer test results for further insights	Requires significant time and effort for panel training and maintenance, on its own does not indicate consumer liking
Degree of difference (DOD)/Difference from control (DFC)	Quantifies how different a sample is from control, rapid analysis	Requires some training, does not indicate consumer liking, or provide a profile of the product.
Consumer tests	Quantify degree of liking between samples, understand consumer preferences	Minimal insight into what consumers like about the products, expensive and necessitates recruitment of target consumers
Threshold tests	Determine the lowest concentration of a compound where a sensory difference is detected	Thresholds are matrix dependent, most useful to supplement analytical flavor data and DA

Identical to control					Extremely different from control
0	1	2	3	4	5

Figure 1.1 Unidirectional DOD scale

Extremely weaker than control				Identical to control				Extremely stronger than control
-4	-3	-2	-1	0	1	2	3	4

Figure 1.2 Bidirectional DOD scale

Table 1.2 Summary of existing sensory lexicons for dairy products and where they can be found

Lexicon	Source
Cheddar cheese (flavor)	Drake et al. (2001)
Dried milk powders and dairy ingredients	Drake et al. (2003)
Chocolate milk	Thompson et al. (2004)
Butter	Krause et al. (2007)
Sour cream	Shepard et al. (2013)
Cottage cheese	Drake et al. (2009)
Processed and imitation cheeses	Drake et al. (2010)
Yogurt	Coggins et al. (2008); Desai et al., 2013
French cheeses (flavor)	Rétiveau et al. (2005)
Goat cheese (flavor)	Talavera and Chambers (2016)
Swiss cheese	Liggett et al. (2008)
Fluid milk	McCarthy et al. (2017)

Example 1: Understanding the effect of processing on sensory properties of milk

Process changes or ingredient additions are often used to improve some aspects of the product. However, making a change to optimize a product in one area may have a negative impact on another area. Increasing the thermal heat treatment of fluid milk, for example, improves the shelf-life, but can have a negative impact on the sensory properties. Research by Lee et al. (2017) compared HTST milks to UP milks and found that consumers preferred HTST milks. DA provided objective sensory information about these milks suggesting why that may be the case–UP milks had higher cooked and sulfur/eggy flavors.

Example 2: Determining drivers of liking for Greek yogurts

Drivers of liking refer to the product attributes that drive consumer liking of a product category. This is determined by concurrently documenting the trained panel profile of products representative of the product category and evaluating consumer acceptability. DA and a consumer test were performed with commercial Greek yogurts. Firm, dense texture, moderate sweet aromatic, milkfat, dairy sour flavors, and moderate sour taste are drivers of liking (Desai et al., 2013).

Example 3: Screening milks to identify source of off-flavor

Consumers complained of spoiled whole milk with off-flavors. In response, trained panelists screened 50 individual milk cartons to confirm or deny the presence of off-flavors (defect screening). The panel was also able to characterize the source of off flavor as coming from milk additives rather than microbial spoilage as originally suspected. This allowed for a targeted approach to address consumer complaints.

Example 4: Identifying source of off-flavor in creamers using DOD

A QC panel flagged creamers out of spec with a high DOD score. Cut down DA was implemented, scoring key flavor and texture attributes. The DA panel identified differences in vanilla aromatics between the flagged sample and the control. This result prompted evaluation of the vanilla flavorings (coming from different suppliers). Off-flavor from one of the vanilla samples matched the off flavors in the flagged samples. This process confirmed that the source of off-flavor was from one of the suppliers and follow-up action was addressed with that supplier.

Figure 1.3 Application of DA with dairy products

CHAPTER 2:

Extrinsic Attributes that Drive Oat Milk Purchase

Extrinsic Attributes that Drive Oat Milk Purchase

D. Rovai, S. Gupta, M. Watson, M. A. Drake

Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, North Carolina, USA

Choice of journal/topic: Journal of Food Science, submitted

Abstract

Oat milk is a popular nondairy milk with high sensory appeal. The objective of this study was to understand trial and repeat purchase motivations, and the extrinsic attributes that influence consumer preference for oat milks. An online survey (n=415) and focus groups (n=24) were conducted with oat milk consumers. Survey data were evaluated by univariate and multivariate analyses. Key themes were identified from focus groups. Oat milk was perceived as sustainable and heart healthy ($p<0.05$). Consumers were motivated to try oat milk because it fit their dietary restrictions and was trendy. Repeat consumption was motivated by unique flavor and texture. Two consumer clusters were identified. Both clusters had the same ideal oat milk attributes but differed in price sensitivity. The ideal oat milk was original flavor with claims of excellent source of calcium, vitamins A & D, and free from artificial flavors and colors ($p<0.05$).

Practical Application

Consumers increasingly seek out alternatives to dairy products. In response, plant-based alternatives have grown in popularity. This research examined consumer perception and preferences for oat milk—a fast growing nondairy milk alternative. These findings provide insights into why consumers tried oat milk for the first time, why they continue to purchase it, and how they value tradeoffs for purchase decisions. Understanding these motivations and behaviors will help guide product development and marketing for oat milks. Insights from this research may also aid in the development of other plant-based alternative products.

Key Words: oat milk, plant-based, consumer

2.1 Introduction

Dairy dominates the milk market, responsible for 83% of milk sales in 2024 (Mills, 2024). However, dairy production has raised concerns about animal welfare, human health, and environmental sustainability (Clay et al., 2020). These concerns, among others including perceived lactose intolerance or allergies, have contributed to the rise of non-dairy milks, despite them being more expensive than dairy milk (Ramsing et al., 2013, Moss et al., 2022). Plant-based milks can be made from a variety of substrates including nuts (almond, cashew, coconut), seeds (flax, sunflower, hemp), legumes (soy, pea), cereals (rice, oats), and pseudocereals (chia) (Sethi et al., 2016; Silva et al., 2022). Plant-based milks are perceived by consumers as more sustainable than dairy milks (Schiano et al., 2020) because they require less resources to produce and contribute fewer greenhouse gas emissions (Riofrio & Baykara, 2022). The general process for plant-based milk production involves soaking the substrate, grinding to break down the tissue, separating the oil phase from the aqueous phase, adding enzymes and making pH adjustments, blanching to inactivate enzymes, thermal processing, homogenization to stabilize the oil and aqueous phases, and a final formulation stage to add in functional ingredients like flavors, colors, emulsifiers, and/or stabilizers (McClements et al., 2019). The resulting product is a stabilized oil in water emulsion, similar to milk in texture, produced from plant sources.

A survey of healthcare professionals found they believe consumers are confused about the nutritional differences between dairy and plant-based alternatives (Clark et al., 2022). Numerous studies have compared the nutritional profile of dairy milk to various plant-based milks and concluded that plant-based milks have inferior nutritional composition (less protein, vitamins, minerals) compared to cow’s milk (Sethi et al., 2016; Chalupa-Krebzdak et al., 2018; Yu et al., 2023; Silva et al., 2024). However, studies have also suggested that plant-based milks can reduce risk of cardiovascular and GI disease (Chalupa-Krebzdak et al., 2018) and can deliver probiotics to consumers to boost health (Rasika et al., 2021). Additionally, despite differences in nutrient content, consumers tend to associate plant-based milks with health benefits (Moss et al., 2022). Nutrient fortification or blending of plant-based milk types have also been proposed as solutions to improve the nutritional quality of plant-based milks to be more similar to dairy (Sethi et al., 2016; Silva et al., 2024).

Perhaps the most important factor in the discussion between plant-based and dairy milks is the sensory quality. Research has cited the poor sensory qualities of plant-based milks compared to dairy as a key factor that has limited growth (Alcorta et al., 2021; Jaeger & Giacalone, 2021; Giacalone et al., 2022). However, while consumers overall prefer the sensory properties of dairy, segments of consumers exist and preferences for plant-based milk are driven by sensory, emotional, and situational use perceptions (Cardello et al., 2022). In addition, the sensory characteristics of plant-based milks vary greatly depending on the raw material/substrate they are made from (Mäkinen et al., 2015, Vaikma et al., 2021). Among plant-based milks, almond milk has consistently had the highest sales over the past five years (Mills, 2024). However, after first sold in the U.S. in 2016, sales of oat milk have skyrocketed, surpassing soy milks as the number two plant-based milk sold in the U.S. by 2020 (Ramsing et al., 2023). In research by Martínez-Padilla et al. (2023), oat milk was the most frequently consumed plant-based milk by Danish consumers. Research comparing almond, pea, coconut, soy, cashew, and oat milks also found that consumers rated oat milk the highest in overall liking (Moss et al., 2022). Another study that found oat milk samples scored consistently well compared to other plant-based milks (Jaeger et al., 2024). In addition to sensory attributes, consumer food related behaviors are influenced by other factors such as price, convenience, production technology, branding, personal health, and societal issues (Jaeger, 2006). This study used a survey and focus groups to understand the extrinsic attributes that influence consumer preferences for oat milks and to explore consumer motivations for trial and repeat purchase of oat milk.

2.2 Methods

2.2.1 Online Survey

2.2.1.1 Recruitment and Screening

An online survey was developed using Lighthouse Studio (Sawtooth Software version 9.15.0, Orem, UT). The survey was uploaded, and participants were recruited from a database of over 12,000 individuals, maintained by the Sensory Service Center at North Carolina State University. Participants first answered demographic questions and were screened out to include only panelists that were > 18 years old and consumed oat milk at least occasionally. Additional participant screening based on data reliability was conducted and is discussed further in the statistical analysis section. The final number of participants whose data was used in this study was n=415.

2.2.1.2 Assessing Consumer Perception of Different Types of Milk

A series of sliding scales were used to assess consumer perceptions toward different types of milks (dairy, oat, almond, soy, coconut). For each type of milk, participants indicated on a sliding scale where they believed each of the milks ranked relative to each other. Milks were compared for attributes using the following scale anchors (0–100): not at all sustainable–very sustainable, not at all nutrient rich (vitamins & minerals)–nutrient rich (vitamins & minerals), hardly any sugar–lots of sugar, not at all heart healthy–heart healthy, does not taste great–tastes great, bad texture–great texture, not at all expensive–very expensive, not at all clean label–very clean label, long ingredient list–short ingredient list, made with unfamiliar ingredients–made with familiar ingredients, lots of allergens–allergen free, hardly any protein–lots of protein.

2.2.1.3 Assessing Oat Milk Consumption Behavior

After probing about different types of milks, participants were informed that the remainder of the survey would focus specifically on oat milk. Oat milk consumption and purchase habits were asked via single select, CATA, and free response questions. These questions asked participants how often they consumed oat milk, how they consumed it (drinking on its own, adding to cereal/oatmeal, adding to smoothie, adding to coffee/tea, using as an ingredient in a recipe) and, if applicable, how often they used oat milk for each of these occasions (multiple times per day, once per day, a few times per week, once per week, a few times per month, once per month, less than once per month). Participants were also asked by CATA why they used oat milk over other milk options (better taste, better texture, better color, more sustainable, healthier, availability, easier to digest, other).

2.2.1.4 Assessing the Attributes of an Ideal Oat Milk

The next portion of the survey aimed to understand the attributes of an ideal oat milk. This objective was assessed via an Adaptive Choice-Based Conjoint (ACBC) exercise (Sawtooth Software, 2014). The factors and their respective levels are summarized in Table 2.1. These factors and levels were chosen based on a market survey of oat milks available for purchase locally and online. For price, a base price of $4.35 per half gallon (64 oz) was determined and additional price increases are noted in Table 2.1. During the ACBC activity, the total summed prices (base price + additions) varied +/- 30% (Harwood & Drake, 2018). In the Adaptive version of an ACBC, participants first complete a build your own (BYO) activity where they select the level within each factor that they find most appealing (taking into account the price which changes on the screen based on the attributes chosen). This process is followed by a screening activity where participants are shown a series of product options and asked to indicate whether the product combination is a possibility or would not work for them. Based on their responses to these screening questions, participants may be prompted whether some of these levels are must haves or dealbreakers. Based on the results from the BYO and these screener exercises, the ACBC algorithm results in a good estimate of each participant's most important and least important attributes. The final tournament portion of the ACBC then aims to determine the utility of each factor and level by presenting panelists with a series of oat milk options and forcing participants to make tradeoffs. This process is useful because, in addition to knowing the ideal oat milk (indicated by the BYO), we have additional information about which of the factors and levels are driving preference and which aren't as important to consumers. Additionally, to validate the insights from the ACBC exercise in terms of factor importance, a constant sum exercise followed the ACBC exercise asking participants to allocate a total of 100 chips across each of the attributes, assigning more chips to the attributes of oat milk that were most important when purchasing oat milks: brand, flavor, nutritional profile, other attributes (creamy, sugar free, low fat, etc.), label claim (good source, organic, gluten free, etc.), and price.

Kano analysis questions were also performed to help validate the results from other exercises in the study. Participants were first shown the functional version of attributes and asked to indicate one of the following: I will dislike this oat milk, I can live with this oat milk, I do not care, I expect oat milk to have this, I will like an oat milk more if it has this. Next, participants used the same anchors in response to the dysfunctional version of the same attributes (ex. Functional - an oat milk that tastes great, Dysfunctional - an oat milk that does NOT taste great). Based on the combination of functional and dysfunctional responses, each attribute was classified as either a must have, performance, attractive, indifferent, reverse, or questionable attribute (Sauerwein et al., 1996; Sharif Ullah & Tamaki, 2010).

2.2.1.5 Assessing Ideal Oat Milk Label Claims

While label claims were included on the ACBC, they were limited in the number that could be included to balance the conjoint design. To further explore consumer preferences for different label claims, a Maximum Difference Scaling (MaxDiff) exercise was used which included 26 label claims found on commercial oat milk labels. These labels were presented to consumers five at a time and consumers indicated the label was the most important and least important when purchasing oat milk. A series of these tasks were completed such that each label was presented to participants five times.

2.2.2 Focus Groups

Four moderator-guided focus groups (two hours each) were conducted with a subset of oat milk consumers who completed the online survey (n=24). The focus group discussions were conducted after the survey concluded (within 3 weeks) and were used as a tool to further explore the insights from the survey and supplement them with qualitative data to capture consumer sentiment and emotions that would otherwise not be captured through quantitative survey data alone. The discussions were moderated to first explore trial motivations for why oat milk consumers looked to try oat milk for the first time–what were they drinking before, why did they consider switching, what influenced them to try oat milk, what was the hesitation? The discussions were then moderated to understand repeat motivations for oat milk consumption–after you tried oat milk for the first time, what made you decide to keep purchasing it? Embedded in these discussions were the comparisons between other types of milk (dairy and other plant-based alternatives) and the occasions for when different types of milk work well and do not work well to understand the applications and limitations of oat milk and explore opportunities for improvement within the oat milk product space.

2.2.3 Statistical Analysis

For each sliding scale, the mean scores for milks were compared to each other using ANOVA with Fisher's least significant difference for means separation (α=0.05). A principal component analysis (PCA) biplot was also constructed using Pearson correlations to visualize differences in sliding scale ratings among milks. Importance scores for the ACBC factors were determined via hierarchical Bayesian (HB) estimation (Sawtooth Software, 2014). The importance scores were compared via ANOVA with Fisher's least significant difference (α=0.05). Utility scores for the levels within each factor were also determined using HB estimation and the utility scores were compared via ANOVA with Fisher's least significant difference (α=0.05) for levels *within* a factor (not *between* factors). Similarly, utility scores for the label claims MaxDiff exercise were determined via hierarchical Bayesian (HB) estimation (Orme, 2009). The utility scores for each label claim were compared via ANOVA with Fisher's least significant difference (α=0.05). Segmentation of ACBC importance scores was performed using AHC with K-means clustering.

Kano results were evaluated using the model proposed by Kano (1984). For each attribute, panelists fell into one of the following buckets: must have, performance, attractive, indifferent, reverse, questionable. "Must have" attributes are expected by the consumer. "Performance" attributes are one dimensional where consumers like having these theme and dislike not having them. "Attractive" attributes are not expected but are liked by respondents when they are present. "Indifferent" attributes are self-explanatory, while "reverse" attributes are those where consumers like it when the feature is absent and dislike it when it is present.

“Questionable” attributes indicate conflicting responses from consumers. The percentage of participants who were assigned to each of these categories was visualized using correspondence analysis symmetric row plots where attributes closer to categories are more correlated with those categories.

To ensure reliability of the survey data collected, a few methods were implemented. Two attention check questions were included in the survey which asked participants to select a specific response, and individuals who did not read the directions and selected the wrong answer were disqualified from the survey. In addition, to identify individuals who were just clicking through the MaxDiff exercises without putting meaningful thought into their responses, data (n = 300) were randomly generated through Sawtooth for the MaxDiff exercise. This randomly generated data was then run using hierarchical Bayesian (HB) estimation and the 95th percentile root likelihood (RLH) score for these responses was used as a cutoff value. Data from participants with RLH scores at or below this value were removed from the data set (Orme, 2019). A total of 455 participants completed the survey in its entirety, however, based on the criteria above, data from 40 participants were discarded. Data from n=415 participants were used for all survey analyses.

For each focus group, two note takers were present listening to the discussion and writing down key themes, insights, and quotes. Following each focus group and after the conclusion of all focus groups, notes were compiled, and key themes and insights were discussed between note takers and the moderator. Themes that were consistent across multiple focus groups were reported. Quotes were selected that expressed sentiment in the consumers’ voice that was consistent with the themes and insights identified.

2.3 Results and Discussion

2.3.1 Demographics and Milk Consumption

The median survey completion time was 32.5 minutes. Survey participants (n=415) were majority female (75.9%), white/Caucasian (66.7%), with an average age of 36.1 ± 14.2 years, who resided in the state where the affiliated university is located–North Carolina (88.2%), and had 4 or more years of college (77.3%), employed full-time (54.0%), with a household size of 1 person (25.8%) or 2 people (36.4%) and no children (77.8%). All participants reported consuming oat milk at least occasionally. In addition, only one consumer claimed to drink exclusively oat milk, while all other consumers indicated that they consume some combination of dairy milk and other plant-based milks in addition to oat milk. Only 6.3% of consumers claimed to never consume dairy milk, suggesting that oat milk is competing for market share against both dairy milk and plant-based milks.

2.3.2 Consumer Perception of Different Types of Milks

Comparison of different milks via sliding scales (Table 2.2) illustrated distinct differences in consumer perception of milks. Dairy milk was perceived as the cheapest and the most nutritious, with the best taste and texture, a clean label with a short ingredient list, familiar ingredients, and the highest protein content (Table 2.2). These consumer perceptions are consistent with research which has established that bovine milk had greater nutritional value and higher sensory scores than almond milk, oat milk, and soy milk (Vashisht et al., 2024). However, dairy milk was perceived as the least sustainable, least heart healthy, and having the most allergens (Figure 2.1). Compared to dairy, plant-based milks in general were perceived as more sustainable, heart healthy, and having fewer allergens, but are less nutritious, more expensive, less clean label, made with a longer ingredient list, less familiar, and containing less protein.

Among plant-based milks, consumers perceive oat milk as the most sustainable and having the least amount of allergens (Table 2.2). Literature supports the lack of allergenicity associated with oat milks (Bocker & Silva 2022). A life cycle assessment in Ecuador found that oat milk production emits less greenhouse gas emissions than bovine milk and uses less water than almond milk (Riofrio & Baykara, 2022), supporting consumer perception that oat milk is sustainable. In terms of taste and texture, oat milk and almond milk were the highest rated plant-based milks (Table 2.2). Preferences for oat milk and almond milk reflect sales data showing that almond milk is the highest selling plant-based milk, followed by oat milk (Mills, 2024). Oat milk, alongside coconut milk, was also rated as the most expensive (Table 2.2).

2.3.3 Trial Motivations

While the sliding scales provided insight into the perception of different types of milk among oat milk consumers, it does not fully describe the factors that influenced their decision to try oat milk in the first place. During focus groups, a few key themes emerged when consumers talked about the factors that influenced their decision to try oat milk for the first time (Figure 2.2). The first theme was that oat milk fit within their dietary restrictions. This is a prerequisite for trying any new food product, and the lack of allergens associated with oat milk (Figure 2.2) made it accessible to most consumers. Sustainability is another important aspect for many consumers, who felt that they wanted to try an alternative to dairy milk that was more environmentally friendly. Consumers have a positive impression of the sustainability of oat milks and a negative perception of dairy milks (Table 2.2), which encourages trial of oat milk based on messaging that oat milk is healthier for the environment. Another theme that emerged was that consumers like having a variety of milks in the fridge to choose from, and they were interested in trying a new product because it is trendy and they had heard good reviews (Figure

2.2). Many focus group participants described their first experience with oat milk in a coffee from Starbucks based on a recommendation or because they heard about it through advertisements. While many consumers were eager to try a new product, others were hesitant to try oat milk for the first time because they worried about wasting money on an expensive product that they may not enjoy. Despite these hesitations, all consumers in this study overcame their initial hesitations to try oat milk for the first time because it fit their dietary restrictions, added variety to their diet, was trendy, and sustainable.

2.3.4 Repeat Purchase Motivations

Encouraging consumers to try a product is vital to product success, but consumers must also enjoy the product enough to want to continue to purchase it. Consumers have a positive perception of oat milk, ranking it highly on taste and texture compared to other plant milks (Table 2.2). Consumers in focus groups described the flavor of oat milk as having a unique oaty, nutty flavor and a thick and creamy texture (Figure 2.2). Consumers also explained that they liked the fact that oat milk is different from dairy milk. When consumers crave oat milk, they want the unique flavor and sweetness that it adds. In addition, most consumers do not consume oat milk exclusively, but like rotating through different types of milks and enjoy adding oat milk to their line-up. Additionally, while plant-based milks in general (and oat milk especially) are perceived as more expensive than dairy milk (Table 2.2), many consumers justified the higher price of non-dairy milks because they believe that these milks last longer in the fridge than dairy milk (Figure 2.2). The shelf life of milks (plant or dairy) is primarily determined by the quality of the raw materials and the heat treatment applied to the milks (Boor, 2001, Jo et al., 2018). When commercial milks are processed, they typically undergo either pasteurization (minimum 72°C for 15 s) or ultra pasteurization (minimum 138°C for 2 s) (21CFR131.3). Because ultra pasteurization is a higher heat treatment than pasteurization, a product that undergoes ultra pasteurization will have a longer shelf life than the same product that only undergoes pasteurization. In the U.S. most bovine milk undergoes pasteurization, while plant-based milks undergo ultra pasteurization. This may explain why consumers perceive plant-based milks to have a longer shelf life than dairy. While there are many brands of dairy milk available to buy that also undergo ultra pasteurization, ultra pasteurized dairy milk is generally more expensive than pasteurized dairy milk and research has found that consumers tend to not read the labels on dairy products and do not understand the different methods of dairy processing (Schiano & Drake, 2021). Finally, consumers explained that buying oat milk makes them feel better about themselves because they are doing good for the environment and are being healthier (Figure 2.2).

During the focus groups, a portion of the discussion focused on the application and limitations of oat milk compared to other types of milk. Consumers explained that oat milk works well for certain applications where the sweetness and flavor of oat is appreciated such as in coffee, smoothies, cereal, oatmeal, or drinking on its own (Figure 2.2). While adding to coffee/tea, adding to smoothie, and adding to cereal/oatmeal are the most common applications for oat milk consumers, a large portion also indicated that they occasionally use it as an ingredient in a recipe (Table 2.3). However, consumers in focus groups complained that there are many recipes and applications where they would like oat milk to work, but it simply falls short (Figure 2.2). In general, consumers observed that oat milk did not work well for cooked applications because it breaks down during cooking. This observation is supported by literature which found that plant milks have lower stability to heat compared to dairy milk, although oat milk has higher stability than almond or soy milk (Vashisht et al., 2024). In addition, the flavor of oat milk does not work well in most savory dishes. Many consumers attempted to use oat milk as a dairy alternative in cream sauces and macaroni and cheese but were disappointed by some combination of breakdown during cooking and a poor flavor complement to the dish (Figure 2.2). However, consumer desire for an oat-based product that does perform well and works in these savory applications suggests that there is room for innovation to better meet consumer desires.

2.3.5 Ideal Oat Milk Attributes

Table 2.1 shows selections from the BYO portion of the ACBC exercise reported for the total population and by cluster. Consumer clusters were determined by segmentation of the ACBC importance scores (Figure 2.3) and revealed two groups: 1) a Price Driven Cluster (n=272) and 2) a Balanced Cluster (n=143). Both clusters assigned the highest importance to price, followed by flavor, type, label claim, free-from claim, and brand (Figure 2.3). Similar importance was assigned to these attributes during the chip allocation exercise where price was the most important factor, followed by flavor, nutritional profile, product attributes such as texture, then followed by brand and label claim (Table 2.4). Previous research has confirmed that brand is minimally important to oat milk consumers (McCarthy et al., 2017; Halme et al., 2023). Comparing the two clusters, the Price Driven Cluster assigned very high importance to price compared to the other attributes, while the difference in importance between price and the other attributes was less drastic for the Balanced Cluster. This result is reflected in the BYO results where the average price was \$0.69 lower for the Price Driven Cluster compared to the Balanced Cluster (Table 2.1). Price was still important to the Balanced Cluster (Figure 2.3), but this group of consumers was willing to pay more to ensure they obtained product attributes that were important to them (Table 2.1).

While the importance scores compare the relative importance of attributes, the utility scores show the relative utility of levels within an attribute. Figure 2.4 shows the ACBC utility scores for the total population. For the flavor attribute, the highest utility was assigned to original/plain flavor, followed by vanilla. For type, consumers preferred original or extra creamy oat milks. Label claims boasting a good/excellent source of vitamins, minerals, as well as protein received the highest utility, and free from artificial flavors and colors received the highest utility amongst the free from claims. Another research study used ACBC with label claims as an attribute for non-dairy milk and found that "fortified with vitamin A/D, calcium" was the most preferred label claim (McCarthy et al., 2017). Label claims will be discussed further and in more detail in the following section. For brand, store brand received the lowest utility, while Oatly, Chobani, and Planet oat received the highest utility. High utility for these brands corresponds with oat milk market share reported by Mintel with Planet Oat having 33.6% of market share, Oatly 23.4%, Chobani 20.7%, and private label only having 8.8% (Mills, 2024). Comparing the utility scores for the total population to the Price Driven Cluster (Figure 2.5) and the Balanced Cluster (Figure 2.6), the utility scores follow a consistent pattern. Consumers from both clusters assigned the highest utility to the same top levels for the flavor, type, label claim, and free-from claim attributes. Only slight differences were observed in the brands with highest utility where the Price Driven Cluster assigned the highest utility to Planet Oat, Chobani, Oatly in that order (Figure 2.5), and the Balanced Cluster assigned highest utility to Oatly, Chobani, Califia Farms in that order (Figure 2.6). Thus, while these two clusters differed slightly in how much they were willing to pay for their ideal oat milk, the ideal oat milk for all consumers surveyed was an original/plain flavored oat milk with an original or extra creamy texture with good/excellent source of calcium, vitamins A and D, and protein, free from artificial flavors and colors, and is not a store brand.

The ideal oat milk was further explored using results from the Kano exercise. Kano analysis categorizes product attributes based on satisfaction/dissatisfaction when the attributes are present and absent from a product. Must have attributes are basic requirements that consumers expect to be present. Performance attributes are features that, when present, increase satisfaction, while reverse attributes are the opposite–reverse attributes decrease satisfaction. Attractive attributes are not expected by consumers, but they are liked when they are present. Consumers feel neutral about indifferent attributes, and questionable attributes mean that participants gave conflicting or illogical responses for these attributes. The correspondence analysis symmetric row plot summarizes the categorization of various oat milk attributes (Figure 2.7). Looking at this plot, positive associations with attributes increase as you move left along the x-axis, with reverse attributes on the right and must have attributes on the left. Great taste and great texture are positive attributes that consumers expect and desire from oat milk. This is expected and is consistent with research by McCarthy et al. (2017) which found that flavor/taste was a prerequisite to purchase for fluid milk and milk alternative users. Additional desirable attributes included “creamy texture”, “vitamins and minerals”, “recognizable ingredients”, “good source of protein”, “preferred flavor”, “sustainably made”, “found in refrigerated section”, “short ingredient list”, “same nutrition as dairy milk”, “same protein and calcium content as milk”. On the other hand, consumers do not like oat milks that are sweetened with added sugars, whether they come from sugar or non-nutritive sweeteners. Previous research has examined consumer preferences for sweeteners in yogurt and found that aspirational desires do not always align with taste preferences (Crown et al., 2024). Similarly, while consumers in the current study indicate that they do not want sweetened oat milk, previous research suggests that increased sweetness increases consumer satisfaction in plant-based milks (Sethi et al., 2016; Moss et al., 2022; Gaider et al., 2024; Jaeger et al., 2024).

2.3.6 Desirable Label Claims

Some label claims were included in the ACBC exercise, but a more comprehensive list of label claims was evaluated using a MaxDiff exercise (Figure 2.8). The utility scores assigned to label claims via MaxDiff suggests that consumers were more influenced by label claims about the nutritional value of the oat milk, rather than allergens or farming practices. This result is observed by the highest utility assigned to "excellent source of calcium, vitamins A and D", "excellent source of protein", "good source of calcium, vitamins A and D", and "good source of protein". Low utility scores were assigned to allergen and free-from claims such as "nut free", "gluten free", "soy free", "vegan", "cholesterol free". In addition, "organic" and "non-GMO" received low utility scores (Figure 2.8). Referring back to the sliding scales, consumers perceived oat milk to contain few allergens and be sustainable but contained low amounts of protein and less nutritious than dairy milk or almond milk (Table 2.2, Figure 2.1). These perceptions may play into the desirable label claims, as consumers want the claims to tell them something they don't know to make them like the product more–if oat milk consumers already know that oat milk doesn't contain nuts, they don't need the package to say so. While free from claims generally received low utility, "free from artificial flavors and colors" was the exception and received relatively high utility. Previous literature has also suggested that consumers have a negative perception of any ingredients termed "artificial". Research by Schiano et al. (2021) found that dried dairy ingredients were perceived as healthier and more natural when they were attached to the claim "no artificial sweeteners, flavors, or colors". Free from artificial ingredients also fits within the definition of clean label (Osborne, 2015; Cao and Miao, 2023), and products with a clean label are perceived by consumers as healthier, more socially responsible, and having greater sensory appeal (Cao and Miao, 2023).

To claim an excellent or good source of a nutrient, the product must meet certain nutritional thresholds. The “good source” claim means the product must contribute 10-19% of the RDI or the DRV per reference amount customarily consumed (21CFR101.54c). For the “excellent source” claim, the product must contribute at least 20% of the RDI or the DRV per reference amount customarily consumed (21CFR101.54b). According to the USDA ARS, unsweetened commercial oat milk contains 1.9 g protein, 355.2 mg calcium, 204 mcg vitamin A, and 4.1 mcg vitamin D per 240 mL (USDA ARS, 2024). Based on these values, oat milk can make the claims for “excellent source” of calcium (27.3%), vitamin A (22.7%), and vitamin D (20.5%) (FDA, 2024). However, oat milks have too low a protein content (3.8%) to make any claims without protein fortification (FDA, 2024). Previous research examined consumer acceptability of oat milks fortified with B-glucan and protein and found that consumers conceptually like the idea of more protein, but they did not like the sensory experience of products with higher protein content (Alsado et al., 2023). Thus, trying to reformulate oat milk to be able to include a protein claim may result in lower product acceptability.

2.3.7 Limitations and Future Work

This research explored consumer perceptions about oat milk attributes through an online survey, but did not include tasting. Future research should explore the drivers of liking for oat milk through consumer tasting and trained panel profiling. Some literature exists which suggests that plant specific flavors in milk evoke negative emotions and results in lower consumer acceptance (Gaider et al., 2024). In addition, sensory differences have been observed between plant milks with different raw materials and across brands (Gaider et al., 2024), and research has shown that particle size can affect the perceived whiteness. Processing and packaging differences may contribute to aroma differences in oat milks (McCarron et al., 2024). Many factors influence the physicochemical and sensory properties of oat milk, such as cultivar, the type of oats, the composition, and the processing steps used (Paul et al., 2020; Patra et al., 2022; Cui et al., 2023; Zhou et al., 2023). Future work should continue to explore the impact of raw ingredients and processing parameters on the sensory quality and acceptability of oat milk. Understanding consumer preferences for both extrinsic and intrinsic oat milk attributes will help provide guidance to industry.

2.4 Conclusion

Consumers perceive oat milk to be one of the most appealing plant-based milks with a good taste and texture. Oat milk is also perceived as sustainable and is not associated with allergens. However, consumers perceive oat milk as expensive and low in protein, vitamins, and minerals. Consumers were motivated to first try oat milk because it fit their dietary restrictions, was trendy, sustainable, and something new to try. When consumers use oat milk, they find it to be the perfect addition to coffee, smoothies, cereal, oatmeal, or drinking on its own because of the subtle sweetness, the oaty/nutty flavor, and the creamy texture it provides. However, oat milk does not work in savory dishes or cooked applications. The attributes of oat milk that keeps consumers buying it–despite the high price–is the unique flavor and texture, the variety it adds, and the fact that it lasts longer in the fridge than dairy milk. Consumers do not want oat milks that contain added sweeteners, whether they be sweetened with sugar or non-nutritive sweeteners. When shopping for oat milk, consumers are most influenced by the price, followed by the flavor, type, label claims, and brand. Consumers are homogeneous in their preferences for oat milks, though one segment of consumers will cede their preferences in favor of the lowest price. The ideal oat milk is original/plain flavor, original or extra creamy in texture, and from a national brand rather than store brand. The ideal oat milk also boasts label claims touting the nutritional value of the product (high in vitamins, minerals, protein), rather than claims about the lack of allergenicity. These findings provide context about consumer motivations for trying and purchasing oat milk and provide insights into desirable attributes for oat milks.

Acknowledgements

Funding provided in part by Novonesis (Franklinton, NC). The use of tradenames does not imply endorsement nor lack of endorsement by those not mentioned.

References

Alcorta, A., Porta, A., Tárrega, A., Alvarez, M. D., & Pilar Vaquero, M. (2021). Foods for plant-based diets: Challenges and innovations. In *Foods* (Vol. 10, Issue 2). MDPI AG. https://doi.org/10.3390/foods10020293

Alsado, C., Lopez-Aldana, L., Chen, L., & Wismer, W. (2023). Consumer Perception and Sensory Drivers of Liking of Fortified Oat Milks. *Foods*, *12*(22). https://doi.org/10.3390/foods12224097

Bocker, R., & Silva, E. K. (2022). Innovative technologies for manufacturing plant-based non-dairy alternative milk and their impact on nutritional, sensory and safety aspects. In *Future Foods* (Vol. 5). Elsevier B.V. https://doi.org/10.1016/j.fufo.2021.100098

Boor, K. J. (2001). ADSA foundation scholar award fluid dairy product quality and safety: Looking to the future. *Journal of Dairy Science*, *84*(1), 1–11. https://doi.org/10.3168/jds.S0022-0302(01)74445-1

Cao, Y., & Miao, L. (2023). Consumer perception of clean food labels. *British Food Journal*, *125*(2), 433–448. https://doi.org/10.1108/BFJ-03-2021-0246

Cardello, A. V., Llobell, F., Giacalone, D., Roigard, C. M., & Jaeger, S. R. (2022). Plant-based alternatives vs dairy milk: Consumer segments and their sensory, emotional, cognitive and situational use responses to tasted products. *Food Quality and Preference*, *100*. https://doi.org/10.1016/j.foodqual.2022.104599

Chalupa-Krebzdak, S., Long, C. J., & Bohrer, B. M. (2018). Nutrient density and nutritional value of milk and plant-based milk alternatives. In *International Dairy Journal* (Vol. 87, pp. 84–92). Elsevier Ltd. https://doi.org/10.1016/j.idairyj.2018.07.018

Clark, B. E., Pope, L., & Belarmino, E. H. (2022). Perspectives from healthcare professionals on the nutritional adequacy of plant-based dairy alternatives: results of a mixed methods inquiry. *BMC Nutrition*, *8*(1). https://doi.org/10.1186/s40795-022-00542-7

Clay, N., Garnett, T., & Lorimer, J. (2020). Dairy intensification: Drivers, impacts and alternatives. In *Ambio* (Vol. 49, Issue 1, pp. 35–48). Springer. https://doi.org/10.1007/s13280-019-01177-y

Crown, E., D. Rovai, C. M. Racette, D. M. Barbano, and M. A. Drake. 2024. Consumer perception of sweeteners in yogurt. J Dairy Sci. In press.

Cui, L., Jia, Q., Zhao, J., Hou, D., & Zhou, S. (2023). A comprehensive review on oat milk: from oat nutrients and phytochemicals to its processing technologies, product features, and potential applications. In *Food and Function* (Vol. 14, Issue 13, pp. 5858–5869). Royal Society of Chemistry. https://doi.org/10.1039/d3fo00893b

FDA. (2024). Daily value on the nutrition and supplement facts labels. Accessed 9/23/24. https://www.fda.gov/food/nutrition-facts-label/daily-value-nutrition-and-supplement-facts-labels

Gaider, M., & Majchrzak, D. (2024). Dynamic sensory evaluation of selected sensory attributes and food-related emotions during consumption of plant-based milk alternatives applying temporal dominance of sensations and temporal dominance of emotions methods. *Journal of Sensory Studies*, *39*(3). https://doi.org/10.1111/joss.12931

Giacalone, D., Clausen, M. P., & Jaeger, S. R. (2022). Understanding barriers to consumption of plant-based foods and beverages: insights from sensory and consumer science. In *Current Opinion in Food Science* (Vol. 48). Elsevier Ltd. https://doi.org/10.1016/j.cofs.2022.100919

Halme, M., Pirttilä-Backman, A. M., & Pham, T. (2023). The perceived value of oat milk and the food-choice motives of young, urban people. *British Food Journal*, *125*(13), 375–389. https://doi.org/10.1108/BFJ-03-2022-0238

Harwood, W. S., & Drake, M. A. (2018). Identification and characterization of fluid milk consumer groups. *Journal of Dairy Science*, *101*(10), 8860–8874. https://doi.org/10.3168/jds.2018-14855

Jaeger, S. R. (2006). Non-sensory factors in sensory science research. *Food Quality and Preference*, *17*(1–2), 132–144. https://doi.org/10.1016/j.foodqual.2005.03.004

Jaeger, S. R., Dupas de Matos, A., Frempomaa Oduro, A., & Hort, J. (2024). Sensory characteristics of plant-based milk alternatives: Product characterisation by consumers and drivers of liking. *Food Research International*, *180*. https://doi.org/10.1016/j.foodres.2024.114093

Jaeger, S. R., & Giacalone, D. (2021). Barriers to consumption of plant-based beverages: A comparison of product users and non-users on emotional, conceptual, situational, conative and psychographic variables. *Food Research International*, *144*. https://doi.org/10.1016/j.foodres.2021.110363

Jo, Y., Benoist, D. M., Barbano, D. M., & Drake, M. A. (2018). Flavor and flavor chemistry differences among milks processed by high-temperature, short-time pasteurization or ultra-pasteurization. *Journal of Dairy Science*, *101*(5), 3812–3828. https://doi.org/10.3168/jds.2017-14071

Mäkinen, O. E., Uniacke-Lowe, T., O'Mahony, J. A., & Arendt, E. K. (2015). Physicochemical and acid gelation properties of commercial UHT-treated plant-based milk substitutes and lactose free bovine milk. *Food Chemistry*, *168*, 630–638. https://doi.org/10.1016/j.foodchem.2014.07.036

Martínez-Padilla, E., Faber, I., Petersen, I. L., & Vargas-Bello-Pérez, E. (2023). Perceptions toward Plant-Based Milk Alternatives among Young Adult Consumers and Non-Consumers in Denmark: An Exploratory Study. *Foods*, *12*(2). https://doi.org/10.3390/foods12020385

McCarron, R., Methven, L., Grahl, S., Elliott, R., & Lignou, S. (2024). Oat-based milk alternatives: the influence of physical and chemical properties on the sensory profile. *Frontiers in Nutrition*, *11*. https://doi.org/10.3389/fnut.2024.1345371

McCarthy, K. S., Parker, M., Ameerally, A., Drake, S. L., & Drake, M. A. (2017). Drivers of choice for fluid milk versus plant-based alternatives: What are consumer perceptions of fluid milk? *Journal of Dairy Science*, *100*(8), 6125–6138. https://doi.org/10.3168/jds.2016-12519

McClements, D. J., Newman, E., & McClements, I. F. (2019). Plant-based Milks: A Review of the Science Underpinning Their Design, Fabrication, and Performance. In *Comprehensive Reviews in Food Science and Food Safety* (Vol. 18, Issue 6, pp. 2047–2067). Blackwell Publishing Inc. https://doi.org/10.1111/1541-4337.12505

Mills, J. (2024). Milk and non-dairy milk – US – 2024. Mintel. https://www.mintel.com/

Moss, R., Barker, S., Falkeisen, A., Gorman, M., Knowles, S., & McSweeney, M. B. (2022). An investigation into consumer perception and attitudes towards plant-based alternatives to milk. *Food Research International*, *159*. https://doi.org/10.1016/j.foodres.2022.111648

Orme, B. 2009. Maxdiff analysis: Simple counting, individual-level logit, and hb. Sawtooth Software Research Paper Series. Sawtooth Software, Sequim, WA.

Orme, B. (2019). *Consistency Cutoffs to Identify "Bad" Respondents in CBC, ACBC, and MaxDiff*. www.sawtoothsoftware.com

Osborn, S. (2015). Labelling relating to natural ingredients and additives. In *Advances in Food and Beverage Labelling: Information and Regulations* (pp. 207–221). Elsevier Inc. https://doi.org/10.1533/9781782420934.3.207

Patra, T., Axel, C., Rinnan, Å., & Olsen, K. (2022). The physicochemical stability of oat-based drinks. *Journal of Cereal Science*, *104*. https://doi.org/10.1016/j.jcs.2022.103422

Paul, A. A., Kumar, S., Kumar, V., & Sharma, R. (2020). Milk Analog: Plant based alternatives to conventional milk, production, potential and health concerns. In *Critical Reviews in Food Science and Nutrition* (Vol. 60, Issue 18, pp. 3005–3023). Bellwether Publishing, Ltd. https://doi.org/10.1080/10408398.2019.1674243

Ramsing, R., Santo, R., Kim, B. F., Altema-Johnson, D., Wooden, A., Chang, K. B., Semba, R. D., & Love, D. C. (2023). Dairy and Plant-Based Milks: Implications for Nutrition and Planetary Health. In *Current Environmental Health Reports* (Vol. 10, Issue 3, pp. 291–302). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s40572-023-00400-z

Rasika, D. M., Vidanarachchi, J. K., Rocha, R. S., Balthazar, C. F., Cruz, A. G., Sant'Ana, A. S., & Ranadheera, C. S. (2021). Plant-based milk substitutes as emerging probiotic carriers. In *Current Opinion in Food Science* (Vol. 38, pp. 8–20). Elsevier Ltd. https://doi.org/10.1016/j.cofs.2020.10.025

Riofrio, A., & Baykara, H. (2022). Techno-environmental and life cycle assessment of 'oat-milk' production in Ecuador: A cradle-to-retail life cycle assessment. *International Journal of Food Science and Technology*, *57*(8), 4879–4886. https://doi.org/10.1111/ijfs.15828

Sauerwein, E., Bailom, F., Matzler, K., & Hinterhuber, H.H. (1996). The kano model: how to delight your customers. ResearchGate, https://www.researchgate.net/publication/240462191

Sawtooth Software. (2014). *ACBC Technical Paper The Adaptive Choice-Based Conjoint (ACBC) Technical Paper*. www.sawtoothsoftware.com

Schiano, A. N., & Drake, M. A. (2021). Consumer understanding of fluid milk and cheese processing and composition. *Journal of Dairy Science*, *104*(8), 8644–8660. https://doi.org/10.3168/jds.2020-20057

Schiano, A. N., Harwood, W. S., Gerard, P. D., & Drake, M. A. (2020). Consumer perception of the sustainability of dairy products and plant-based dairy alternatives. *Journal of Dairy Science*, *103*(12), 11228–11243. https://doi.org/10.3168/jds.2020-18406

Sethi, S., Tyagi, S. K., & Anurag, R. K. (2016). Plant-based milk alternatives an emerging segment of functional beverages: a review. In *Journal of Food Science and Technology* (Vol. 53, Issue 9, pp. 3408–3423). Springer India. https://doi.org/10.1007/s13197-016-2328-3

Sharif Ullah, and A.M.M., & Tamaki, J. (2010). Analysis of Kano-model-based customer needs for product development. The Journal of the International Council on Systems Engineering, 14(2), 154-172, https://doi.org/10.1002/sys.20168

Silva, A. R. A., Silva, M. M. N., & Ribeiro, B. D. (2021). Plant-based milk products. In *Future Foods: Global Trends, Opportunities, and Sustainability Challenges* (pp. 233–249). Elsevier. https://doi.org/10.1016/B978-0-323-91001-9.00025-6

Silva, A. R. A. da, Santelli, R. E., Braz, B. F., Silva, M. M. N., Melo, L., Lemes, A. C., & Ribeiro, B. D. (2024). A Comparative Study of Dairy and Non-Dairy Milk Types:

Development and Characterization of Customized Plant-Based Milk Options. *Foods*, *13*(14), 2169. https://doi.org/10.3390/foods13142169

USDA ARS. (2024). FoodData central: oat milk, unsweetened, plain, refrigerated. Accessed 9/23/24. https://fdc.nal.usda.gov/fdc-app.html#/food-details/2257046/nutrients

Vaikma, H., Kaleda, A., Rosend, J., & Rosenvald, S. (2021). Market mapping of plant-based milk alternatives by using sensory (RATA) and GC analysis. *Future Foods*, *4*. https://doi.org/10.1016/j.fufo.2021.100049

Vashisht, P., Sharma, A., Awasti, N., Wason, S., Singh, L., Sharma, S., Charles, A. P. R., Sharma, S., Gill, A., & Khattra, A. K. (2024). Comparative review of nutri-functional and sensorial properties, health benefits and environmental impact of dairy (bovine milk) and plant-based milk (soy, almond, and oat milk). *Food and Humanity*, *2*, 100301. https://doi.org/10.1016/j.foohum.2024.100301

Yu, Y., Li, X., Zhang, J., Li, X., Wang, J., & Sun, B. (2023). Oat milk analogue versus traditional milk: Comprehensive evaluation of scientific evidence for processing techniques and health effects. In *Food Chemistry: X* (Vol. 19). Elsevier Ltd. https://doi.org/10.1016/j.fochx.2023.100859

Zhou, S., Jia, Q., Cui, L., Dai, Y., Li, R., Tang, J., & Lu, J. (2023). Physical–Chemical and Sensory Quality of Oat Milk Produced Using Different Cultivars. *Foods*, *12*(6). https://doi.org/10.3390/foods12061165

Table 2.1 Conjoint factors, levels, and consumer selections for oat milk during the Build-Your-Own (BYO) exercise for total population and clusters.

	Total Population (n=415)	Price Driven Cluster (n=272)	Balanced Cluster (n=143)
Price per half gallon (64 oz)			
Average	$5.61	$5.37	$6.06
Standard Deviation	$1.24	$1.12	$1.34
Flavor			
Original/Plain	63.4%	59.6%	70.6%
Vanilla	30.4%	34.6%	22.4
Chocolate/Other Flavor	6.3%	5.9%	7.0%
Type			
Original	31.8%	33.5%	28.7%
Extra creamy	28.0%	26.1%	31.5%
Unsweetened	24.8%	25.4%	23.8%
Zero sugar	8.2%	8.8%	7.0%
Low fat	7.2%	6.3%	9.1%
Label Claim			
Excellent source of calcium, vitamins A and D	42.9%	48.2%	32.9%
Plant based	17.3%	16.9%	18.2%
Good/excellent source of protein (+$2.60)	15.7%	12.9%	21.0%
Good source of calcium, vitamins A and D	11.3%	12.9%	8.4%
Organic (+$2.60)	6.3%	3.3%	11.9%
Non GMO	4.1%	3.7%	4.9%
Vegan	2.4%	2.2%	2.8%
Free-from Claim			
Free from artificial flavors and colors	44.6%	44.5%	44.8%
Dairy free	18.1%	19.5%	15.4%
Lactose free	16.1%	17.3%	14.0%
No added oil	12.8%	11.8%	14.7%
Gluten free	4.6%	3.3%	7.0%
Nut free	2.2%	2.2%	2.1%
Soy free	1.7%	1.5%	2.1%
Brand			
Store brand	26.3%	31.3%	16.8%
Silk (+$0.30)	22.7%	23.9%	20.3%
Chobani (+$1.15)	16.4%	14.7%	19.6%
Oatly (+$1.25)	15.4%	12.1%	21.7%
Planet Oat (+$1.00)	12.5%	14.0%	9.8%
Califia Farms (+$1.70)	6.7%	4.0%	11.9%

Data represents n=415 participants. One option was selected from each category (flavor, type, label claim, free-from claim, brand) based on participants' preferred oat milk attributes. Price per half gallon was calculated and provided to participants using a base price of $4.35. Levels that incurred increased cost were clearly indicated during this exercise.

Table 2.2 Consumer perception of milks by attributes via sliding scales

Attribute	Almond milk	Coconut milk	Milk (dairy)	Oat milk	Soy milk
Sustainable[1]	56.2c	61.6b	47.4d	70.3a	64.5b
Nutritious[2]	70.2b	62.5d	79.9a	65.6c	63.6cd
Sugar[3]	53.5b	59.5a	54.8b	54.5b	49.4c
Heart Healthy[4]	73.8a	62.6b	59.4c	72.4a	65.1b
Taste[5]	68.8b	59.7c	76.6a	69.4b	50.1d
Texture[6]	66.1b	59.4c	77.8a	67.1b	55.2d
Price[7]	65.5b	69.4a	40.8c	69.3a	62.9b
Clean Label[8]	65.2a	64.4a	64.8a	67.1a	60.3b
List Length[9]	58.4b	61.1b	80.6a	58.0b	53.1c
Familiarity[10]	71.3b	66.8c	83.4a	67.9c	58.4d
Allergens[11]	43.7d	56.5b	33.1e	62.7a	48.3c
Protein[12]	62.9b	47.7d	80.5a	53.8c	63.4b

Data represents n=415 participants. Milks were placed on a sliding scale for each attribute relative to each other and the attribute anchors. Different letters following means in a row indicate significant differences ($p<0.05$). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

Attribute anchors (0–100): [1]not at all sustainable–very sustainable, [2]not at all nutrient rich (vitamins & minerals)–nutrient rich (vitamins & minerals), [3]hardly any sugar–lots of sugar, [4]not at all heart healthy–heart healthy, [5]does not taste great–tastes great, [6]bad texture–great texture, [7]not at all expensive–very expensive, [8]not at all clean label–very clean label, [9]long ingredient list–short ingredient list, [10]made with unfamiliar ingredients–made with familiar ingredients, [11]lots of allergens–allergen free, [12]hardly any protein–lots of protein.

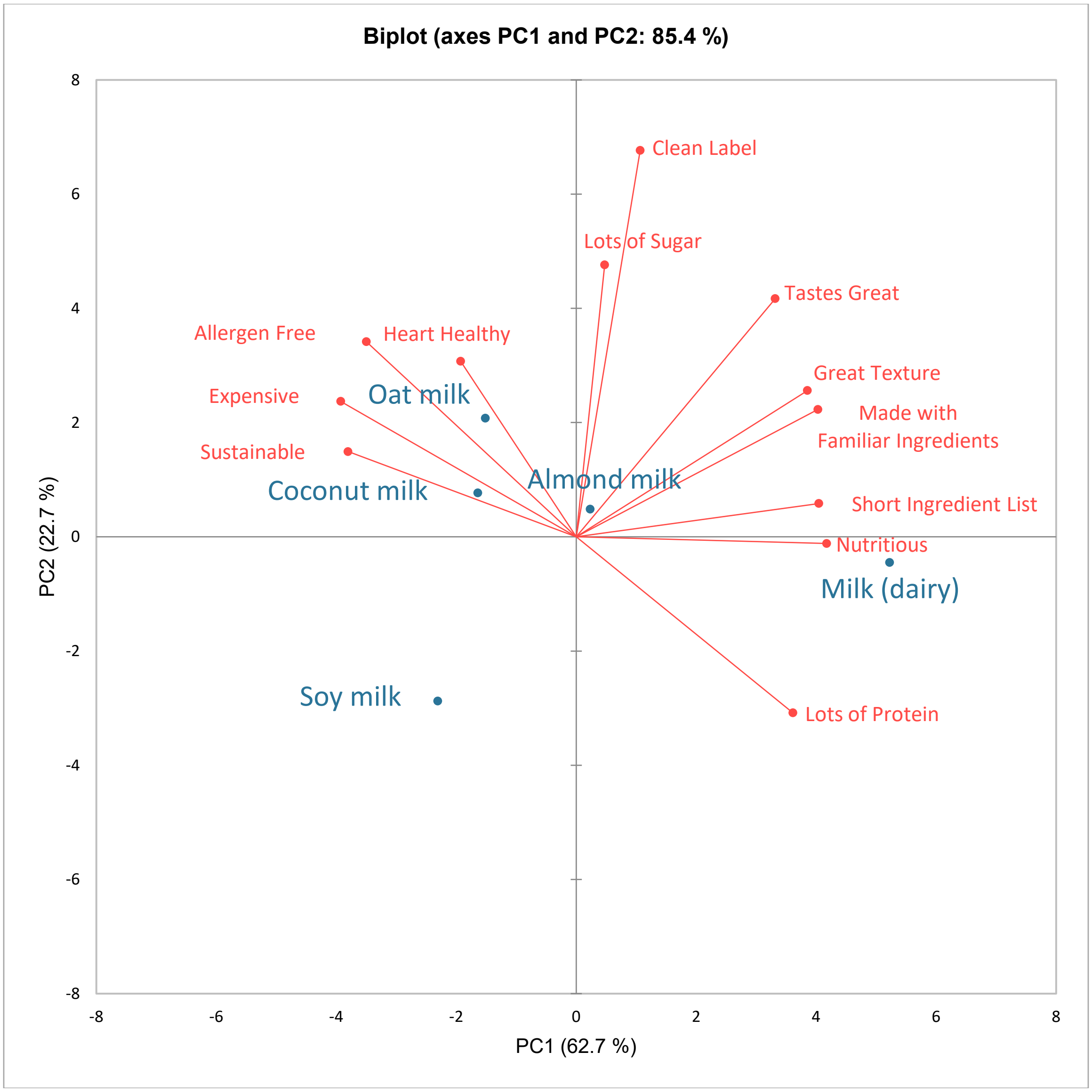


Figure 2.1 PCA biplot of consumer perception (n=415) of different types of milk (blue) scored on a sliding scale (0-100) for 12 attributes (red). This figure is a graphical representation of the data from Table 2.2.

Trial Motivations

Allergies/Sensitivity

Oat does not contain allergens, unlike other plant-based milks (nut, soy)

"Easy on my stomach, sometimes dairy milk doesn't work for me."

Variety of Milks

Consumers like trying new products

"I was purchasing something new that I had never tried before."

Trendy, Fear of Missing Out

"When it started to be more trendy a year or two ago, I decided to try it."

"Everyone was talking about it [Starbucks oat milk drink]."

Health

Shift away from dairy, perception of less fat, healthier for environment

"I was trying to get away from drinking so much dairy…now I only really drink plant milk"

Hesitation

"When I see new stuff I am scared. I don't want to waste my money…waited until I had a coupon"

Alternative milks are *"definitely more expensive…when it comes to your health [maybe it's worth it]…[considering tradeoffs] you have to balance your budget."*

Trial and Error

Perfect addition to ...

- Coffee
- Smoothies
- Cereal
- Oatmeal
- Drinking on Own

Because it provides...

Flavor that is unique (oaty, nutty), delicious, subtly sweet, and has a thick & creamy texture.

"A little nutty, creamy…it added more dimension to the coffee than regular milk because it has a different flavor."

"I really like the taste of oats… cereal, drinking on its own."

Does NOT work for...

- Cooked applications
- Savory dishes
- Cream sauces
- Mac & Cheese

Because it...

Breaks down during cooking, adds oat flavor for an application it is unwanted, and doesn't pair well with savory applications.

"As creamy as oat milk is, it does NOT work when it is cooked…somehow just breaks down entirely."

"If I was going to make a chicken alfredo, I don't want oat dancing up in that."

Opportunity

Oat Milk for Cooking and Baking

"I do not buy cow's milk for anything other than baking."

"If it worked well I might make that switch."

Clarification of Flavor Profile

"Milk bought anywhere is normal milk…there are all these different types and varieties of oat milk… oh god, there are so many different choices."

If a package could guarantee *"no secondary tastes, just oat"* and it was a *"predictable flavor, that would be ideal"*

Repeat Motivations

Shelf-life

"Alternative milks [open in the refrigerator] have a lot longer shelf life than milk [dairy]"

Longer shelf life is…*"saving me money because it's not expiring."*

Comparing shelf-stable to refrigerated...

"Shelf stable oat milk tastes like oat milk. I don't like the taste of shelf stable milk [dairy]."

Variety

Enjoy having multiple types of milk in the fridge at a time. Each have their own unique flavor and uses.

"I get bored. I like to rotate…three separate types of milk in the fridge."

"I like having a variety [of milks]. I feel like a mad scientist [adding different types of milk]"

Flavor and Texture

Oat milk *"is supposed to be different. I don't want it to be like milk… I'm using it specifically to add sweetness and flavor."*

When they crave oat milk, they want the unique oaty/nutty flavor and like that it is thicker/creamier than other plant-based options

"Oat milk is thicker and creamier than other dairy free options."

Positive Impact of Plant

"I feel better about myself when I buy plant milk."

Figure 2.2 Insights from four focus group, each two hours long, with oat milk consumers (n=24). Themes represent consumer sentiment that was consistent across focus groups. Selected quotes from consumers support the themes in the consumers' voice.

Table 2.3 Frequency of oat milk consumption by application

	Drinking on its own (n=147)	Adding to cereal/ oatmeal (n=266)	Adding to smoothie (n=282)	Adding to coffee/tea (n=296)	Using as an ingredient in a recipe (n=241)
Multiple times per day	0.7%	0.5%	0.2%	2.4%	0.5%
Once per day	2.9%	4.8%	5.1%	11.6%	1.9%
A few times per week	8.0%	14.7%	10.1%	15.9%	9.4%
Once per week	3.4%	7.2%	10.1%	7.7%	8.2%
A few times per month	8.9%	22.9%	22.4%	19.8%	17.1%
Once per month	5.5%	7.7%	9.6%	5.5%	9.6%
Less than once per month	6.0%	6.3%	10.4%	8.4%	11.3%

Data represents n=415 consumers. Frequency was only asked to participants who indicated they use oat milk for the applications listed. The number of consumers who use oat milk for each occasion is indicated in the table header.

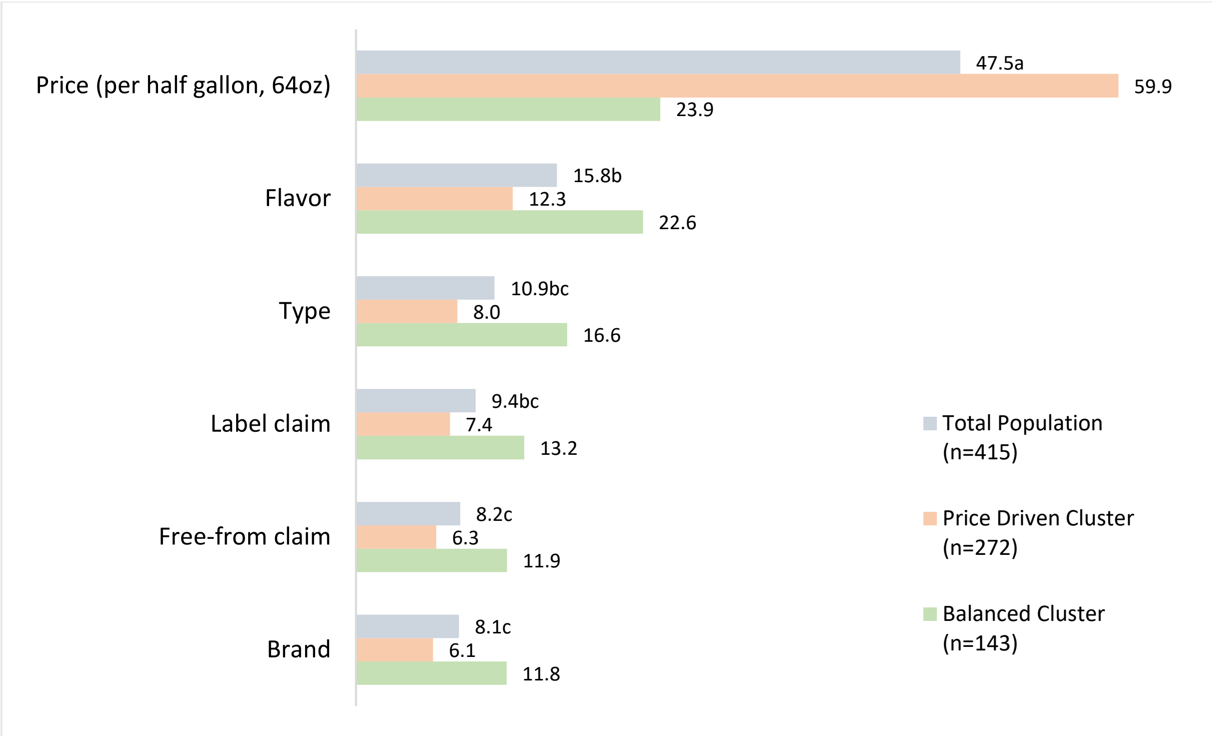


Figure 2.3 Importance scores from ACBC exercise for the total population and each cluster. Higher numbers indicate that the attribute was more important to consumers when purchasing oat milks. Statistical lettering is reported for the total population ($p<0.05$). Segmentation was performed using AHC with K-means cluster analysis.

Table 2.4 Chip allocation for oat milk purchase factors

Attribute	Mean
Price	37.2
Flavor	20.7
Nutritional profile	14.9
Other attributes (creamy, sugar free, low fat, etc.)	12.7
Brand	7.5
Label claim (good source, organic, gluten free, etc.)	7.1

Data represents n=415 participants. Attributes were evaluated using a constant sum (chip allocation) exercise where participants allocated a total of 100 points across the six attributes specified. Mean point allocation for each attribute is reported.

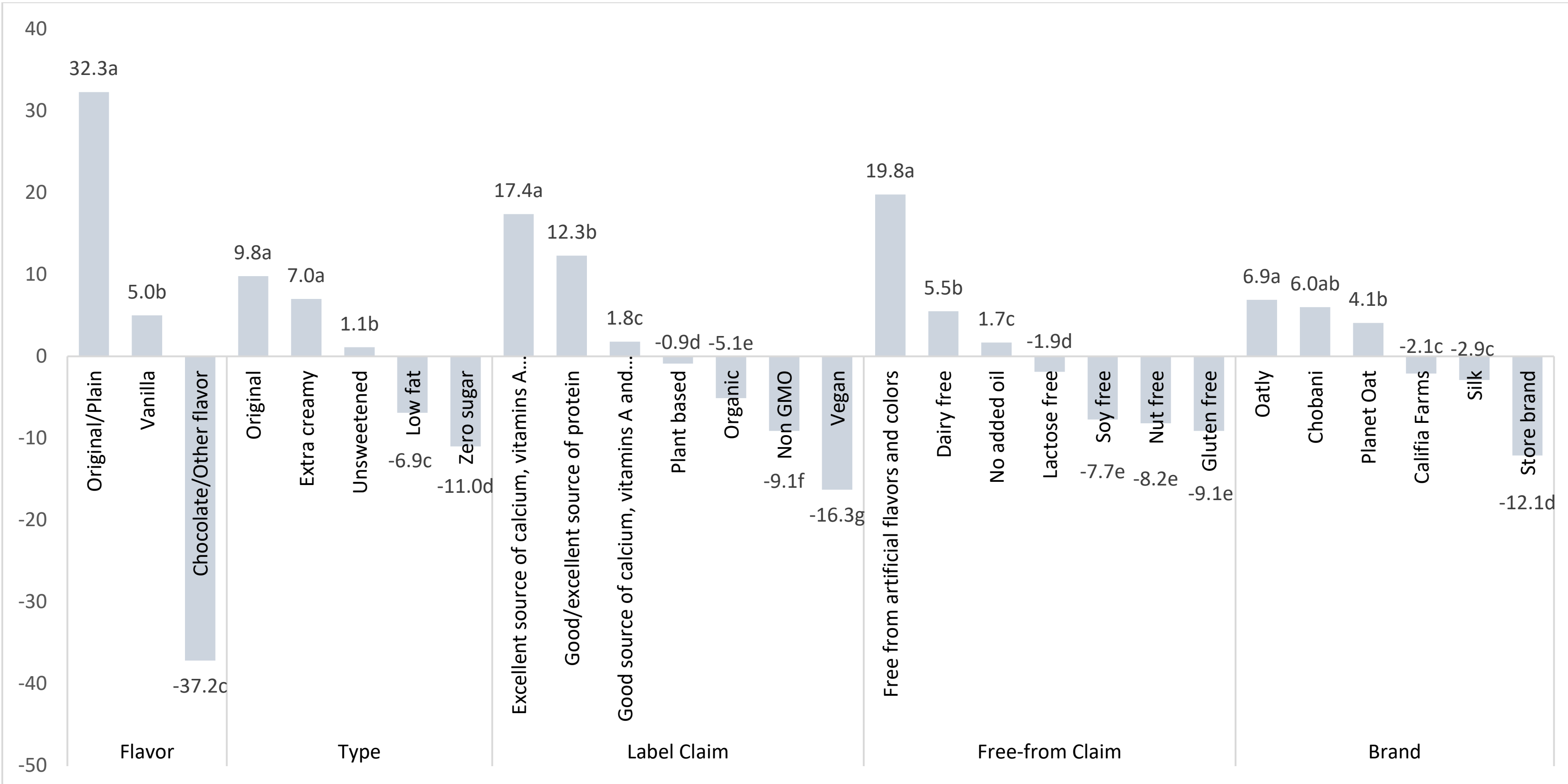


Figure 2.4 Utility Scores from ACBC Exercise for Total Population (n=415).

Utility scores are zero-centered interval scaled. Higher numbers indicate that the level was more desirable to consumers. Different letters within an attribute indicate significant differences between the levels ($p<0.05$). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

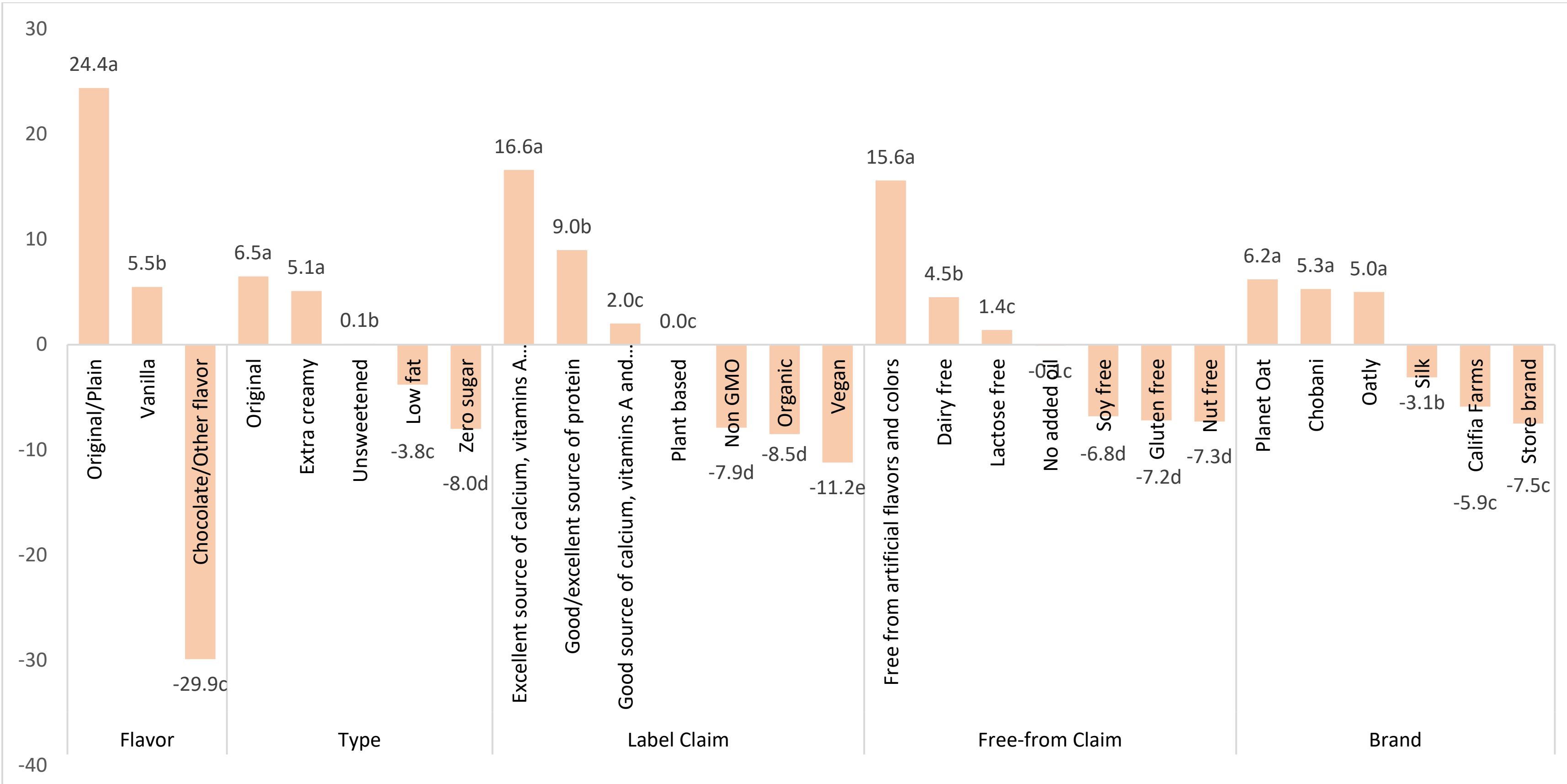


Figure 2.5 Utility Scores from ACBC Exercise for Price Driven Cluster (n=272).

Utility scores are zero-centered interval scaled. Higher numbers indicate that the level was more desirable to consumers. Different letters within an attribute indicate significant differences between the levels ($p<0.05$). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

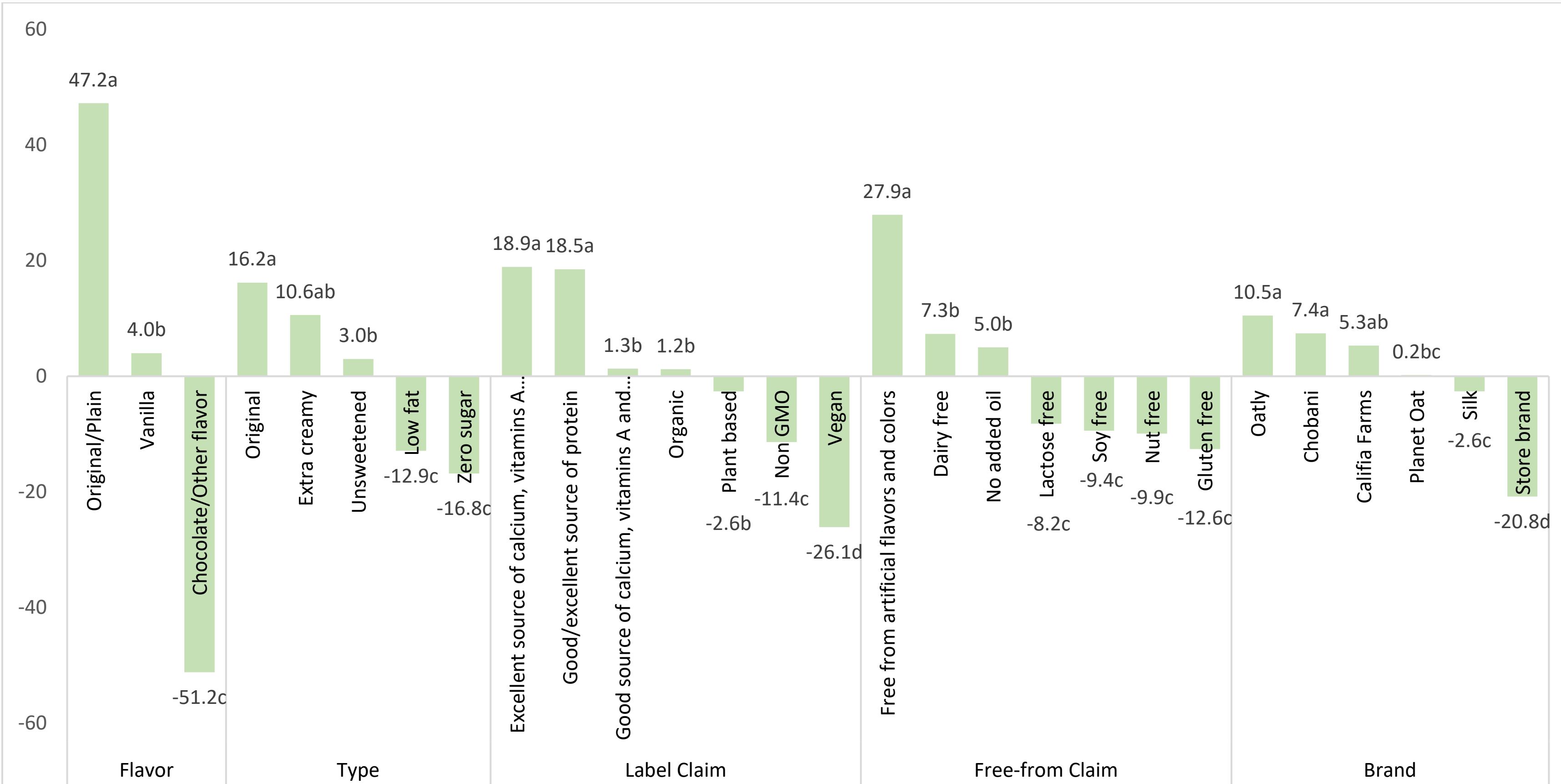


Figure 2.6 Utility Scores from ACBC Exercise for Balanced Cluster (n=143).

Utility scores are zero-centered interval scaled. Higher numbers indicate that the level was more desirable to consumers. Different letters within an attribute indicate significant differences between the levels (p<0.05). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

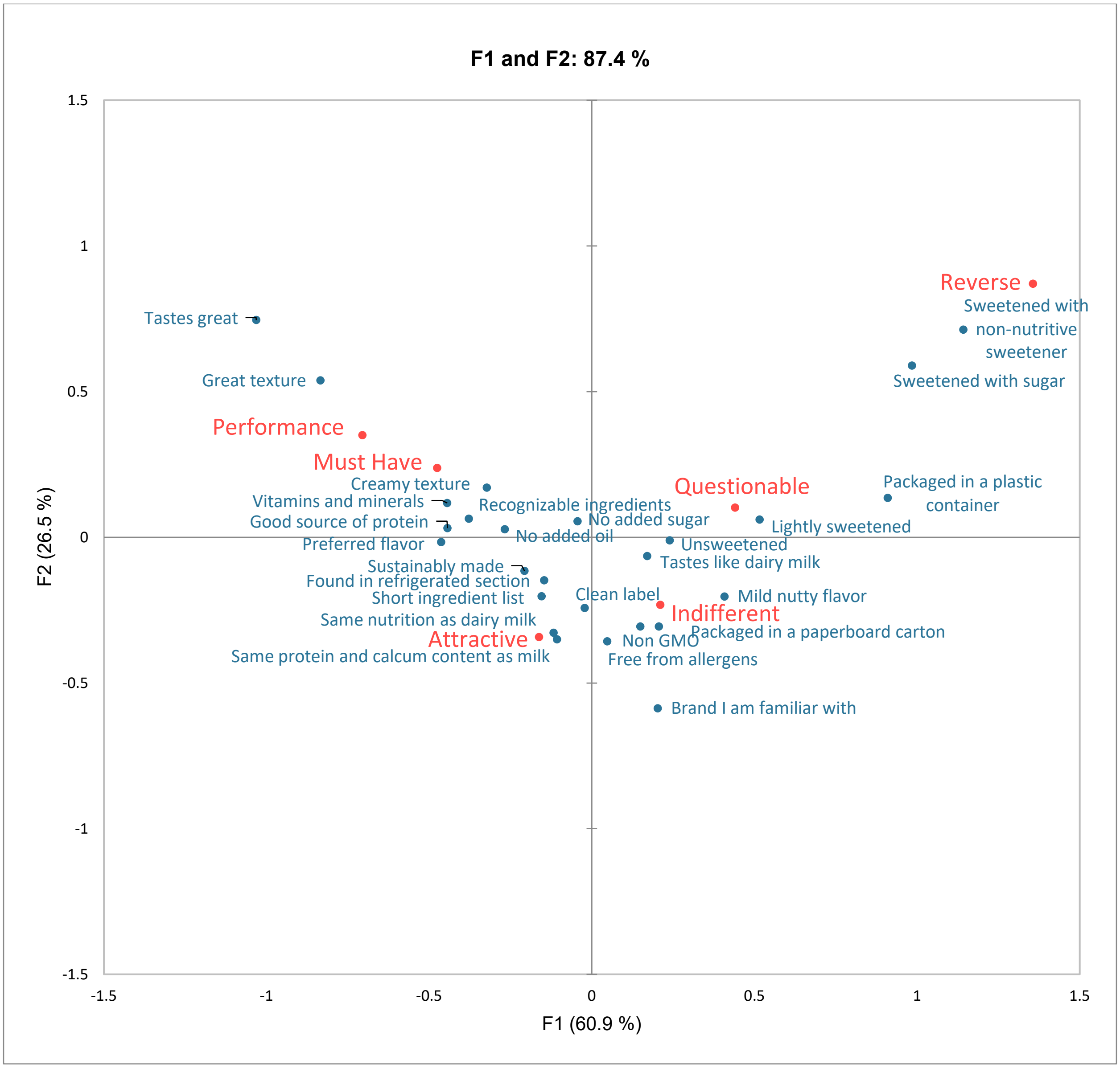


Figure 2.7 Correspondence analysis biplot of Kano modeling classifications (red) of oat milk attributes (blue) for all participants (n=415).

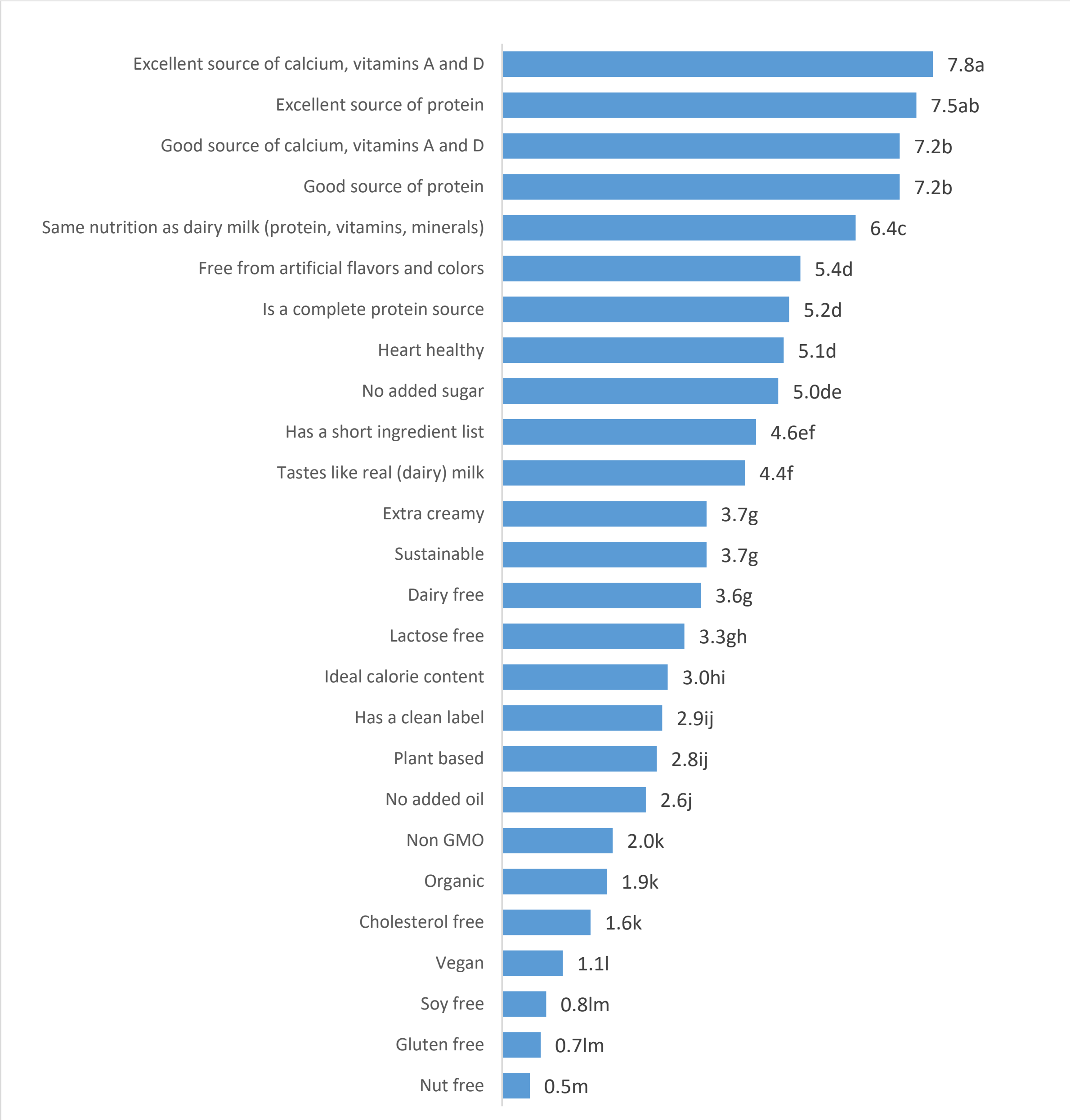


Figure 2.8 Utility Scores from MaxDiff Label Claim Exercise (n=415).

Utility scores are probability scaled responses. Higher numbers indicate that the label claim was more desirable to consumers. Different letters indicate significant differences between the label claims ($p<0.05$). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

CHAPTER 3:

The Impact of Presentation Format and Coffee on Consumer Preferences for Coffee Creamer

The Impact of Presentation Format and Coffee on Consumer Preferences for Coffee Creamer

D. Rovai, K. Homwongpanich, M. Watson, M. A. Drake

Department of Food, Bioprocessing and Nutrition Sciences, Southeast Dairy Foods Research Center, Raleigh, North Carolina, USA

Choice of journal/topic: Journal of Sensory Studies, submitted

Abstract

Consumer evaluation of coffee creamers necessitates evaluation in a cup of coffee. Two phases of research were used to determine if the method of serving creamers or the coffee they were evaluated with influenced creamer acceptance. In phase I, consumers evaluated creamers (n=4) served using one of three methods: 1) Fixed amount (n=127), 2) Free pour (n=120), and 3) Warm-up (n=122). In phase II, consumers (n=134) evaluated creamers in two coffees (light and dark roast). Data were analyzed by univariate and multivariate statistics. From phase 1, significant method*creamer interactions were observed for appearance and aroma liking ($p<0.05$), but key insights were not impacted by choice of method. From phase 2, no significant coffee*creamer interaction was observed for overall liking ($p>0.05$), but significant interactions ($p<0.05$) were observed for many other liking attributes (appearance, aroma, flavor, mouthfeel, creaminess). All methods are suitable for consumer evaluation of coffee creamers, but researchers should consider the impact of coffee on creamer liking.

Practical Application

Coffee creamers are consumed in the context of a cup of coffee, and sensory evaluation with consumers should seek to mimic the consumer experience. Adding coffee as a carrier for creamers during consumer evaluation introduces extraneous variables. This research evaluated the impact of serving method and coffee on creamer evaluation. Insights from this study can aid in choosing an appropriate method for consumer evaluation of coffee creamers.

Key Words: creamer, coffee, methodology, roast

3.1 Introduction

U.S. retail sales of coffee are expected to reach $19.8 billion by the end of 2024 (Olsen, 2024). Consumers look to coffee as a source of energy, comfort, routine, and enjoyment. The act of coffee consumption is ritualistic, serving as a marker to start the day or signify a transition in the daily routine. Consumers drink their coffee in a variety of forms, but hot coffee is the most popular way to consume coffee (Olsen, 2024). Many factors affect the sensory perception of coffee including the bean origin, preparation technique (Elmaci & Gok, 2021), brewing method (Pereira et al., 2023), and bean roasting temperature (Carcea et al., 2023). In addition, there are many ways for consumers to customize a cup of coffee. Adding cream/creamers is the most popular addition to coffee, with about 88% of consumers indicating they add some form to their coffee (Olsen, 2024). The FDA does not define coffee creamer but does provide a definition for coffee cream: “light cream or ‘coffee cream’ contains not less than 18% but less than 30% milkfat and may contain ingredients such as stabilizers, emulsifiers, nutritive sweeteners, flavoring ingredients” (21CFR131.155). However, this definition does not fully encompass the range of products referred to as coffee creamers in the United States. Otherwise referred to as whiteners or lighteners, creamers were first introduced on the market in 1950 as a powdered shelf-stable non-dairy creamer (Shurtleff & Aoyagi, 2013). Non-dairy creamers are made using oil, sweeteners, emulsifiers, and flavors to create a mouthfeel similar to dairy cream. Today, creamers have expanded from non-dairy powdered creamers to include liquid non-dairy creamers, plant-based creamers, and dairy creamers. Liquid coffee creamers differ from cream or half and half because they contribute added sweetness and/or added flavors in addition to providing whitening and mouthfeel.

Adding creamers or other products like milk to coffee enhances the sensory characteristics of coffee, making it more appealing to consumers (Komes et al., 2015). Coffee consumers tend to enjoy some of the bitterness from coffee, but too much bitterness in coffee decreases consumer liking (Mahmud et al., 2021). The addition of creamer cuts some of the bitterness and acidity from the coffee, as well as cooling down the coffee to make it easier to consume, adding mouthfeel/body, sweetness, and flavor variety. The fat content in creamers affects the release of aroma compounds from the coffee, modifying the taste of coffee when it is added (Parat-Wilhels et al., 2005). In addition, the proteins in milk/creamer have been found to bind to phenolic compounds in coffee that are responsible for bitter and astringent responses in the mouth (Rashidinejad et al., 2022). For these reasons, consumers continue to add creamer to their coffee, and the projected sales of cream and creamer are $9.1 billion in 2024, with $5.6 billion coming from creamer alone (Olsen, 2024).

When performing sensory evaluation of a food product, best practices are to mimic consumer behavior as closely as possible, while simultaneously controlling for external variables (Lawless & Heymann, 2010). For some food products, mimicking consumer behavior entails consuming the product with a carrier (Lawless and Heymann, 2010). This process is common in products like condiments which are rarely eaten plain by consumers. Research evaluating the sensory perception of condiments and carriers found that the perception of both the condiments and the carrier changed when they were combined with each other (van Eck et al., 2019). This is important to understand from a consumer perspective because it emphasizes the need for products to be evaluated in the context of how consumers use them to account for some of the variability that comes with the addition of carriers.

Research by Cherdchu & Chambers (2014) used descriptive analysis to document the profile of soy sauces with different carriers and found that different carriers affected the perceived intensity of product attributes. Another study examined the role of soy sauce carriers with consumers and found that while the carrier and the soy sauce impacted consumer perception, there was no significant interaction between the two (Keefer et al., 2021). The decision to include a carrier in sensory evaluation is influenced by the research objective. For example, descriptive studies that rely on trained panelists to document objective product attributes akin to an instrument may evaluate a product on its own, while consumer studies that document subjective product liking may benefit more from evaluating products with a carrier. This practice is observed in research with butter where trained panelists evaluated butter plain, but consumers evaluated butter in a baked application (Jinjarak et al., 2006), on pita bread (Krause et al., 2007) or on crackers (Garvey et al., 2020).

Creamer is a product that is used almost exclusively as an additive to coffee and is not consumed by consumers on its own. Thus, when evaluating consumer acceptance of creamer, it should be evaluated in the context of a cup of coffee. However, the introduction of coffee brings additional variables such as the type of coffee, the brewing parameters, the strength of the coffee, the coffee:creamer ratio, and the procedure for adding the creamer to the coffee. While research exists which discusses the impact of coffee parameters such as type of bean, coffee:water ratio, roast, and preparation on consumer acceptability of coffee (Andueza et al., 2007; Boeneke et al., 2007; Cotter et al., 2021; Elmaci & Gok, 2021; Carcea et al., 2023; Pereira et al., 2023), minimal research exists that addresses the sensory evaluation of creamers. The only study to the authors' knowledge which examined consumer evaluation of creamers (Fibrianto & Maharani, 2021) does not provide specifics about how the creamer was served, the amount served, or the type of coffee that it was evaluated with. Thus, minimal literature exists which explores the role of context of carrier as it pertains to coffee creamer evaluation. The objective of this research was to understand if the methodology of evaluating creamer in coffee has an effect on creamer liking. To address this goal, the study was comprised of two parts. Phase I of the study aimed to understand if the method of serving creamer in coffee impacted liking, and phase II aimed to understand if the type of coffee (light vs dark) impacted creamer liking in coffee.

3.2 Methods

3.2.1 Overview

This study was divided into two phases (Figure 3.1). In each phase, coffee consumers who added creamer to their coffee evaluated four commercial coffee creamers, each in a cup of coffee. Phase I evaluated the impact of serving method on creamer evaluation: coffee creamers were evaluated using three distinct serving methods across different days. Phase II evaluated the impact of the type of coffee on creamer evaluation. All testing with consumers was conducted in compliance with North Carolina State University Institutional Review Board regulations (exempted IRB 26494).

3.2.2 Creamers and Coffees

Four commercial vanilla-flavored coffee creamers were chosen to represent brands with high market share (Olsen, 2024) as well as different types of creamers. Creamer A and C were dairy creamers, Creamer B was a nondairy creamer, and Creamer D was a plant-based creamer (almond). All creamers were stored at refrigerated temperatures (4°C) and evaluated no less than 2 weeks of their best by date. The same four creamers were evaluated in Phase I and Phase II of this study.

Three commercial ground coffees were chosen to represent different roasts and flavor profiles. For phase I, a medium roast coffee was chosen, as the objective of this phase was to evaluate the different serving methods, not the coffee. For phase II where the objective was to determine the effect of the coffee on creamer evaluation, two dissimilar coffees were chosen. Coffee A was a light roast coffee characterized by a beany and acidic profile, while coffee B was a dark roast coffee characterized by smoky and burnt flavors. All coffees were purchased pre-ground and brewed per manufacturer recommendations in commercial pour over brewers (Bunn, CWTF15-APS, PF, Bunn-O-Matic Corporation, Springfield, IL) at 82°C into Bunn pump pots and served to order (90 mL coffee) in a 178 mL styrofoam cup (Dart Jcup, PFS Sales, Raleigh, NC).

3.2.3 Recruitment and Consumers

Consumers were recruited from the North Carolina State University Sensory Service Center database of over 11,000 panelists via a screener survey through Compusense Cloud version 23.0.3 (Compusense Inc., Guelph, ON, Canada). Participants (18 - 64 y) were screened to meet the following criteria: consume coffee more than once per week, use creamer in their coffee at least 50% of the time, and were responsible for greater than 25% of household shopping. For phase I, a new set of consumers were recruited for each test day (n=120-127), while for phase II the same consumers participated in both days of testing (n=134). Phase 1 and phase 2 were conducted approximately 5 weeks apart. Consumers that completed a day of testing in phase I received a $10 gift card and those that participated in the 2-day test in phase II received a $30 gift card.

3.2.4 Phase I - Impact of Method

To evaluate the impact of method on creamer evaluation, three separate consumer taste tests were conducted across three weeks to investigate three different methods of serving creamer with coffee. These methods were chosen to represent industry standards and to explore the impact of standardizing the amount of creamer added to coffee among samples and among consumers. Each of the three methods are summarized in Figure 3.2. The first method of serving creamer was to present a fixed amount of creamer (30 mL creamer, 90 mL coffee) and instruct panelists to add the entire amount of creamer to the coffee (n=127). The fixed volume of creamer served was determined based on historical data from coffee creamer tests conducted by the Sensory Service Center at North Carolina State University. This "fixed amount method" standardized the creamer amount among different creamers and different consumers. The second method will be referred to as the "warm-up method" (n=122). Panelists were given an excess amount (100 mL) of creamer (vanilla, non-dairy, market leader; not one of the four product samples evaluated) in a 237 mL HDPE bottle and instructed to add creamer to 90 mL of a warm-up cup of coffee to their personal ideal amount (based on the color, taste, and any other sensory cues). The bottle of creamer was then returned to the server who calculated the volume of creamer added to this warm-up sample by subtracting the remaining volume of creamer from the initial volume. The calculated volume was then prepared for all samples evaluated by that consumer. Each consumer was instructed to add the entire pre-determined ideal volume of each of the four creamers to each coffee. This method allowed for variability in creamer amount between panelists but standardized the amount of creamer across samples for each individual. The "free pour method" (n=120) provided panelists with an excess amount (100 mL) of each creamer in a 237 mL HDPE bottle and instructed panelists to add each sample to the coffee according to their preference, noting that they may add as much or as little creamer as they liked. This method did not standardize the amount of creamer between samples or panelists, allowing for panelists to add creamer to their discretion, as they would if they were to use the products at home.

3.2.5 Phase II - Impact of Coffee

To evaluate the impact of coffee on creamer evaluation, a two-day consumer taste test (n=134) was chosen to evaluate 8 samples (4 creamers x 2 coffees) across two days (4 samples/day). A randomized complete block design (William's design) was used so that panelists received any 4 of the 8 samples each day. The serving method was constant in this phase of the experiment, and the fixed amount method was chosen to present the same volume of creamers to all consumers.

3.2.6 Questionnaire

An online questionnaire was developed using Compusense Cloud version 23.0.3 (Compusense Inc., Guelph, ON, Canada). The same questionnaire was used for all tests in phase I and phase II. The ballot aimed to capture differences in overall liking (key metric) as well as diagnostic attributes. All liking attributes (appearance creamer, aroma creamer, appearance in coffee, aroma in coffee, overall, sweetness, bitterness, overall flavor, vanilla flavor, creaminess) were scored on a 9-point scale where 1=dislike extremely and 9=dislike extremely. For all questions, consumers were specifically asked to rate attributes for the *creamers* in coffee (i.e. "Which statement best describes your impression of the vanilla flavor of this *creamer* in coffee?"). A three-minute rest was enforced between samples and participants were instructed to cleanse their palate with bottled purified water (Pure Life, 240 ml bottles, purchased locally) between samples.

3.2.7 Data Analysis

For each of the phases described above, a 2-way analysis of variance was conducted to determine the significance (α=0.05) of main effects (phase 1 = creamer, method; phase II = creamer, coffee) and the interaction (phase I = creamer*method, phase II = creamer*coffee). In phase I, individual consumers only saw one of the serving methods—Fixed amount (n=127), Free pour (n=120), Warm-up (n=122). Thus, consumer was considered a random effect nested within method for phase I. In phase II, each individual consumer evaluated all combinations of creamer and coffee. Consumer was considered a random effect for phase II. When applicable, Fisher's LSD was used as a post hoc test for means separation (α=0.05). Principal component analysis biplots were constructed using Pearson correlation coefficients for each phase of the study to visualize liking data for creamers. All statistical analysis was performed in XLSTAT 2023.3.1 (Lumivero, 2024).

3.3 Results and Discussion

3.3.1 Phase I - Impact of Method

For each sensory attribute, the f-value and p-value for the model, main effects, and the two-way interaction for the impact of method is summarized in Table 3.1. The model was significant ($p<0.05$) for all attributes. In every instance where a significant model was found, creamer was a significant main effect. This result was expected since these creamers were chosen to represent distinct differences within the creamer market space. A significant creamer*method interaction was found for appearance liking in coffee and aroma liking in coffee. A significant interaction suggests that for those two specific liking attributes, the serving method influenced liking scores. Table 3.2 summarizes the interactions between creamer and method for these two attributes. For three of the creamers (A, B, C), appearance liking was higher using the free pour and warm-up methods than the fixed amount method ($p<0.05$). However, for Creamer D, appearance liking was not statistically different for any of the serving methods ($p>0.05$). For aroma liking in coffee for creamers A, B, and C there was no significant difference within each creamer across methods ($p<0.05$). However, creamer D prepared using the warm-up method scored lower than creamer D prepared using either the free pour or fixed amount method.

The source of these interactions may come down to inherent differences within the creamers. Creamer D was a plant-based creamer and had an obvious color difference—brown color—compared to the other three samples which were whiter in color. For the creamers that were white in color (A, B, C), the methods with more room for customization of the amount of creamer (free pour, warm-up) had higher ratings for appearance liking because the consumer had some influence on reaching their target color, while they had no influence during the fixed amount method—they had to add the entire portion of creamer regardless of their preferences for amount of creamer. For Creamer D, the brown color limited the range of colors that could be achieved, so appearance liking was less impacted by the method and its impact on color variability. It is worth noting that the warm-up sample used for the warm-up method was a non-dairy creamer which was not brown in color. A method effect might not be evident for Creamer D if a plant-based creamer was used as the warm-up sample.

While a significant creamer*method interaction for appearance and aroma liking exemplifies that the choice of method may influence consumer liking for these attributes, results for overall liking of creamers did not differ based on the serving method used. Thus, any of these three methods are suitable for creamer evaluation and similar results would be expected regarding creamer preferences using all methods. The observed interactions for appearance and aroma liking in coffee may warrant more research, specifically looking at the impact of whitening power among plant-based creamers. A PCA biplot summarizes creamer liking with different methods (Figure 3.3). All liking attributes loaded positively on PC1 suggesting creamers loading in the positive direction on PC1 are more liked than those loading negatively on this component. This plot illustrates that samples were differentiated based on creamers, not based on methods. This observation is also supported by the higher F-values associated with the creamer main effect than the method effect or interaction, suggesting that more of the differences within the model are attributed to the different creamers (Table 3.1). These results support the notion that all methods were effective in differentiating creamers based on liking preference in the same way. Table 3.3 summarizes the mean creamer ratings for each of the attributes. Based on overall liking scores, Creamer B and Creamer C were the most preferred. These creamers also rated high for all liking attributes, although Creamer B has lower appearance and aroma liking of creamer and aroma liking in coffee compared to Creamers A and C.

Because each of these methods have been shown to discriminate between creamers in similar ways, the decision on which method to use depends on the research objective of the research and logistical considerations. These three methods represent an inverse relationship between mimicking consumer behavior and controlling extraneous variables. The fixed amount method has the most control of variability by ensuring that the same volume of creamer is used in all samples, but it is not a good representation of how most consumers add creamer to their coffee. The free pour method encounters the opposite problem. This method does the best job mimicking consumer behavior by allowing participants to add the creamer however they like, but results in a lack of control by allowing for different volumes of creamer to be used. The warm-up method falls somewhere in between by both of these metrics, approximating the volume of creamer used by each participant through a warm-up sample and then standardizing that amount across all samples.

The volume of creamer used in coffee varies among consumers. The mean volume of creamer used for 90 mL of coffee during the free pour method was 33.1±21.0 mL and the median was 27.7 mL. The average amount of creamer used was highest for Creamer D (35.0 mL) and lowest for Creamer B (30.8 mL). For the warm-up method, the mean volume of creamer used was 35.8±18.5 mL and the median was 25.0 mL. The fixed amount (30 mL) method did a good job capturing the average amount of creamer used by all consumers, falling in between the mean and median volumes of both the free pour and warm-up methods. However, the large standard deviations observed for both methods suggest that there is large variability in the amount of creamer added by consumers.

From a logistical perspective, the amount of time, resources, and staff needed to conduct a taste test varies for each of these methods. The fixed amount method is the easiest to implement, requires the smallest volume of creamer, and requires less preparation. Both the free pour and warm-up methods require significantly more volume of creamer to ensure excess creamer for each participant to account for variability in usage level. In addition, the warm-up method requires more staff interaction because of the need to record the volume used and portion sample volumes in real time. Thus, unless one of the research objectives is to understand volume usage or consumer behavior during pouring, the fixed amount method appears to be the most efficient and appropriate method to evaluate creamer liking with minimal volume and sample preparation.

3.3.2 Phase II - Impact of Coffee

For each sensory attribute, the F-value and p-value for the model, main effects, and the two-way interaction for the impact of coffee are summarized in Table 3.4. The model was significant ($p<0.05$) for all attributes. In every instance where a significant model was found, creamer was a significant main effect. This result is consistent with the results from Phase I of the study and further confirms that there are distinct preferences for the creamers used in this study. Significant creamer*coffee interactions were found for appearance liking in coffee, aroma liking in coffee, overall flavor liking, vanilla flavor liking, mouthfeel liking, and creaminess liking. This result indicates that, for these attributes and these representative creamers, creamer preferences differed depending on the coffee it was served with. Table 3.5 summarizes the interactions between creamer and coffee for attributes where the interaction was significant. For these attributes, some of the creamers were not significantly different when prepared in light roast coffee compared to dark roast coffee, while others were. For example, for aroma liking, Creamer A prepared in light roast coffee was not different than Creamer A prepared in dark roast coffee (nor was Creamer C) ($p>0.05$), however, Creamer B prepared in light roast coffee was significantly different than Creamer B in dark roast coffee and Creamer D in light roast coffee was significantly different than Creamer D in dark roast coffee ($p<0.05$). Thus, for these attributes, the type of coffee used to evaluate creamers would have an effect on the experimental conclusions (Table 3.5).

Table 3.6 summarizes the mean creamer ratings for each of the attributes. Creamer A, Creamer B, and Creamer C were liked more overall than Creamer D. Creamer D received consistently low ratings for all attributes. Creamer A and Creamer C received significantly higher ratings than Creamer B for appearance liking creamer, aroma liking creamer, and aroma liking in coffee. In addition, Creamer C had a lower rating for creaminess liking than Creamer A. These results differ slightly from the results for Phase I of the study, where Creamer B and Creamer C were rated significantly higher for overall liking than Creamer A and Creamer D. However, results should not be directly compared across phases because this study was not designed in a way to allow for such comparison.

A PCA biplot summarizes creamer liking with light roast and dark roast coffees (Figure 3.4). All liking attributes are loading positively on PC1 suggesting samples loading in the positive direction on PC1 are more liked than those loading in the negative direction on PC1. For each creamer, liking is higher for the creamer served with light roast coffee compared to the same creamer served with dark roast coffee. The high impact of coffee is exemplified by the high F-values associated with the effect of coffee (Table 3.4). This result is consistent with the findings from Table 3.7 which shows that consumers rated creamers higher when they were served with light roast coffee compared to dark roast coffee for all attributes aside from appearance liking and aroma liking of creamer. Previous research examined the effect of coffee roast on espresso characteristics and found that lighter roasted coffee had higher preference scores than darker roasts (Boeneke et al., 2007). Similarly, research by Harwood et al. (2020) examined consumer perception toward brewed black coffee and reported consumer groups that were differentiated by preferences for light roast or dark roast coffee. Only one of the three consumer groups identified had preferences for flavors consistent with darker roasts while the other two groups had preferences for moderately intense coffees (balanced flavor, balance between sour and bitter taste) and lower intensity coffees (lower in flavors associated with roasting, lower bitter taste), though sample sizes were not reported for each of these three consumer groups (Harwood et al., 2020). Coffee creamer consumers may be consumers that generally prefer lighter roast coffees. Additional research would be needed to explore this hypothesis. The impact of coffee in elevating or lowering creamer liking scores is important to understand because it may influence the setting of benchmark or threshold liking scores for product developers or sensory scientists. The current study suggests that evaluating creamers in a lighter roast coffee may result in universally higher liking scores compared to a darker roast coffee.

3.3.3 Limitations

The findings from this research are limited by the creamers and coffees chosen for evaluation. Creamers were chosen to represent different types of commercial creamers. All creamers were vanilla flavored, and these results may differ with a different flavor of creamer. Only two types of coffee were evaluated during phase II of this study–commercial coffees characterized as light roast or dark roast. Coffee is a very complex product whose flavor profile can be influenced by a variety of factors, in addition to roast, such as origin, variety, composition, storage, processing, brewing parameters, etc. (Andueza et al., 2007). Furthermore, research on drivers of liking for cold brew coffee revealed that the attributes that drive liking for cold brew coffee differ from traditional hot coffee (McCain-Keefer et al., 2020). The current study suggests that the type of coffee may have some impact on creamer evaluation, but further research should explore the interactions between creamer product characteristics and different types of coffee in greater detail.

3.4 Conclusion

When executing a sensory evaluation of creamers, it is important to present creamers in the context that consumers use them–in a cup of coffee. However, our study suggests that the method of serving creamer does not have a significant impact on the sensory results or the actionable insights, apart from perhaps the appearance of the creamer in coffee. The fixed amount serving method optimizes for logistical consumer testing considerations in most scenarios. For the type of coffee used to evaluate coffee creamers, overall liking was not affected by the coffee used. However, the type of coffee did have an effect on creamer liking as light roast coffee was preferred over dark roast coffee by the consumers in this study. In addition, the impact of coffee on some of the attributes such as appearance and aroma in coffee and overall flavor liking did have a significant effect on creamer evaluation. Creamer manufacturers and product developers should keep this item in mind during product development and the creamer evaluation process, as creamer formulations optimized based on one type of coffee may be perceived differently when consumers use them in different types of coffee.

Acknowledgments

Funding provided in part by Chobani (New Berlin, NY).

References

Andueza, S., Vila, M. A., De Peña, M. P., & Cid, C. (2007). Influence of coffee/water ratio on the final quality of espresso coffee. *Journal of the Science of Food and Agriculture*, *87*(4), 586–592. https://doi.org/10.1002/jsfa.2720

Boeneke, C. A., Mcgregor, J. U., & Aryana, K. J. (2007). A dairy-based espresso beverage manufactured using three different coffee bean roasts. *Journal of Food Processing and Preservation*, *31*(1), 1–12. https://doi.org/10.1111/j.1745-4549.2007.00103.x

Carcea, M., Danesi, I., De Gara, L., Diretto, G., Fanali, C., Raffo, A., Sinesio, F., Della Posta, S., Frusciante, S., Moneta, E., Narducci, V., Peparaio, M., Saggia Civitelli, E., & Turfani, V. (2023). Chemical composition and sensory profile of the Italian espresso coffee powder and beverage under different roasting conditions. *European Food Research and Technology*, *249*(5), 1287–1301. https://doi.org/10.1007/s00217-023-04213-0

Cherdchu, P., & Chambers, E. (2014). Effect of carriers on descriptive sensory characteristics: A case study with soy sauce. *Journal of Sensory Studies*, *29*(4), 272–284. https://doi.org/10.1111/joss.12100

Cotter, A. R., Batali, M. E., Ristenpart, W. D., & Guinard, J. X. (2021). Consumer preferences for black coffee are spread over a wide range of brew strengths and extraction yields. *Journal of Food Science*, *86*(1), 194–205. https://doi.org/10.1111/1750-3841.15561

Fibrianto, K., & Maharani, Y. (2021). The effect of different non-dairy creamer on ready-To-drink milk coffee. *IOP Conference Series: Earth and Environmental Science*, *782*(3). https://doi.org/10.1088/1755-1315/782/3/032083

Garvey, E. C., Sander, T., O'callaghan, T. F., Drake, M., Fox, S., O'sullivan, M. G., Kerry, J. P.,

& Kilcawley, K. N. (2020). A cross-cultural evaluation of liking and perception of salted butter produced from different feed systems. *Foods*, *9*(12). https://doi.org/10.3390/foods9121767

Harwood, W. S., McLean, K. G., Ennis, J. M., Ennis, D. M., & Drake, M. A. (2020). Comparison of preference mapping with projective mapping for characterizing consumer perception of brewed black coffees. *Journal of Sensory Studies*, *35*(3). https://doi.org/10.1111/joss.12563

Jinjarak, S., Olabi, A., Jiménez-Flores, R., & Walker, J. H. (2006). Sensory, functional, and analytical comparisons of whey butter with other butters. *Journal of Dairy Science*, *89*(7), 2428–2440. https://doi.org/10.3168/jds.S0022-0302(06)72316-5

Keefer, H. R. M., Harwood, W. S., Ennis, D., & Drake, M. A. (2021). The effect of carrier on consumer liking of soy sauce. *Journal of Sensory Studies*, *36*(6). https://doi.org/10.1111/joss.12700

Komes, D., Bušić, A., Vojvodić, A., Belščak-Cvitanović, A., & Hruškar, M. (2015). Antioxidative potential of different coffee substitute brews affected by milk addition. *European Food Research and Technology*, *241*(1), 115–125. https://doi.org/10.1007/s00217-015-2440-z

Krause, A. J., Lopetcharat, K., & Drake, M. A. (2007). Identification of the characteristics that drive consumer liking of butter. *Journal of Dairy Science*, *90*(5), 2091–2102. https://doi.org/10.3168/jds.2006-823

Lawless, H., & Heymann, H. (2010). *Sensory Evaluation of Food Principles and Practices* (D. Heldman, Ed.; Second). Springer.

Lumivero. (2024). XLSTAT statistical and data analysis solution. https://www.xlstat.com/en.

Mahmud, M. M. C., Shellie, R. A., Mohebbi, M., Mahato, D. K., & Keast, R. (2021). The effect of fat and coffee concentration on the consumer acceptance of iced-coffee beverages. *Journal of Food Science*, *86*(11), 5004–5015. https://doi.org/10.1111/1750-3841.15932

McCain-Keefer, H. R., Meals, S., & Drake, M. A. (2020). The sensory properties and consumer acceptance of cold brew coffee. *Journal of Sensory Studies*, *35*(6). **https://doi.org/10.1111/joss.12604**

Olsen, K. (2024). Coffee and RTD Coffee – US – 2024. Mintel. **https://www.mintel.com/**

Parat-Wilhelms, M., Denker, M., Borcherding, K., Hoffmann, W., Luger, A., & Steinhart, H. (2005). Influence of defined milk products on the flavour of white coffee beverages using static headspace gas chromatography-mass spectrometry/olfactometry and sensory analysis. *European Food Research and Technology*, *221*(3–4), 265–273. https://doi.org/10.1007/s00217-005-1152-1

Pereira, L. L., Guarçoni, R. C., da Luz, J. M. R., de Oliveira, A. C., Moreli, A. P., Filete, C. A., Paiva, G. de, Debona, D. G., Gomes, W. dos S., Cardoso, W. S., Berilli, S. da S., & Oliveira, E. C. da S. (2023). Impacts of brewing methods on sensory perception and organoleptic compounds of coffee. *Food Chemistry Advances*, *2*. https://doi.org/10.1016/j.focha.2023.100185

Rashidinejad, A., Tarhan, O., Rezaei, A., Capanoglu, E., Boostani, S., Khoshnoudi-Nia, S., Samborska, K., Garavand, F., Shaddel, R., Akbari-Alavijeh, S., & Jafari, S. M. (2022). Addition of milk to coffee beverages; the effect on functional, nutritional, and sensorial properties. In *Critical Reviews in Food Science and Nutrition* (Vol. 62, Issue 22, pp. 6132–6152). Taylor and Francis Ltd. https://doi.org/10.1080/10408398.2021.1897516

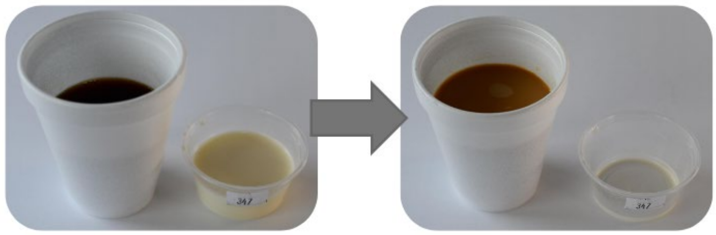
Shurtleff, W., & Aoyagi, A. (2013). *History of non-dairy whip topping, coffee creamer, cottage* *cheese, and icing/frosting (with and without soy)(1900-2013): Extensively annotated bibliography and sourcebook*. Soyinfo Center.

van Eck, A., Fogliano, V., Galindo-Cuspinera, V., Scholten, E., & Stieger, M. (2019). Adding condiments to foods: How does static and dynamic sensory perception change when bread and carrots are consumed with mayonnaise? *Food Quality and Preference*, *73*, 154–170. https://doi.org/10.1016/j.foodqual.2018.11.013

<table>
<tr><th colspan="3">Phase I – Impact of Method</th><th colspan="2">Phase II – Impact of Coffee</th></tr>
<tr><td colspan="3">• Coffee – 90 mL medium roast coffee per creamer sample
• <u>Serving method</u> – three methods
• Design – four samples, evaluated by one method per day, with one-week wash-out period between methods, different consumers</td><td colspan="2">• <u>Coffee</u> – two samples (light roast, dark roast), 90 mL each coffee per creamer
• Serving method – fixed amount of creamer
• Design – eight samples (4 creamers x 2 coffees) randomized over 2 days, same consumers</td></tr>
<tr><td>Week 1 Method:
Fixed amount
(n=127)</td><td>Week 2 Method:
Free pour
(n=120)</td><td>Week 3 Method:
Warm-up
(n=122)</td><td>Coffee:
Light roast
(n=134)</td><td>Coffee:
Dark roast
(n=134)</td></tr>
</table>

Figure 3.1 Experimental overview

Method – Fixed amount

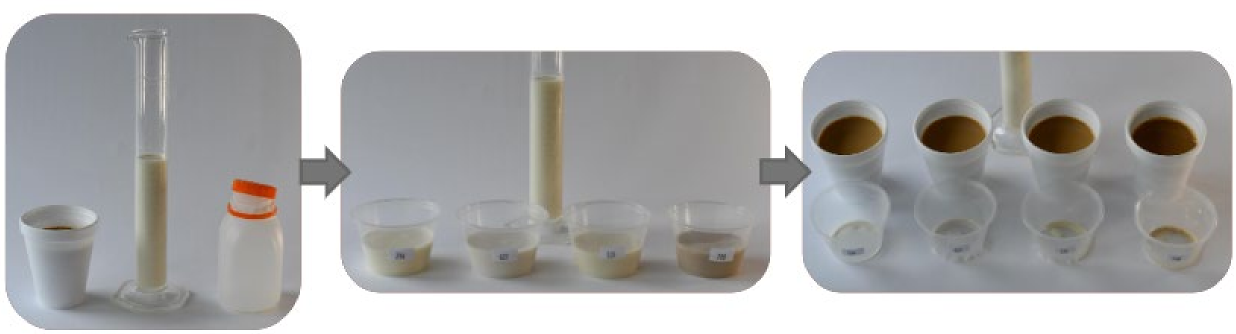


Creamer (30 mL) was pre-portioned for each of the coffees. Panelists were instructed to add the entire amount of each creamer to their coffee.

Method – Free pour

An excess amount of creamer (120 mL) was pre-portioned for each coffee. Panelists were instructed to add creamer according to their preference, adding as much or as little as they like. The amount of creamer added could be the same or different for each coffee, based on preference.

Method – Warm-up



A warm-up coffee with creamer was served with an excess amount of creamer (100 mL). Panelists added creamer according to their preference. The amount of creamer used was calculated based on remaining warm-up sample (initial volume – volume left over). This amount was portioned for all creamers. Panelists were instructed to add the entire amount of each sample (portioned based on warm-up volume) to their coffee.

Figure 3.2 Creamer serving methods. All three methods were compared during phase I of the study. The fixed amount method was used during phase II of the study.

Table 3.1 Analysis of variance of the effect of creamer, method, and their interaction on creamer attributes (phase I).

Attribute	Model		Creamer		Method		Creamer*Method	
	F-value	P-value	F-value	P-value	F-value	P-value	F-value	P-value
Appearance Liking Creamer	5.8	**<0.0001**	314.8	**<0.0001**	2.4	0.091	0.2	0.988
Aroma Liking Creamer	2.5	**<0.0001**	24.2	**<0.0001**	1.9	0.148	2.1	0.055
Appearance Liking in Coffee	3.1	**<0.0001**	24.2	**<0.0001**	3.7	**0.025**	3.4	**0.002**
Aroma Liking in Coffee	2.3	**<0.0001**	16.5	**<0.0001**	0.7	0.476	2.2	**0.045**
Overall Liking	1.7	**<0.0001**	13.0	**<0.0001**	0.5	0.629	0.7	0.636
Sweetness Liking	1.8	**<0.0001**	4.6	**0.003**	0.1	0.928	0.9	0.480
Bitterness Liking	2.0	**<0.0001**	5.9	**0.001**	0.0	0.988	0.4	0.902
Overall Flavor Liking	1.6	**<0.0001**	12.5	**<0.0001**	0.4	0.690	1.5	0.191
Vanilla Flavor Liking	1.5	**<0.0001**	9.2	**<0.0001**	1.5	0.235	1.4	0.196
Mouthfeel Liking	2.3	**<0.0001**	7.8	**<0.0001**	0.6	0.531	0.8	0.554
Creaminess Liking	2.2	**<0.0001**	5.2	**0.001**	0.1	0.870	1.1	0.373

Data represents n=369 consumers nested within each method—Fixed amount (n=127), Free pour (n=120), Warm-up (n=122). F-values and P-values are presented for each attribute and effect. Values are bolded if statistically significant ($p<0.05$).

Table 3.2 Liking means for creamer*method interactions in phase I.

Appearance Liking in Coffee

	Creamer A	Creamer B	Creamer C	Creamer D
Free Pour	7.3a	7.1abc	7.3ab	6.3ef
Fixed Amount	7.0cd	6.7de	6.7e	6.6ef
Warm-up	7.1abc	7.0bcd	7.2abc	6.4ef

Aroma Liking in Coffee

	Creamer A	Creamer B	Creamer C	Creamer D
Free Pour	6.3abcd	6.1bcde	6.4ab	6.0cde
Fixed Amount	6.4abc	6.0de	6.3abcd	5.8e
Warm-up	6.2bcd	6.1bcde	6.6a	5.4f

Data represents n=369 consumers nested within each method—Fixed amount (n=127), Free pour (n=120), Warm-up (n=122). Mean values are reported for each attribute where a significant creamer*method interaction was found ($p<0.05$). Liking was scored on a 9-point scale where 1=dislike extremely, 9=like extremely. Different letters following means within a matrix signify significant differences by ANOVA with Fisher's LSD post hoc comparison ($\alpha=0.05$).

Figure 3.3 PCA biplot for creamers (A, B, C, D) evaluated using each serving method (Fixed method, n=127; Warm-up (WU) method, n=122; Free Pour method, n=120) during phase I. Creamers were evaluated using the attributes shown in red by consumers.

Table 3.3 Mean values for creamer liking averaged across serving method in phase I.

Attribute	Creamer A	Creamer B	Creamer C	Creamer D	Creamer
Appearance Liking Creamer	7.0a	6.8b	7.0a	4.5c	**<0.0001**
Aroma Liking Creamer	6.8a	6.2b	6.8a	6.1b	**<0.0001**
Appearance Liking in Coffee	7.1a	7.0a	7.0a	6.5b	**<0.0001**
Aroma Liking in Coffee	6.3a	6.1b	6.4a	5.7c	**<0.0001**
Overall Liking	5.7b	6.4a	6.4a	5.9b	**<0.0001**
Sweetness Liking	5.7c	6.1ab	6.1a	5.9bc	**0.003**
Bitterness Liking	5.4b	5.8a	5.8a	5.6b	**0.001**
Overall Flavor Liking	5.6b	6.3a	6.3a	5.8b	**<0.0001**
Vanilla Flavor Liking	5.5b	6.1a	6.1a	5.6b	**<0.0001**
Mouthfeel Liking	6.3b	6.7a	6.6a	6.3b	**<0.0001**
Creaminess Liking	6.1bc	6.5a	6.3ab	6.1c	**0.001**

Data represents n=369 consumers nested within each method—Fixed amount (n=127), Free pour (n=120), Warm-up (n=122). Mean ratings are reported. Liking was scored on a 9-point scale where 1=dislike extremely, 9=like extremely. Different letters following means across a row signify significant differences by ANOVA with Fisher's LSD post hoc comparison (α=0.05). Statistical lettering is only reported for attributes where creamer was a significant main effect ($p<0.05$), indicated in bold.

Table 3.4 Analysis of variance of the effect of creamer, coffee, and their interaction on creamer attributes (phase II).

Attribute	Model		Creamer		Coffee		Creamer*Coffee	
	F-value	P-value	F-value	P-value	F-value	P-value	F-value	P-value
Appearance Liking Creamer	10.5	**<0.0001**	293.3	**<0.0001**	0.5	0.473	0.8	0.503
Aroma Liking Creamer	3.7	**<0.0001**	29.1	**<0.0001**	0.7	0.417	1.4	0.233
Appearance Liking in Coffee	4.9	**<0.0001**	16.1	**<0.0001**	5.6	**0.018**	7.9	**<0.0001**
Aroma Liking in Coffee	5.1	**<0.0001**	32.2	**<0.0001**	47.6	**<0.0001**	7.5	**<0.0001**
Overall Liking	3.7	**<0.0001**	9.6	**<0.0001**	50.1	**<0.0001**	2.2	0.086
Sweetness Liking	3.8	**<0.0001**	6.2	**0.0004**	41.3	**<0.0001**	1.4	0.240
Bitterness Liking	3.2	**<0.0001**	7.4	**<0.0001**	74.5	**<0.0001**	2.4	0.069
Overall Flavor Liking	3.5	**<0.0001**	11.8	**<0.0001**	61.3	**<0.0001**	3.7	**0.012**
Vanilla Flavor Liking	3.4	**<0.0001**	14.2	**<0.0001**	66.1	**<0.0001**	3.0	**0.030**
Mouthfeel Liking	5.0	**<0.0001**	13.0	**<0.0001**	27.3	**<0.0001**	3.0	**0.031**
Creaminess Liking	4.3	**<0.0001**	18.9	**<0.0001**	37.8	**<0.0001**	4.5	**0.004**

Data represents n=134 consumers. F-values and P-values are presented for each attribute and effect. Values are bolded if statistically significant ($p<0.05$).

Table 3.5 Mean values for creamer*coffee interactions in phase II.

Appearance Liking in Coffee

	Creamer A	Creamer B	Creamer C	Creamer D
Light Roast	6.7ab	6.8ab	6.9a	6.6b
Dark Roast	6.9ab	6.6ab	6.9a	5.9c

Aroma Liking in Coffee

	Creamer A	Creamer B	Creamer C	Creamer D
Light Roast	6.5a	6.2ab	6.5a	5.9b
Dark Roast	6.4a	5.2c	6.1ab	4.9c

Overall Flavor Liking

	Creamer A	Creamer B	Creamer C	Creamer D
Light Roast	6.0bc	6.7a	6.3ab	5.9bcd
Dark Roast	5.7cd	5.6cd	5.5d	4.5e

Vanilla Flavor Liking

	Creamer A	Creamer B	Creamer C	Creamer D
Light Roast	6.0ab	6.3a	6.2a	5.6cd
Dark Roast	5.6bc	5.2d	5.3cd	4.3e

Mouthfeel Liking

	Creamer A	Creamer B	Creamer C	Creamer D
Light Roast	6.5ab	6.8a	6.5ab	6.3bc
Dark Roast	6.4bc	6.3bc	6.1c	5.4d

Creaminess Liking

	Creamer A	Creamer B	Creamer C	Creamer D
Light Roast	6.4a	6.6a	6.3ab	6.0bc
Dark Roast	6.3ab	6.1bc	5.8c	5.0d

Data represents n=134 consumers. Mean ratings matrices are reported for each attribute where a significant creamer*method interaction was found ($p<0.05$). Liking was scored on a 9-point scale where 1=dislike extremely, 9=like extremely. Different letters following means within a matrix signify significant differences by ANOVA with Fisher's LSD post hoc comparison ($\alpha=0.05$).

Table 3.6 Mean liking scores of creamers averaged across coffee in phase II.

Attribute	Creamer A	Creamer B	Creamer C	Creamer D	Creamer
Appearance Liking Creamer	7.2a	6.8b	7.2a	4.4c	**<0.0001**
Aroma Liking Creamer	6.9a	6.2b	7.0a	6.0b	**<0.0001**
Appearance Liking in Coffee	6.8a	6.7a	6.9a	6.2b	**<0.0001**
Aroma Liking in Coffee	6.4a	5.7b	6.3a	5.4c	**<0.0001**
Overall Liking	6.0a	6.1a	6.0a	5.3b	**<0.0001**
Sweetness Liking	5.9a	6.0a	5.9a	5.4b	**0.0004**
Bitterness Liking	5.5a	5.4a	5.3a	4.8b	**<0.0001**
Overall Flavor Liking	5.9a	6.1a	5.9a	5.2b	**<0.0001**
Vanilla Flavor Liking	5.8a	5.7a	5.8a	5.0b	**<0.0001**
Mouthfeel Liking	6.5ab	6.5a	6.3b	5.8c	**<0.0001**
Creaminess Liking	6.4a	6.3ab	6.1b	5.5c	**<0.0001**

Data represents n=134 consumers. Mean ratings are reported. Liking was scored on a 9-point scale where 1=dislike extremely, 9=like extremely. Different letters following means across a row signify significant differences by ANOVA with Fisher's LSD post hoc comparison (α=0.05). Statistical lettering is only reported for attributes where creamer was a significant main effect ($p<0.05$), indicated in bold.

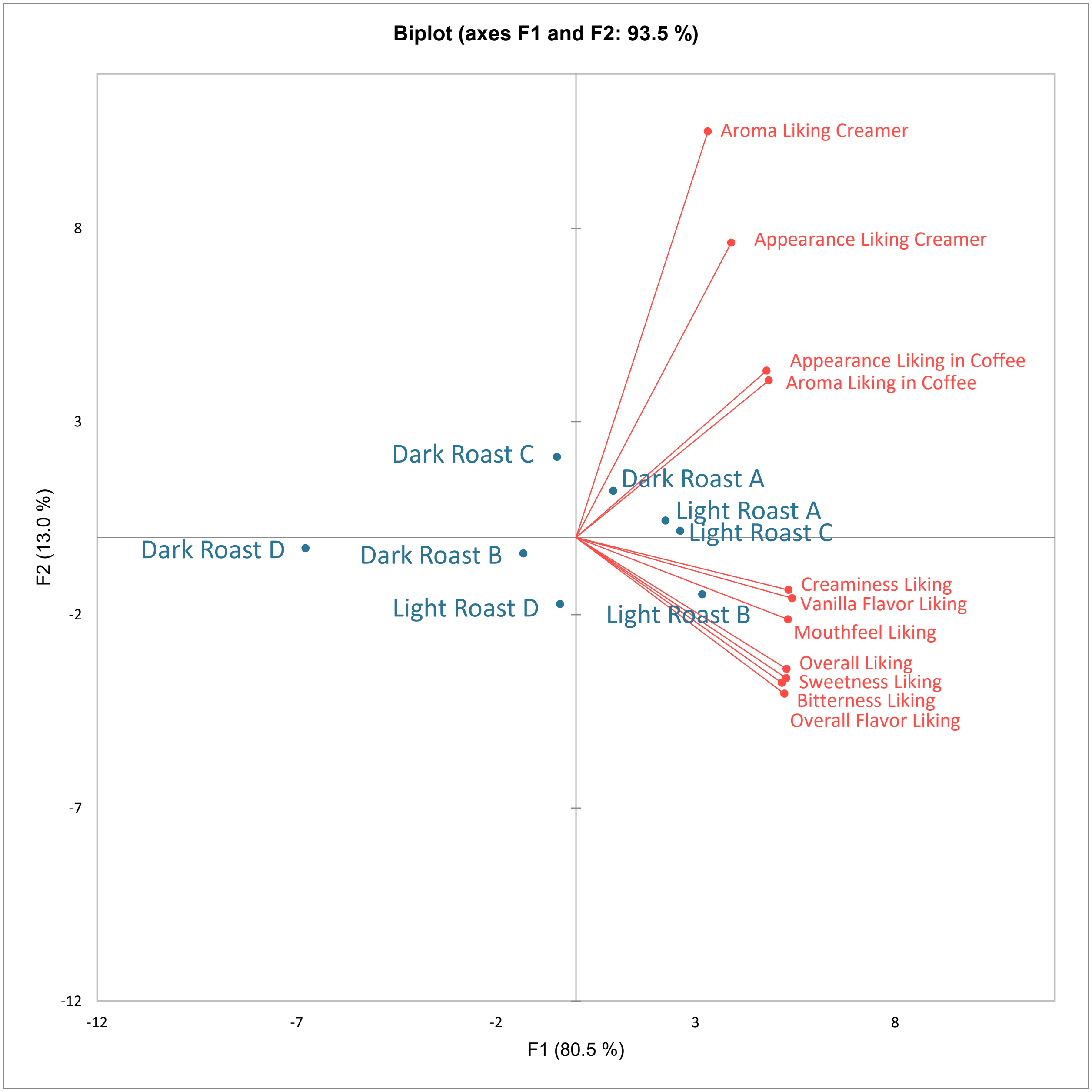


Figure 3.4 PCA biplot for creamers (A, B, C, D) evaluated using in each type of coffee (light roast, dark roast) during phase II (n=134). Creamers were evaluated using the attributes shown in red by consumers (n=134).

Table 3.7 Mean liking scores of creamers averaged across coffee roast in phase II.

	Light Roast	Dark Roast	Coffee
Appearance Liking Creamer	6.4	6.4	0.473
Aroma Liking Creamer	6.6	6.5	0.417
Appearance Liking in Coffee	6.7a	6.6b	**0.018**
Aroma Liking in Coffee	6.3a	5.7b	**<0.0001**
Overall Liking	6.2a	5.4b	**<0.0001**
Sweetness Liking	6.2a	5.5b	**<0.0001**
Bitterness Liking	5.7a	4.8b	**<0.0001**
Overall Flavor Liking	6.2a	5.3b	**<0.0001**
Vanilla Flavor Liking	6.0a	5.1b	**<0.0001**
Mouthfeel Liking	6.5a	6.1b	**<0.0001**
Creaminess Liking	6.3a	5.8b	**<0.0001**

Data represents n=134 consumers. Mean ratings are reported. Liking was scored on a 9-point scale where 1=dislike extremely, 9=like extremely. Different letters following means across a row signify significant differences by ANOVA with Fisher's LSD post hoc comparison ($\alpha=0.05$). Statistical lettering is only reported for attributes where coffee was a significant main effect ($p<0.05$), indicated in bold.

CHAPTER 4:

Comparison of Gins Using Temporal Dominance of Sensations (TDS), Temporal Check-All-That-Apply (TCATA), and Temporal Ranking (TR)

Comparison of gins using temporal dominance of sensations (TDS), temporal check-all-that-apply (TCATA), and temporal ranking (TR)

D. Rovai[1], H. McCain Keefer[1], J. C. Castura[2], M. A. Drake[1]

[1] Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, North Carolina, USA

[2] Compusense, Inc., Guelph, Ontario, Canada

Choice of journal/topic: Journal of Sensory Studies, submitted

Abstract

Gin is a distilled spirit characterized by juniper flavor, often accompanied by other botanical flavors. The objective of this research was to use gin as a case study for comparing temporal dominance of sensations (TDS), temporal check-all-that-apply (TCATA), and temporal ranking (TR). Trained panelists used each of these methods to evaluate the temporal profile of six gins (standardized to 30% ABV). Differences in temporality were visualized by difference graphs for pairwise gin comparisons across methods. TDS differentiated gins based on dominant flavor attributes but did not fully describe the complexity of the gins. TCATA was effective in capturing differences in subtle flavor attributes but did not discriminate samples based on juniper flavor. TR differentiated gins based on differences in juniper flavor intensity and documented additional differences in complex gins but did not discriminate subtle flavor attributes. Researchers should consider these tradeoffs when selecting a temporal method.

Practical Application

Temporal evaluations are a tool to document attributes in products over the course of evaluation. These methods can be used to help companies and researchers better understand complex products such as gin. Understanding the usefulness and limitations of each of the temporal methods described in this research can help scientists make decisions about which method(s) to use based on the product, study objective(s), and resources. Insights from these studies can aid product development and optimization opportunities.

Key Words: TDS, TCATA, temporal ranking, gin

4.1 Introduction

Temporal sensory methods offer a dynamic evaluation of food products by taking into account how a product changes over the course of evaluation (Cliff & Heymann, 1993; Lawless & Heymann, 2010). New temporal methods and variations of existing methods have emerged to aid in understanding the sensory characteristics of complex food products (Keefer et al., 2023). Popular temporal methods that evaluate multiple attributes simultaneously include Temporal Dominance of Sensations (TDS) and Temporal Check-All-That-Apply (TCATA). TDS documents which attribute panelists perceive as dominant and how that selection changes over the course of evaluation (Pineau et al., 2009), while TCATA documents all attributes present throughout the course of evaluation (Castura et al., 2016). Each method provides different insights into product characteristics (change in dominant attribute vs change in all attributes over time) and offer tradeoffs (time, amount of training, number of attributes, cognitive difficulty, etc.) (Keefer et al., 2023). Temporal Ranking (TR) is a newer method in which panelists continuously rank the top three prominent attributes during evaluation (Keefer et al., 2022). TR was found to be more discriminating than TCATA in a study investigating sensory properties of protein beverages but the usefulness of TR has not been validated in other food matrices.

Alcoholic beverages are complex food matrices with interesting temporal profiles. Fermenting simple sugars or broken down starches (wheat, barley, rye, etc.) in anaerobic conditions yield ethanol, carbon dioxide, and other flavor-active compounds such as carbonyls (aldehydes/ketones), higher/fusel alcohols, esters, organic acids, and sulfur compounds among others (Nykanen, 1986; Cacho & Lopez, 2005; Antalick et al., 2010; Olaniran et al., 2017). Alcoholic spirits have inherent complexity due to the fermentation process, as well as the burning/numbing sensations associated with alcohol. Temporal methods have been used to evaluate a variety of alcoholic beverages including beer, wine (red, white, sparkling, Chinese rice wine), brandy, rum, Baijiu, cocktails, and carbonated alcoholic beverages (Meillon et al., 2009; Sokolowsky & Fischer, 2012; Baker et al., 2016; Fiches et al., 2016; Vidal et al., 2016; McMahon et al., 2017; Carla Correa Simioni et al., 2018; Ramsey et al., 2018; Kemp et al., 2019; Poveromo & Hopfer, 2019; Harwood et al., 2020; He et al., 2021; Medel-Maraboli et al., 2021; Wang et al., 2021; Pierguidi et al., 2021; Yu et al., 2023; Stoffel et al., 2023a; Stoffel et al., 2023b; Wakihira et al., 2023). However, only a handful of studies have used alcoholic beverages to compare TDS to TCATA (Medel-Maraboli et al., 2021; Yu et al., 2023) and no studies have done so using distilled spirits as the matrix.

Gin is a distilled spirit containing no less than 40% alcohol by volume which is characterized by the addition of juniper berries during the distillation process (27 CFR § 5.144 Gin). During or after distillation, various botanicals may also be added to gin, such as coriander seeds, citrus peel, licorice, and angelica root (Riu-Aumatell, 2012). The addition of these botanicals results in a wide range of aroma and flavor compounds in commercial gins. The sensory complexity of gin makes it an interesting candidate for sensory evaluation, but to the authors' knowledge the temporal profile of gin has not yet been investigated in any research application. This research evaluates the temporal profile of commercial gins and compares three temporal methods: TDS, TCATA, and TR.

4.2 Methods

4.2.1 Overview

Commercial gins were first evaluated by descriptive analysis (DA). Following DA, a subset of gins was selected and evaluated by the same trained panelists using the three different temporal methods. A graphic overview of the study is illustrated in Figure 4.1. This research was reviewed and approved by the North Carolina State University Institutional Review Board for Human Subjects (IRB 25149).

4.2.2 Sample Selection

Ten commercial gins (40% - 48% ABV) were selected for descriptive analysis (DA). The selected gins represented a range of sensory attributes, price points and market share. Based on DA, six gins were selected for further evaluation with three temporal methods (TDS, TCATA, TR). These six gins were selected to represent a range of sensory attributes and intensities (Figure 4.2).

4.2.3 Sample Preparation

Gins were standardized to 30% ABV using deionized water in one liter Pyrex glass bottles No. 1395 to prevent sensory fatigue, reduce odor suppression effects of ethanol, and enhance flavor release (McDonnel et al., 2001; Ickes & Cadwallader, 2018; Harwood et al., 2020). Gins were served in 163 mL Libbey Embassy brandy glasses covered with a watch glass and labeled with three-digit blinding codes. For DA, panelists were served 60 mL of each sample to allow for retasting. For temporal evaluations, 12 mL of each sample (approximately one mouthful), were presented to panelists (Harwood et al., 2020; Keefer et al., 2022). Prepared samples were held at 21 °C for at least one hour prior to evaluation to establish headspace equilibrium.

4.2.4 Panelists and Training

Seven trained and experienced DA panelists—each with at least 40 hours of prior DA experience with gin or other spirits and 25 hours of prior experience with temporal methods—were selected to evaluate the gins in this study. Panelists were trained in DA using references and a 0 to 15 point universal intensity scale consistent with the Spectrum method (Meilgaard et al., 2016). Six of the seven DA panelists proceeded with temporal evaluation of the gins.

Panel protocols, training, and lexicon development followed ASTM guidelines (ASTM, 1981; ASTM, 2011). Table 4.1 outlines the lexicon and references used for panel training. References consisted of spices in 30% ethanol solution (Fisher Scientific, Waltham, MA), natural references, and chemical references. Reference solutions (30% ethanol) were prepared according to Riu-Aumatell et al. (2008). Select spices were ground in a pestle and mortar until visibly broken up and extracted in ethanol over 24 h. Ethanol extracts were then filtered and diluted to 30% ABV with deionized water. All chemical references were obtained from Sigma-Aldrich (Rockville, MD) and prepared in a sniff jar with filter paper. Natural references were prepared by putting fresh herbs (mint, cilantro, parsley) in a lidded 118 mL souffle cup 1 h before training.

In addition to previous experience with DA, spirits, and temporal methods, panelists underwent an additional nine 45-min sessions of training/calibration using commercial gins and references. Analysis of variance of preliminary data confirmed that panelists could consistently differentiate the gins using the identified lexicon. Prior to temporal evaluations, the same trained panelists took part in six 30-min practice sessions for each temporal method to familiarize themselves with the procedures, samples, and interfaces for each method using a modified list of the same sensory attributes used for DA (described below).

4.2.5 Descriptive Analysis

Descriptive analysis (DA) was performed on the 10 commercial gins in triplicate. Panelists evaluated orthonasal aroma, flavor, basic taste, and aftertaste intensities using a 0-15 point universal intensity scale (Spectrum descriptive analysis method). Orthonasal aroma was evaluated by lifting the watch glass slightly to smell the headspace. For flavor and basic taste, panelists were instructed to take a mouthful of sample at a time and swish in the mouth before expectorating. Samples were evaluated across nine sessions, with three to four samples evaluated per session. Compusense Cloud version 23.0.3 (Compusense, Guelph, Canada) was used for data collection. A four-minute rest was enforced between each sample and panelists rinsed their palates with bottled purified water (Pure Life, 8 oz).

4.2.6 Temporal Evaluation

The six gins selected based on DA profiling were evaluated using the three temporal methods: TDS, TCATA, and TR. A lexicon modified from the DA panel was selected for temporal evaluations. The lexicon consisted of 11 attributes: alcohol burn/numbing, anise/fennel, bitter taste, citrus, coriander/floral, fusel, grassy/herbaceous, juniper/pine, mint, sweet aromatic, and sweet taste. This lexicon included the key attributes from DA without presenting too many attribute options as 10 - 12 attributes have been cited as the upper limit for TDS (Di Monaco et al., 2014; Pineau et al., 2012; Schlich & Pineau, 2017) and up to 15 attributes have been shown to be acceptable for TCATA (Jaeger et al., 2018). Data were collected in Compusense Cloud version 23.0.3 (Compusense Inc., Guelph, ON, Canada). Panelists performed evaluations across nine sessions (3 reps each for TR, TDS, TCATA) where all permutations were presented blocked by method (Table 4.2). In each session, panelists evaluated all six samples with an enforced three-minute rest between samples. In addition, gins were presented to panelists using a Williams Latin square design (Williams, 1952) to minimize sample position bias.

For each temporal method, samples were evaluated over the course of 155 s. Samples were evaluated orthonasally (0-20 s), followed by a 15-s break before in-mouth evaluation (135 s). In-mouth evaluation began when the panelist took the mouthful of gin (12 mL), followed by expectoration after 15 s and aftertaste evaluation for the duration of the in-mouth portion (Harwood et al., 2020). Attribute order was randomized using a Williams Latin square design across the entire panel to reduce attribute position bias, but consistent within individual panelists across all sessions (Di Monaco et al., 2014; Meyners & Castura, 2016; Pineau et al., 2012).

4.2.7 Data Analysis

4.2.7.1 Descriptive Analysis (DA)

Analysis of variance with Fisher's least significant difference post hoc test ($\alpha = 0.05$) was used to determine significant differences in attribute intensities among the gins profiled by DA. A principal component analysis (PCA) biplot was constructed using Pearson correlation to visualize differences between gins.

4.2.7.2 Temporal Evaluations

During TDS evaluation, panelists continuously check the attribute that most catches their attention. During TCATA evaluation, panelists continuously check all attributes that are present (and uncheck attributes that aren't present). Data for TDS and TCATA were reported as proportions for each attribute (TDS = Dominance Proportion, TCATA = Citation Proportion), with 0 indicating that the attribute was not selected, and 1 indicating the attribute was selected at that time by all panelists over all reps. During TR evaluation, panelists ranked up to three attributes that are the most relevant and update the attributes continuously. Graphically, data for TR were reported as modified borda counts, where ranked 1st attributes were coded to = 3, second = 2, third = 1, and all unranked attributes = 0 (Keefer et al., 2022). For the determination of product difference curves, ranks were not modified, and attributes in TR that were not ranked were "tied for last" and assigned the average score of the remaining ranks (Keefer et al., 2022). Smoothed product curves and difference curves were constructed in R version 4.3.2 (R Core

Team, 2023) using the R package “tempR” (Castura, 2022). Significant differences in the difference curves were determined via McNemar’s exact test ($\alpha = 0.05$).

4.2.7.3 Panel Consistency

Panel consistency during temporal evaluations was determined via the intraclass correlation coefficient (ICC). The ICC is the ratio of between-groups variance to total variance. ICC calculations in this study followed the approach for obtaining ICC from ranking data used by Bi & Kuesten (2012) but applied in a new way. For TR, results were treated as dependent ranking data where up to three attributes were ranked and the rest tied for last. For TDS and TCATA, the results were binary data (checked or not checked). TR data were analyzed using Friedman’s test (Friedman, 1937), while TDS and TCATA data were analyzed using Cochran’s Q (Cochran, 1950). For binary data, results from Friedman’s test and Cochran’s Q test coincide (Koch et al., 1980). ICC was reported graphically, showing the panel average for each method (across three reps) over the course of evaluation, and then plotted in R version 4.3.2 (R Core Team, 2023). There are currently no agreed upon standard values for acceptable reliability using ICC (Koo & Li, 2016), but a higher ICC reflects higher panel performance. When evaluating the consistency of the temporal results, we interpreted the ICC to having the following interpretations: 0 to 0.39 bad, 0.4 to 0.59 fair, 0.6 to 0.75 good, and 0.75 to 1 excellent (Cicchetti, 1994).

4.3 Results and Discussion

4.3.1 Descriptive Analysis (DA)

DA results are summarized in Figure 4.2. The PCA biplot visualizes the first two principal components which account for 54.7% of the variability and was the basis for choosing the six gins selected for temporal evaluations. Gin A was characterized by anise/fennel, minty, and sweet taste. Gins B and C were characterized by terpene/juniper, bitter taste, fusel, alcohol burn, and numbing, with gin C also having herbal flavor. Gins D and E were characterized by high aroma intensity with citrus and coriander/floral notes, with gin E defined more by coriander/floral flavor and gin D more by citrus flavor. Gin F was characterized by herbal and cilantro flavors.

4.3.2 Temporal Profile of Gins

For each temporal method, the data were used to construct one plot per gin. Plots show TDS curves (Figure 4.3), TCATA curves (Figure 4.4), and TR curves (Figure 4.5). These figures show the panel average for each attribute over the course of evaluation. Looking across gins, trends were observed where the temporal profile for gin consisted of the prominent flavor attribute(s) detected in the headspace orthonasally, with those flavors also present retronasally during the beginning of in-mouth evaluation. During the beginning of in-mouth evaluation (0-15 s), sweet taste was also prevalent, dropping off sharply after expectoration (15 s). Alcohol burn/numbing became present after expectoration and lingered until the end of evaluation. Bitter taste also became prevalent after expectoration, peaking after 30-60 s in-mouth and lingering in the aftertaste. Juniper and fusel were present in all gins, receiving high citation rates in TCATA for all gins (Figure 4.4) that tended to peak around the time of expectoration before slowly dropping off. Juniper—the characteristic flavor associated with gin (27 CFR § 5.144 Gin)—was present in all gins. However, citations for juniper were low in TR and TDS for gins where numerous other flavors were present, such as gin A and gin E (Figure 4.3, Figure 4.5). Perhaps juniper flavor is expected to be present in all gins and was thus not considered the dominant flavor that "caught panelists' attention" (Le Révérend et al., 2008; Labbe et al., 2009; Lenfant et al., 2009; Pineau et al., 2009; Varela et al., 2018; Hutchings et al., 2022). Other flavor attributes (anise, coriander, citrus, grassy, mint, sweet aromatic) tended to follow a similar trend of peaking near 15 s and dissipating quickly after expectoration (Figure 4.3, Figure 4.4, Figure 4.5), likely due to the volatile nature of these compounds (Dou et al., 2023; Vichi et al., 2005).

In this study, aroma was a good predictor of in-mouth flavors. Othonasal and retronasal smells are qualitatively similar (Lawless & Heymann, 2010), meaning the aromas you smell are typically the same volatile flavors perceived retronasally when the sample is placed in the mouth. However, aroma evaluation provided no insight into the basic tastes or burn experienced during the in-mouth evaluation because sensory perception of basic tastes is triggered through gustation (Lawless & Heymann, 2010).

For all gins, the attributes with the highest citations/borda counts determined by TDS (Figure 4.3), TCATA (Figure 4.4), and TR (Figure 4.5) aligned with the DA intensity scores (Figure 4.2). In addition, the temporal product graphs (Figure 4.3, Figure 4.4, Figure 4.5) provided information about when these attributes were prominent over the course of evaluation. The following section will discuss how these temporal methods compared in terms of differentiating gins based on their temporal profiles.

4.3.3 Comparisons of Temporal Methods

Comparing temporal methods has been likened to "comparing apples to oranges" because each method has different objectives (Meyners, 2020). With this in mind, this section will focus on highlighting the usefulness and shortcomings of each method—rather than choosing which method is "best"—in order to better understand how to evaluate complex food matrices such as gin.

Difference graphs are useful to compare the temporal profile of two products because they highlight attributes that are statistically different ($p<0.05$) between two products and show in which product the attribute was more prominent, and over what time during evaluation it was statistically significant. Results from gin comparisons for all 3 methods resulted in 45 difference graphs (15 pairwise comparisons x 3 methods). All difference graphs were carefully evaluated (results not shown). Three figures (one figure per method – Figure 4.6, TDS; Figure 4.7, TCATA; Figure 4.8, TR) will be used to discuss how temporal methods differ in characterizing gins that are similar (gin B, gin C), gins that are moderately different (gin C, gin F), and gins that are very different (gin B, gin E), as determined by DA (Figure 4.2).

4.3.3.1 TDS vs TCATA

Difference graphs for the same gins were used to compare insights from TDS (Figure 4.6) and TCATA (Figure 4.7). When comparing gins that are similar to each other (gin B, gin C), both methods discriminated the gins by identifying significant attributes. Based on TDS results, juniper/pine dominated the aroma evaluation and early in the in-mouth evaluation more often in gin B than in gin C (Figure 4.6a). TCATA results showed gin C had a higher citation of citrus aroma, coriander/floral aroma, and coriander/floral flavor in the early in-mouth evaluation than gin B (Figure 4.7a). Comparing gins that are moderately different to each other (gin C, gin F), both methods identified higher grassy/herbaceous in the aroma and early flavor perception associated with gin F compared to gin C, though the citation duration was longer when panelists used TCATA (Figure 4.6b, Figure 4.7b). Based on TDS results, gin C had a significantly higher dominance rate for juniper/pine and fusel in the aroma compared to gin F (Figure 4.6b). TCATA results did not show any significant differences in juniper/pine or fusel, but gin C had a higher citation of coriander/floral aroma and flavor in the early in-mouth evaluation than gin F (Figure 4.7b). Finally, comparing gins that are very different from each other (gin B, gin E), both methods identified higher coriander/floral perception in the aroma and early flavor perception associated with gin E compared to gin B (Figure 4.6c, Figure 4.7c). This citation duration for coriander/floral was longer when panelists used TCATA compared to TDS. Based on TDS results, gin B had a significantly higher dominance rate for juniper/pine and fusel in the aroma compared to gin E (Figure 4.6c). TCATA results did not show any differences in citation for juniper/pine or fusel but did show that gin E was perceived as having a higher citation of citrus aroma and higher sweet aromatic, citrus flavor, and sweet taste in the early in-mouth evaluation than gin B (Figure 4.7c).

TDS and TCATA methods both discriminated the gins, but the interpretations of the results from each method would lead to different conclusions about how the gins differed. TCATA better discriminated subtle flavor attributes such as the coriander/floral in gin C and the citrus and sweet aromatic in gin E. These findings are consistent with earlier research showing TCATA product profiles were more detailed than TDS product profiles, perhaps because TCATA provides insight into secondary sensations whereas TDS provides insight only into the dominant attribute (Ares et al., 2015; Esmerino et al., 2017; Nguyen et al., 2018). Additionally, in gins differing in more than two or three attributes (gin B & gin E), TCATA was able to discriminate more attributes than TDS (Figure 4.6c, Figure 4.7c). A similar result was observed when TCATA and TDS were used to evaluate nutritive and nonnutritive sweeteners in water, where TCATA identified secondary attributes that were not characterized by TDS (Reyes et al., 2017).

Although TDS failed to discriminate subtle flavor attributes, it documented differences in juniper. Because juniper was present in all gins, TCATA was unable to discriminate this attribute because it was universally checked for all gins. However, TDS was able to differentiate when the juniper flavor was more dominant in one gin compared to another, regardless of the fact that it was present in both gins. From this perspective, TDS differentiated flavors that were present in all gins, while TCATA discriminated gins on a larger number of subtle flavor attributes.

4.3.3.2 TDS vs TR

Difference graphs for the same gins were used to compare insights from TDS (Figure 4.6) and TR (Figure 4.8). When comparing gins that are similar to each other (gin B, gin C), both methods agreed in finding few differences but disagreed in what those few differences were. Based on TDS results, juniper/pine had a higher dominance rate for gin B during the aroma and beginning of in-mouth evaluation compared to gin C (Figure 4.6a). However, based on TR results, gin B had significantly higher sweet taste during the beginning of in-mouth evaluation and higher bitter taste between 70-90 s than gin C (Figure 4.8a). The comparison of gins that are moderately different (gin C, gin F) showed that gin F was significantly higher in grassy/herbaceous aroma and grassy/herbaceous flavor during the beginning of in-mouth evaluation than gin C by both TDS (Figure 4.6b) and TR (Figure 4.8b). TDS results showed gin C had higher dominance rates for both juniper/pine aroma and fusel aroma than gin F, whereas TR did not detect any significant aroma differences between these gins but showed gin C had significantly higher fusel flavor between 10-20 s. In gins that were very different from each other, more attributes were found to be significantly different between the two gins via TR compared to TDS. Both TDS (Figure 4.6c) and TR (Figure 4.8c) showed gin B had higher juniper/pine and fusel aroma and higher juniper/pine flavor during the beginning of in-mouth evaluation, whereas gin E had higher coriander/floral aroma and flavor during the beginning of in-mouth evaluation. However, TR results showed gin B had higher juniper/pine flavor than gin E for a longer duration in-mouth (until 50 s) compared to TDS (until 10 s). Based on TR results, gin B had higher fusel flavor until 30 s, and higher bitter taste from 60-80 s than gin E. TR results showed gin E had higher sweet aromatic and citrus aroma and citrus flavor during the beginning of in-mouth evaluation than gin B.

The findings in this study suggest that the insights obtained from comparing different gins would differ depending on whether you used TR or TDS. There is not enough data to suggest that one method was more discriminating regarding the pairs of gins that were similar or moderately different, but it was clear looking at the results that TR was a more discriminating method than TDS when comparing two very different gins (Figure 4.6c, Figure 4.8c). The lower discrimination observed by TDS occurs because TDS only allows for the selection of the one most dominant attribute at a time, while TR allows for ranking up to three attributes at a time. The ability to rank multiple attributes in TR allows for the documentation of more attributes simultaneously compared to TDS.

4.3.3.3 TCATA vs TR

Difference graphs for the same gins were used to compare insights from TCATA (Figure 4.7) and TR (Figure 4.8). Comparing gins that are similar to each other (gin B, gin C), TCATA found gin C had higher citation for citrus aroma and coriander/floral flavor and citrus flavor during the beginning of in-mouth evaluation than gin B (Figure 4.7a). TR results showed that gin B had higher sweet taste during the beginning of in-mouth evaluation and bitter taste from 70-90 s than gin C (Figure 4.8a). For the gins that were moderately different, both TCATA (Figure 4.7b) and TR (Figure 4.8b) found gin F had higher grassy/herbaceous aroma and flavor during the beginning of in-mouth evaluation than gin C. Additionally, TR found gin C had higher fusel flavor between 10-20 s, whereas TCATA found gin C had higher coriander/floral aroma and flavor during the beginning of in-mouth evaluation. In gins that were very different (gin B, gin E), TCATA (Figure 4.7c) and TR (Figure 4.8c) both showed gin E had higher coriander/floral and citrus aroma and coriander/floral and citrus flavor during the beginning of in-mouth evaluation. TCATA results showed that gin E had higher sweet aromatic and sweet taste during the beginning of in-mouth evaluation than gin E (Figure 4.7c). TR results showed gin E had higher sweet aromatic aroma than gin B (Figure 4.8c). Additionally, while TCATA did not identify any significant attributes associated with gin B (Figure 4.7c), TR results showed gin B had higher juniper/pine and fusel in the aroma and beginning of in-mouth evaluation, and bitter between 60-70 s than gin E (Figure 4.8c).

When TR was investigated previously, it was found to be more discriminating than TCATA when evaluating the temporal profile of protein beverages (Keefer et al., 2022). The current study did not find overwhelming differences in discriminatory power of these two methods. Because TR allows for ranking of attributes (as opposed to presence/absence with TCATA), it is better able to differentiate attributes that were present in all gins (juniper/pine, fusel) based on differing intensities/ranking compared to other attributes. However, because TR involves ranking of the top three attributes, it is not sensitive to differences in attributes that may rank 4th, 5th, etc. For example, while gin C had subtle notes of coriander and citrus, the attributes ranked 1st, 2nd, and 3rd were the same as gin B for most of the evaluation: juniper, fusel, and some combination of sweet, bitter, and burn (Figure 4.4, Figure 4.5). As a result, coriander and citrus were not differentiated in these gins because, although present, they were rarely a top-three attribute. On the other hand, TCATA does not document whether the attribute is subtle or dominant, but simply reflects whether the attribute was present or not at each time (in any perceivable amount). Thus, the subtle background flavors (coriander/floral, citrus) in gin C were captured by TCATA evaluation (Figure 4.7b).

This research provides novel insight about the strengths of these methods based on the attributes of interest. If one of the research objectives is to discriminate relative intensity/dominance of flavors that are ubiquitous across a product class (e.g. juniper flavor in gin) then TR is a better option than TCATA. Conversely, if the objective is to discriminate subtle flavor differences in a product with many different flavor components, TCATA would be a better option than TR. While TR is a new method, this research highlights some of the potential usefulness of this method as a tool for evaluating the temporal profile of complex food products such as gin.

4.3.4 Panel Consistency

The ICC plots (Figure 4.9) show the panel consistency for each method over the course of evaluation (aroma and in-mouth). The ICC is lowest for all methods during the beginning of in-mouth evaluation (5 – 25 s), which includes the time of expectoration (15 s). This may be due to a dithering effect during this time, where panelists experience uncertainty and indecisive behavior before selecting/ranking an attribute (Varela et al., 2018), made more difficult by the presence of competing basic tastes, volatile flavor compounds, and burning/numbing sensations all present at this time. Dithering is particularly associated with TDS (Varela et al., 2018) due to the need to choose the one dominant attribute, as opposed to selecting/ranking multiple attributes, which may explain the lower ICC scores for TDS compared to TR or TCATA (Figure 4.9).

The average ICC for each method shows that TDS provided fair repeatability (0.46 aroma, 0.59 in-mouth), while TCATA (0.69 aroma, 0.71 in-mouth) and TR (0.61 aroma, 0.70 in-mouth) provided good repeatability. These ICC values and the ability of the panel to detect statistically significant differences among gins across all temporal methods (section 3.3) indicate a well-performing panel.

4.3.5 Limitations and Future Work

One of the limitations of this research is that findings may be specific to the gins chosen. The findings from this research, in terms of the differences between temporal methods and the extent to which they differentiated gins, may differ depending on the food matrix. Future research should continue to explore the effect of matrix and complexity on the usefulness of these temporal methodologies, especially with newer methodologies like TR.

Another limitation of this research is that the in-mouth evaluation only focused on flavor, taste, and mouth burn/numbing attributes. Previous work with wines has found temporal differences in mouthfeel and texture (Kemp et al., 2019). Future research should explore using these three methods to compare the temporality of additional texture and mouthfeel attributes. A multi-sip approach was proposed to evaluate wines using TDS (Silva et al., 2018). TCATA evaluations using a multi-sip protocol were also used to evaluate milkshakes (Maheeka et al., 2021) and Syrah wines (Baker et al., 2016; Castura et al., 2022). These studies found temporal differences across sips, and future research may explore how the temporality of gin evolves with a multi-sip approach. Previous work has also suggested that TDS and TCATA can be combined to get a deeper understanding of what is perceived and what most catches the attention (Kawasaki et al., 2019). Future research may explore the difference in insights obtained from these dominance-highlighted TCATA curves, compared to TR curves.

4.4 Conclusion

All three temporal methods (TDS, TCATA, TR) were able to discriminate the sensory attributes of commercial gins but had different strengths and limitations in terms of the differences they identified. TDS was able to differentiate gins based on differences in juniper intensity (the attribute characteristic of gin and present in all gins) but was less effective than TCATA or TR at characterizing the differences between gins that differed from each other by many different attributes. TCATA excelled in characterizing the differences between gins that differed by many different attributes, capturing differences even in subtle flavor attributes. However, TCATA did not differentiate attributes like juniper which were present in all gins but at different intensities. Like TDS, TR differentiated gins based on juniper, and was able to document many different attributes in gins that were very different from each other. The shortcoming of TR is that it did not differentiate gins by some of the subtle flavor attributes. When evaluating complex products such as gin, researchers should consider these applications and limitations in order to choose an appropriate temporal method based on the overall research objective and product characteristics of interest.

References

Antalick, G., Perello, M. C., & de Revel, G. (2010). Development, validation and application of a specific method for the quantitative determination of wine esters by headspace-solid-phase microextraction-gas chromatography-mass spectrometry. Food Chemistry, 121(4), 1236–1245. https://doi.org/10.1016/j.foodchem.2010.01.011

Ares, G., Jaeger, S. R., Antúnez, L., Vidal, L., Giménez, A., Coste, B., Picallo, A., & Castura, J. C. (2015). Comparison of TCATA and TDS for dynamic sensory characterization of food products. Food Research International, 78, 148–158. https://doi.org/10.1016/J.FOODRES.2015.10.023

ASTM STP 758. (1981). Committee E-18, guidelines for the selection and training of sensory panel members, ASTM special technical publication (p.758). West Conshohocken, PA: ASTM International.

ASTM Stock #DS72. (2011). Lexicon for sensory evaluation: Aroma, flavor, texture, and appearance. West Conshohocken, PA: Stock #DS72. ASTM International.

Baker, A. K., Castura, J. C., & Ross, C. F. (2016). Temporal Check-All-That-Apply Characterization of Syrah Wine. Journal of Food Science, 81(6), S1521–S1529. https://doi.org/10.1111/1750-3841.13328

Bi, J., & Kuesten, C. (2012). Intraclass Correlation Coefficient (ICC): A Framework for Monitoring and Assessing Performance of Trained Sensory Panels and Panelists. Journal of Sensory Studies, 27(5), 352–364. https://doi.org/10.1111/j.1745-459X.2012.00399.x

Cacho, J. F., & Lopez, R. (2005). Alcoholic Beverages. Food and Nutritional Analysis. Encyclopedia of Analytical Science, 285–291. https://doi.org/https://doi.org/10.1016/B0-12-369397-7/00181-3

Carla Corrêa Simioni, S., Nayara Ribeiro, M., Rios de Souza, V., Antônio Nunes, C., & Carla Marques Pinheiro, A. (2018). Multiple-sip temporal dominance of sensations associated with acceptance test: a study on special beers. Journal of Food Science and Technology, 55(3), 1164–1174. https://doi.org/10.1007/s13197-018-3032-2

Castura, J. C., Antúnez, L., Giménez, A., & Ares, G. (2016). Temporal Check-All-That-Apply (TCATA): A novel dynamic method for characterizing products. Food Quality and Preference, 47, 79–90. https://doi.org/10.1016/j.foodqual.2015.06.017

Castura, J. C. (2022). tempR: Temporal sensory data analysis. R. R package version 0.9.9.20.

Castura, J.C. (2022). Discriminability and uncertainty in principal component analysis (PCA) of temporal check-all-that-apply (TCATA) data. Food Quality and Preference, 96: 104370. https://doi.org/10.1016/j.foodqual.2021.104370

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6(4): 284-290. https://psycnet.apa.org/doi/10.1037/1040-3590.6.4.284

Cliff, M., & Heymann, H. (1993). Development and use of time-intensity methodology for sensory evaluation: A review. Food Research International, 26(5), 375–385. https://doi.org/10.1016/0963-9969(93)90081-S

Cochran, W.G. (1950). The comparison of percentages in matched samples. Biometrika, 37(3/4): 256-266. https://doi.org/10.2307/2332378

Di Monaco, R., Su, C., Masi, P., & Cavella, S. (2014). Temporal dominance of sensations: A review. Trends in Food Science and Technology, 38(2), 104–112. http://dx.doi.org/10.1016/j.tifs.2014.04.007

Dou, Y., Mäkinen, M., & Jänis, J. (2023). Analysis of Volatile and Nonvolatile Constituents in

Gin by Direct-Infusion Ultrahigh-Resolution ESI/APPI FT-ICR Mass Spectrometry. Journal of Agricultural and Food Chemistry. https://doi.org/10.1021/acs.jafc.3c00707

Esmerino, E. A., Castura, J. C., Ferraz, J. P., Tavares Filho, E. R., Silva, R., Cruz, A. G., Freitas, M. Q., & Bolini, H. M. A. (2017). Dynamic profiling of different ready-to-drink fermented dairy products: A comparative study using Temporal Check-All-That-Apply (TCATA), Temporal Dominance of Sensations (TDS) and Progressive Profile (PP). Food Research International, 101, 249–258. https://doi.org/10.1016/J.FOODRES.2017.09.012

Fiches, G., Saint Eve, A., Jourdren, S., Déléris, I., Brunerie, P., & Souchon, I. (2016). Temporality of perception during the consumption of French grape brandies with different aging times in relation with aroma compound release. Flavour and Fragrance Journal, 31(1), 31–40. https://doi.org/10.1002/ffj.3265

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the American Statistical Association, 32(200): 675-701. https://doi.org/10.1080/01621459.1937.10503522

Harwood, W. S., Parker, M. N., & Drake, M. A. (2020). Influence of ethanol concentration on sensory perception of rums using temporal check-all-that-apply. Journal of Sensory Studies, 35(1). https://doi.org/10.1111/JOSS.12546

He, Y., Chen, S., Tang, K., Qian, M., Yu, X., & Xu, Y. (2021). Sensory characterization of Baijiu pungency by combined time-intensity (TI) and temporal dominance of sensations (TDS). Food Research International, 147. https://doi.org/https://doi.org/10.1016/j.foodres.2021.110493

Hutchings, S. C., Cha, W., Dunshea, F.R., Sharma, C., & Torrico, D.D. (2022). Understanding dominance: The effect of changing the definition of dominance when using TDS with consumers. Journal of Sensory Studies, 37: e12750. https://doi.org/10.1111/joss.12750

Ickes, C. M., & Cadwallader, K. R. (2018). Effect of ethanol on flavor perception of Rum. Food Science and Nutrition, 6(4), 912–924. https://doi.org/10.1002/fsn3.629

Jaeger, S. R., Alcaire, F., Hunter, D. C., Jin, D., Castura, J. C., & Ares, G. (2018). Number of terms to use in temporal check-all-that-apply studies (TCATA and TCATA Fading) for sensory product characterization by consumers. Food Quality and Preference, 64, 154–159. https://doi.org/10.1016/J.FOODQUAL.2017.09.013

Kawasaki, H., Yoshimura, W., Wakita, A., & Kasamatsu, C. (2019). Visualization of temporal differences between dominant perceptions in temporal dominance of sensations (TDS) and temporal check-all-that-apply (TCATA) perceptions using dominance-highlighted TCATA (dTCATA) curves. Food Quality and Preference, 77, 166-171. https://doi.org/10.1016/j.foodqual.2019.05.009

Keefer, H. M., Harwood, W. S., Castura, J. C., & Drake, M. A. (2022). Temporal ranking for characterization and improved discrimination of protein beverages. Journal of Sensory Studies, 37(4). https://doi.org/10.1111/joss.12751

Keefer, H. R. M., Rovai, D., & Drake, M. A. (2023). A Timely Application—Temporal methods, past, present, and future. In Journal of Food Science. John Wiley and Sons Inc. https://doi.org/10.1111/1750-3841.16491

Kemp, B., Trussler, S., Willwerth, J., & Inglis, D. (2019). Applying temporal check-all-that-apply (TCATA) to mouthfeel and texture properties of red wines. Journal of Sensory Studies, 34(4). https://doi.org/10.1111/joss.12503

Koch, G. G., Amara, I. A., Stokes, M. E., & Gillings, D. B. (1980). Some views on parametric and non-parametric analysis for repeated measurements and selected bibliography. *International Statistical Review/Revue Internationale de Statistique*, 48(3): 249-265. https://doi.org/10.2307/1402940

Koo, T., & Li, M. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability. Journal of Chiropractic Medicine, 15(2): 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Labbe, D., Schlich, P., Pineau, N., Gilbert, F., & Martin, N. (2009). Temporal dominance of sensations and sensory profiling: A comparative study. Food Quality and Preference, 20(3), 216–221. https://doi.org/10.1016/j.foodqual.2008.10.001

Lawless, H., & Heymann, H. (2010). Sensory Evaluation of Food Principles and Practices (D. Heldman, Ed.; Second). Springer.

Le Révérend, F. M., Hidrio, C., Fernandes, A., & Aubry, V. (2008). Comparison between temporal dominance of sensations and time intensity results. Food Quality and Preference, 19(2), 174–178. https://doi.org/10.1016/j.foodqual.2007.06.012

Leksrisompong, P., Lopetcharat, K., Guthrie, B., Drake, M. (2012). Descriptive analysis of carbonated regular and diet lemon-lime beverages. Journal of Sensory Studies, 27(4), 247-263. 10.1111/j.1745-459X.2012.00389.x

Lenfant, F., Loret, C., Pineau, N., Hartmann, C., & Martin, N. (2009). Perception of oral food breakdown. The concept of sensory trajectory. Appetite, 52(3), 659–667. https://doi.org/10.1016/j.appet.2009.03.003

Maheeka, W. N. R. P., Godfrey, A. J. R., Ellis, A., & Hort, J. (2021). Comparing temporal sensory product profile data obtained from expert and consumer panels and evaluating the value of a multiple sip TCATA approach. Food Quality and Preference, 89, 104141. https://doi.org/10.1016/J.FOODQUAL.2020.104141

McDonnell, E., Hulin-Bertaud, S., Sheehan, E. M., & Delahunty, C. M. (2001). Development and learning process of a sensory vocabulary for the odor evaluation of selected distilled beverages using descriptive analysis. Journal of Sensory Studies, 16(4), 425–445. https://doi.org/10.1111/j.1745-459x.2001.tb00311.x

McMahon, K. M., Culver, C., Castura, J. C., & Ross, C. F. (2017). Perception of carbonation in sparkling wines using descriptive analysis (DA) and temporal check-all-that-apply (TCATA). Food Quality and Preference, 59, 14–26. https://doi.org/10.1016/j.foodqual.2017.01.017

Medel-Marabolí, M., López-Solís, R., Valenzuela-Prieto, D., Vargas-Silva, S., & Obreque-Slier, E. (2021). Limited relationship between temporality of sensory perception and phenolic composition of red wines. LWT, 142. https://doi.org/10.1016/j.lwt.2021.111028

Meilgaard, M.C., Civille, G.V., & Carr, B.T. (2016). *Sensory Evaluation Techniques*. 5th Edition. CRC Press.

Meillon, S., Urbano, C., & Schlich, P. (2009). Contribution of the Temporal Dominance of Sensations (TDS) method to the sensory description of subtle differences in partially dealcoholized red wines. Food Quality and Preference, 20(7), 490–499. https://doi.org/10.1016/j.foodqual.2009.04.006

Meyners, M. (2020). Temporal methods: Are we comparing apples and oranges? Food Quality and Preference, 79. https://doi.org/10.1016/j.foodqual.2018.11.022

Meyners, M., & Castura, J. C. (2016). Randomization of CATA attributes: Should attribute lists be allocated to assessors or to samples? Food Quality and Preference, 48, 210–215. https://doi.org/10.1016/j.foodqual.2015.09.014

Nguyen, Q. C., Næs, T., & Varela, P. (2018). When the choice of the temporal method does make a difference: TCATA, TDS and TDS by modality for characterizing semi-solid foods. Food Quality and Preference, 66, 95–106. https://doi.org/10.1016/j.foodqual.2018.01.002

Nykanen, L. (1986). Formation and Occurrence of Flavor Compounds in Wine and Distilled Alcoholic Beverages. American Journal of Enology and Viticulture, 37(1), 84-96. DOI: 10.5344/ajev.1986.37.1.84

Olaniran, A. O., Hiralal, L., Mokoena, M. P., & Pillay, B. (2017). Flavour-active volatile compounds in beer: production, regulation and control. In Journal of the Institute of Brewing (Vol. 123, Issue 1, pp. 13–23). John Wiley and Sons Inc. https://doi.org/10.1002/jib.389

Pierguidi, L., Spinelli, S., Monteleone, E., & Dinnella, C. (2021). The combined use of temporal dominance of sensations (TDS) and discrete time-intensity (DTI) to describe the dynamic sensory profile of alcoholic cocktails. Food Quality and Preference, 93. https://doi.org/https://doi.org/10.1016/j.foodqual.2021.104281

Pineau, N., de Bouillé, A. G., Lepage, M., Lenfant, F., Schlich, P., Martin, N., & Rytz, A. (2012). Temporal Dominance of Sensations: What is a good attribute list? Food Quality and Preference, 26(2), 159–165. https://doi.org/https://doi.org/10.1016/j.foodqual.2012.04.004

Pineau, N., Schlich, P., Cordelle, S., Mathonnière, C., Issanchou, S., Imbert, A., Rogeaux, M., Etiévant, P., & Köster, E. (2009). Temporal Dominance of Sensations: Construction of the TDS curves and comparison with time–intensity. Food Quality and Preference, 20(6), 450–455. https://doi.org/10.1016/J.FOODQUAL.2009.04.005

Poveromo, A. R., & Hopfer, H. (2019). Temporal check-all-that-apply (TCATA) reveals matrix interaction effects on flavor perception in a model wine matrix. Foods, 8(12). https://doi.org/10.3390/foods8120641

R Core Team. (2021). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Ramsey, I., Ross, C., Ford, R., Fisk, I., Yang, Q., Gomez-Lopez, J., & Hort, J. (2018). Using a combined temporal approach to evaluate the influence of ethanol concentration on liking and sensory attributes of lager beer. Food Quality and Preference, 68, 292–303. https://doi.org/10.1016/j.foodqual.2018.03.019

Reyes, M. M., Castura, J. C., & Hayes, J. E. (2017). Characterizing dynamic sensory properties of nutritive and nonnutritive sweeteners with temporal check-all-that-apply. Journal of Sensory Studies, 32(3). https://doi.org/10.1111/JOSS.12270

Riu-Aumatell, M., Vichi, S., Mora-Pons, M., López-Tamames, E., & Buxaderas, S. (2008). Sensory characterization of dry gins with different volatile profiles. Journal of Food Science, 73(6). https://doi.org/10.1111/j.1750-3841.2008.00820.x

Riu Aumatell, M. (2012). Gin: production and sensory properties. In Alcoholic beverages: Sensory evaluation and consumer research. Oxford: Woodhead Publishing. 267-280. https://doi.org/10.1533/9780857095176.3.267.

Schlich, P., & Pineau, N. (2017). Temporal Dominance of Sensations. Time-Dependent

Measures of Perception in Sensory Evaluation, 283–320. https://doi.org/10.1002/9781118991640.CH11

Silva, A. P., Voss, H. P., van Zyl, H., Hogg, T., de Graaf, C., Pintado, M., & Jager, G. (2018). Temporal dominance of sensations, emotions, and temporal liking measured in a bar for two similar wines using a multi-sip approach. Journal of Sensory Studies, 33(5). https://doi.org/10.1111/JOSS.12459

Sokolowsky, M., & Fischer, U. (2012). Evaluation of bitterness in white wine applying descriptive analysis, time-intensity analysis, and temporal dominance of sensations analysis. Analytica Chimica Acta, 732, 46–52. https://doi.org/10.1016/j.aca.2011.12.024

Stoffel, E., Robertson, T., Catania, A., & Casassa, F.L. (2023a). The impact of fermentation temperature and cap management on selected volatile compounds and temporal sensory characteristics of grenache wines from the central coast of California. Molecules, 28(10). https://doi.org/10.3390/molecules28104230

Stoffel, E., Lesniauskas, R., Anderson, S., Krystoff, C., & Casassa, F.L. (2023b). Temporal evaluation of retronasal and mouthfeel sensations in cofermented and blended red wines from California. American Journal of Enology and Viticulture, 74(1). DOI: 10.5344/ajev.2023.22065

Varela, P., Antúnez, L., Carlehög, M., Alcaire, F., Castura, J. C., Berget, I., Giménez, A., Næs, T., & Ares, G. (2018). What is dominance? An exploration of the concept in TDS tests with trained assessors and consumers. Food Quality and Preference, 64, 72–81. https://doi.org/10.1016/j.foodqual.2017.10.014

Vichi, S., Riu-Aumatell, M., Mora-Pons, M., Buxaderas, S., & López-Tamames, E. (2005).

Characterization of volatiles in different dry gins. Journal of Agricultural and Food Chemistry, 53(26), 10154–10160. https://doi.org/10.1021/jf058121b

Vidal, L., Antúnez, L., Gimenez, A., Medina, K., Boido, E., & Ares, G. (2016). Dynamic characterization of red wine astringency: Case study with Uruguayan Tannat wines. Food Research International, 82, 128-135. https://doi.org/10.1016/j.foodres.2016.02.002

Wakihira, T., Visalli, M., & Schlich, P. (2023). Temporal drivers of liking by period: A case study on lemon-flavored carbonated alcoholic drinks with consumers in natural settings. Food Quality and Preference, 104793. https://doi.org/10.1016/j.foodqual.2022.104793

Wang, Q. J., Niaura, T., & Kantono, K. (2021). How does wine ageing influence perceived complexity? Temporal-Choose-All-That-Apply (TCATA) reveals temporal drivers of complexity in experts and novices. Food Quality and Preference, 92. https://doi.org/10.1016/j.foodqual.2021.104230

Williams, R.M. (1952). Experimental designs for serially correlated observations. Biometrika, 39 (1-2): 151-167. https://psycnet.apa.org/doi/10.1093/biomet/39.1-2.151

Yu, H., Guo, W., Xie, J., Ai, L., Chen, C. (2023). Evaluation of taste characteristics of Chinese rice wine by quantitative description analysis, dynamic description sensory and electronic tongue. Journal of Food Measurement and Characterization, 17(1): 824-835. http://dx.doi.org/10.1007/s11694-022-01637-2

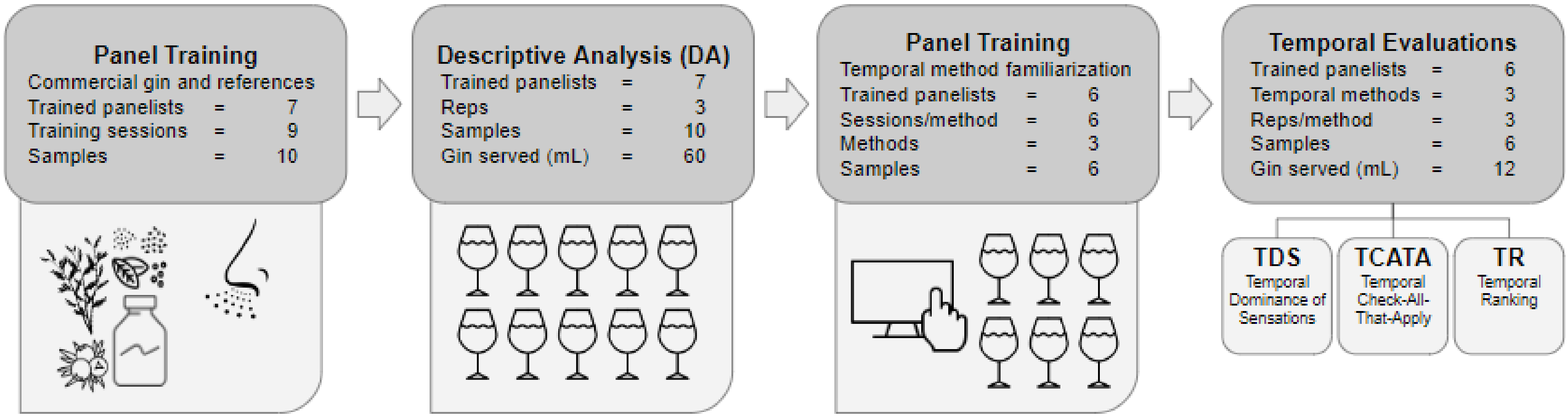


Figure 4.1 Overview of experimental design and flow.

Commercial gins and references (natural references, chemical references, and flavor-infused ethanol solutions) were used to train panelists and aid in lexicon development. A trained panel evaluated the ten gins across three reps. Six representative gins were chosen for temporal evaluations. Panelists were trained on each temporal evaluation method prior to evaluating six gins across three reps for each of the three methods.

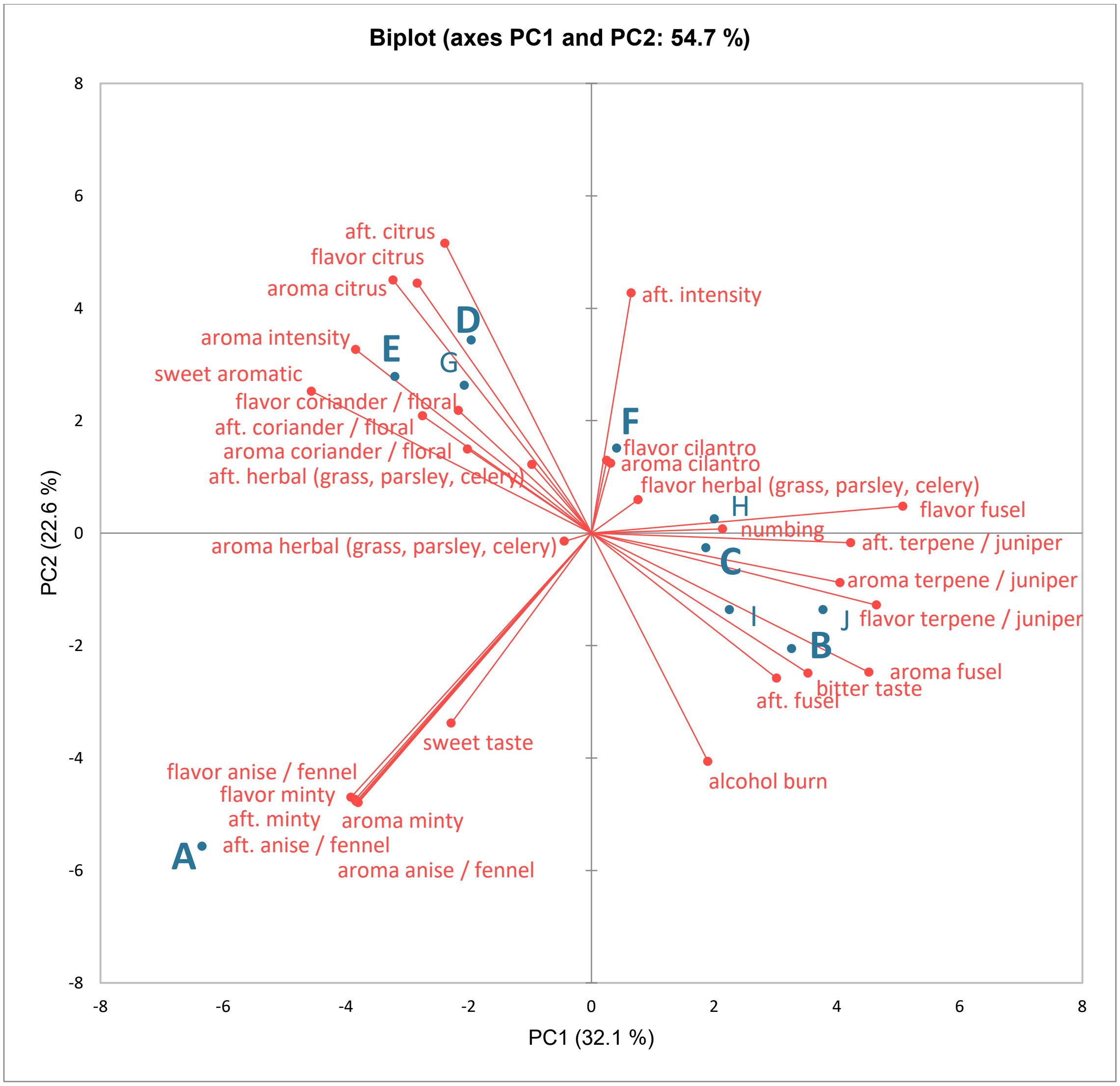


Figure 4.2 PCA biplot for 10 gins evaluated by descriptive analysis. Ten commercial gins in blue (A – J) were profiled by seven trained panelists using the attributes in red. Bolded samples (A – F) were chosen for temporal evaluation because they were well separated. aft. = aftertaste

Table 4.1 Lexicon used for descriptive analysis of gins.

Attribute	Description	Reference
aroma description	adjectives that describe the orthonasal headspace	-
[4]aroma intensity	overall intensity of the aroma in the orthonasal headspace	-
[4]sweet aromatic	aromatics associated with sweet foods such as vanilla, sweet solutions, syrups	-
[1,2]coriander / floral	aromatics associated with coriander, flowers, perfumes	3% w/w ground coriander (Frontier co-op) in 30% solution of ethanol in DI water; Chemical references: linalool, limonene (Sigma-Aldrich, Rockville, MD)
[1,2]citrus	aromatics of fruits such as oranges, lemons and limes	6% w/w lemon peel in 30% solution of ethanol (Fisher Scientific, Waltham, MA, USA) in DI water
herbal (grass, parsley, celery)	green aromatics associated with grass, parsley, celery	natural reference - parsley
cilantro	green aromatics specific to cilantro	natural reference - cilantro
fusel	aromatics of alcohols such as acetone	acetone
[1,2]anise / fennel	aromatics of anise seed, licorice, fennel	3% w/w ground fennel (Morton & Bassett Spices) in 30% solution of ethanol in DI water
minty	aromatics of peppermint, spearmint	natural reference - mint
[1,2]terpene / juniper	aromatics of juniper berries, pine needles, evergreen conifers	3% w/w ground juniper berries (Frontier co-op) in 30% solution of ethanol in DI water; Chemical references: myrcene, γ-terpinene, α-pinene (Sigma-Aldrich, Rockville, MD)
[4]sweet taste	basic taste elicited by sugars	sweet 2 solution - 2% sucrose
[4]bitter taste	basic taste elicited by caffeine, quinine	bitter 2 solution - 0.05% caffeine
[5]alcohol burn	perception of trigeminal pain/irritation from exposure to ethanol while sample is in the mouth	cinnamon oil (500 ppm)
aft. intensity	overall intensity of the aftertaste up to 1 minute after expectoration	-
[3]numbing	loss of tactile sensation in the oral cavity	1 cough drop (Halls sugar-free mentho-lyptus honey-lemon)/300 mL boiling water = numbing 3

Adapted from [1]McDonnell et al. (2001), [2]Riu-Aumatell et al. (2008), [3]Leksrisompong et al. (2012), [4]Meilgaard et al. (2016), [5]Harwood et al. (2020).

Table 4.2 Presentation design for temporal evaluation of gin across 9 sessions (3 methods x 3 reps) where all permutations were presented blocked by method. In each session, panelists evaluated all six samples of gin (presented using a Williams Latin square design to minimize sample position bias). TDS = Temporal Dominance of Sensations, TCATA = Temporal Check-All-That-Apply, TR = Temporal Ranking.

Session	Panelist 1	Panelist 2	Panelist 3	Panelist 4	Panelist 5	Panelist 6
1	TR	TDS	TR	TCATA	TDS	TCATA
2	TR	TDS	TR	TCATA	TDS	TCATA
3	TR	TDS	TR	TCATA	TDS	TCATA
4	TDS	TCATA	TCATA	TR	TR	TDS
5	TDS	TCATA	TCATA	TR	TR	TDS
6	TDS	TCATA	TCATA	TR	TR	TDS
7	TCATA	TR	TDS	TDS	TCATA	TR
8	TCATA	TR	TDS	TDS	TCATA	TR
9	TCATA	TR	TDS	TDS	TCATA	TR

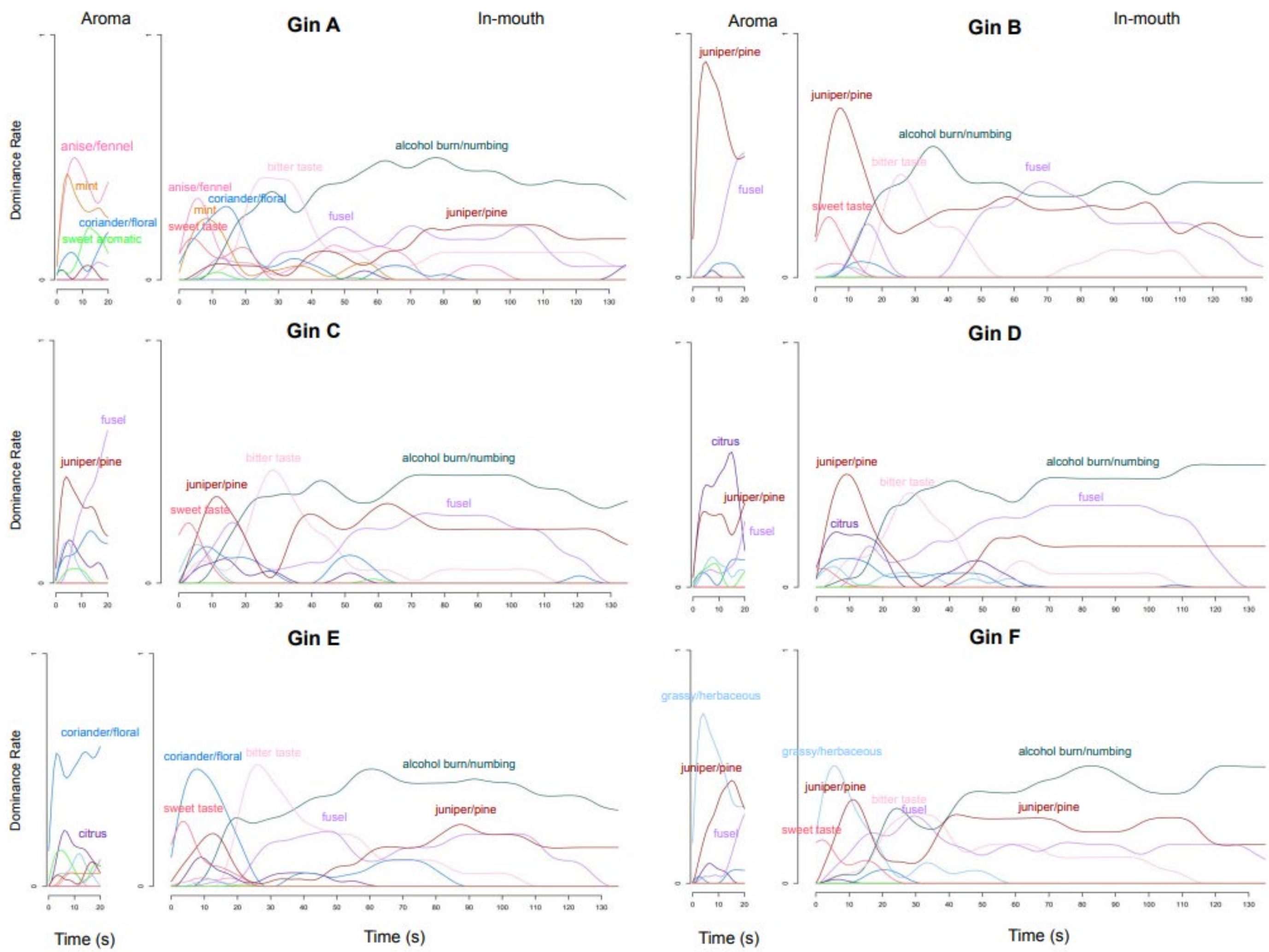


Figure 4.3 TDS curves for gin products A-F. Aroma was evaluated for 20 s before a 15-s pause, followed by 135 s total of in-mouth evaluation with expectoration after 15 s and aftertaste persisting until the end. Curves represent the average dominance rate across all participants and reps.

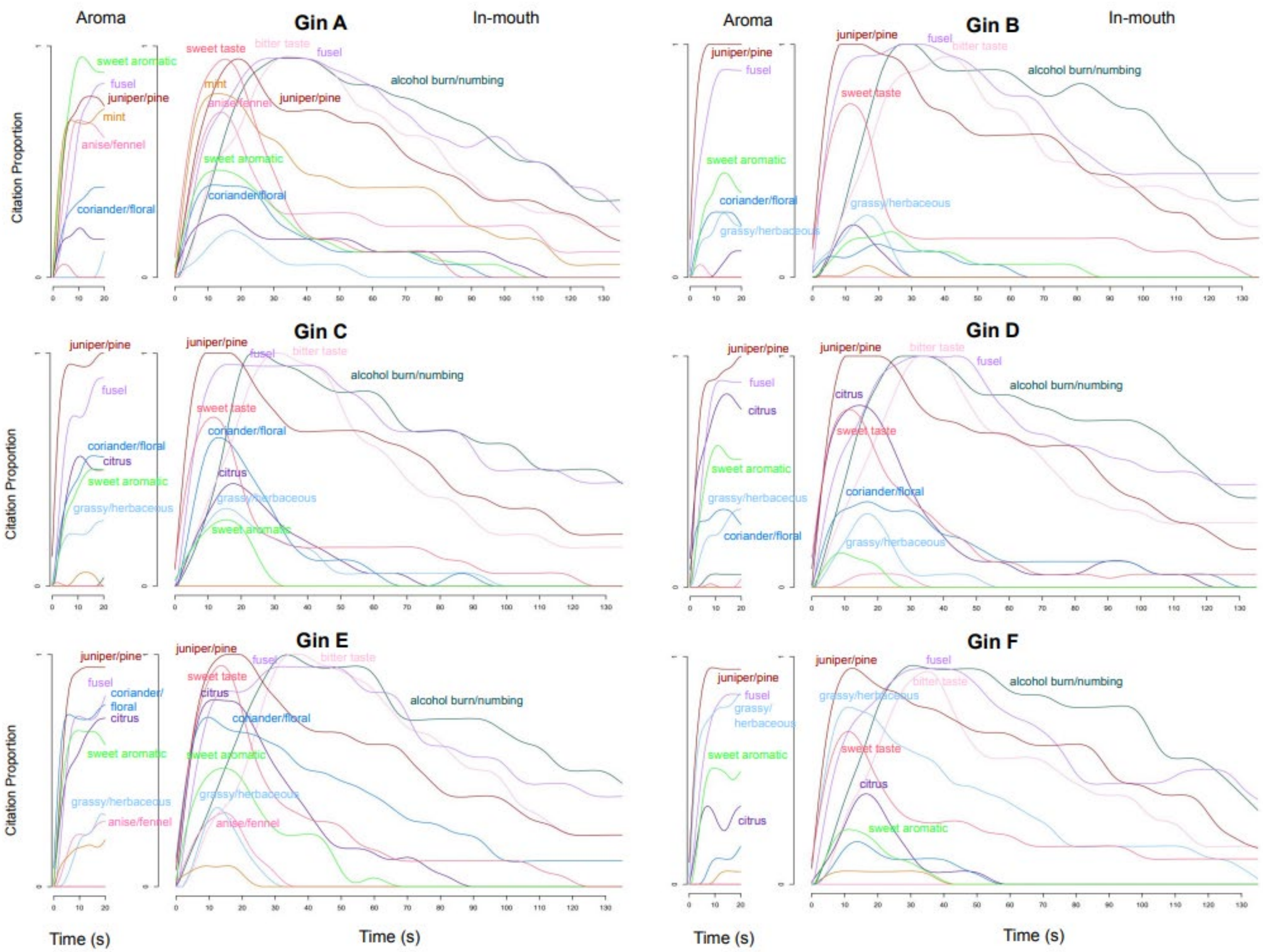


Figure 4.4 TCATA curves for gin products A-F. Aroma was evaluated for 20 s before a 15-s pause, followed by 135 s total of in-mouth evaluation with expectoration after 15 s and aftertaste persisting until the end. Curves represent the average citation proportion across all participants and reps.

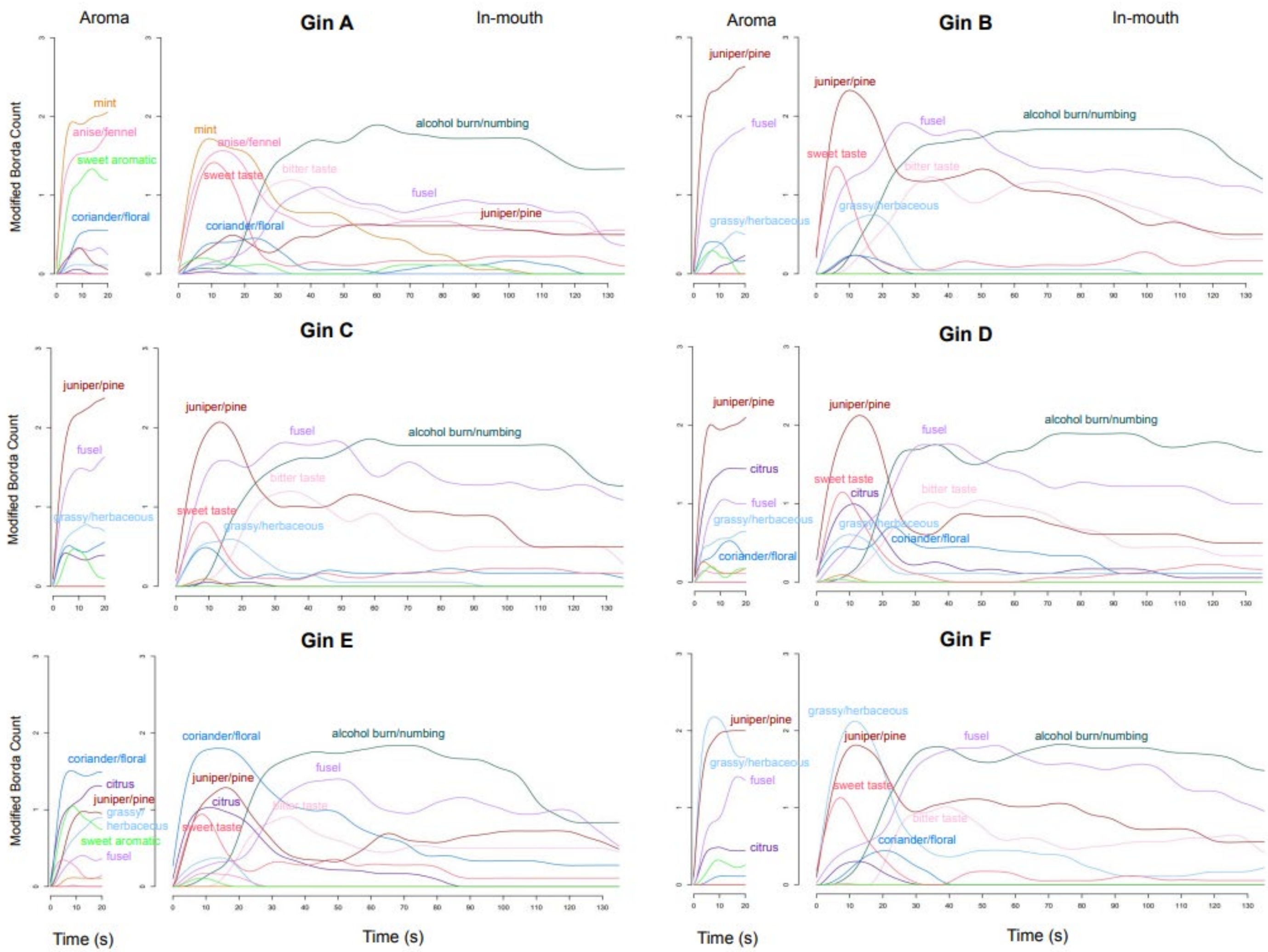


Figure 4.5 TR curves for gin products A-F. Aroma was evaluated for 20 s before a 15-s pause, followed by 135 s total of in-mouth evaluation with expectoration after 15 s and aftertaste persisting until the end. Curves represent the average modified borda count across all participants and reps.

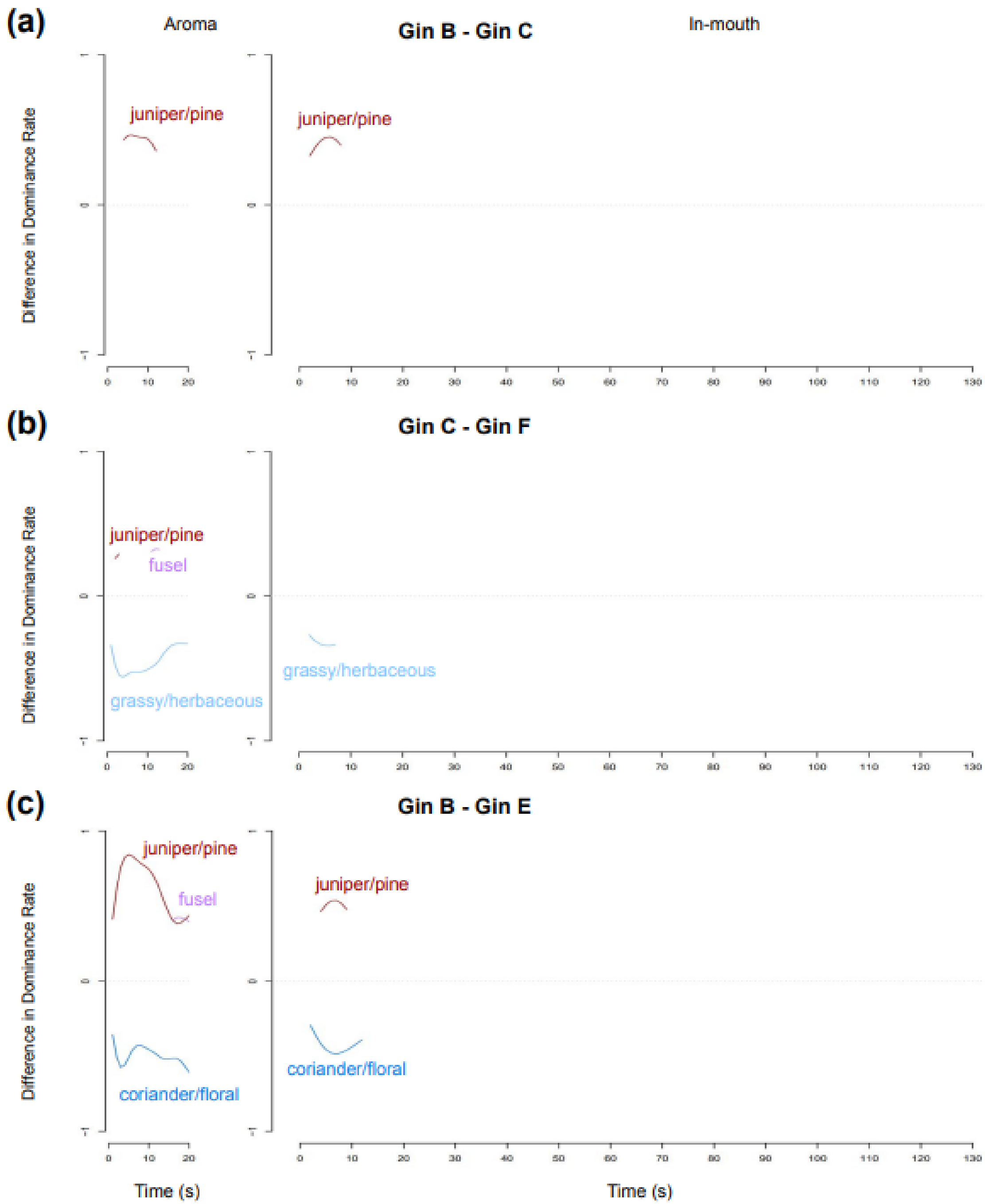


Figure 4.6 TDS difference curves for gins that were (a) only slightly different, (b) moderately different, and (c) very different. Aroma was evaluated for 20 s before a 15-s pause, followed by 135 s total of in-mouth evaluation with expectoration after 15 s and aftertaste persisting until the end. Positive values are associated with the gin listed first, while negative values are associated with the gin listed second. Significant differences determined via McNemar's test ($\alpha = 0.05$).

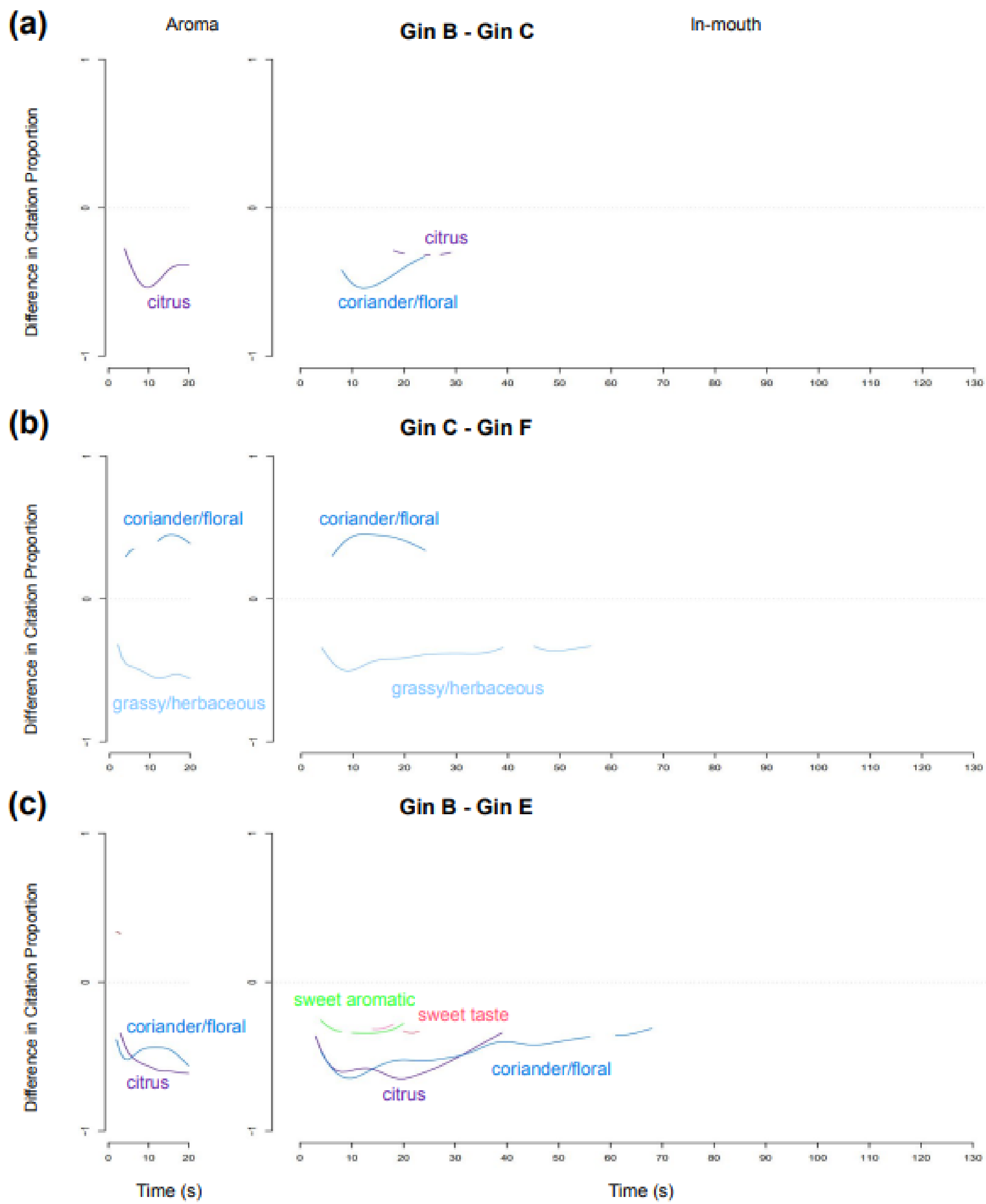


Figure 4.7 TCATA difference curves for gins that were (a) only slightly different, (b) moderately different, and (c) very different. Aroma was evaluated for 20 s before a 15-s pause, followed by 135 s total of in-mouth evaluation with expectoration after 15 s and aftertaste persisting until the end. Positive values are associated with the gin listed first, while negative values are associated with the gin listed second. Significant differences determined via McNemar's test ($\alpha = 0.05$).

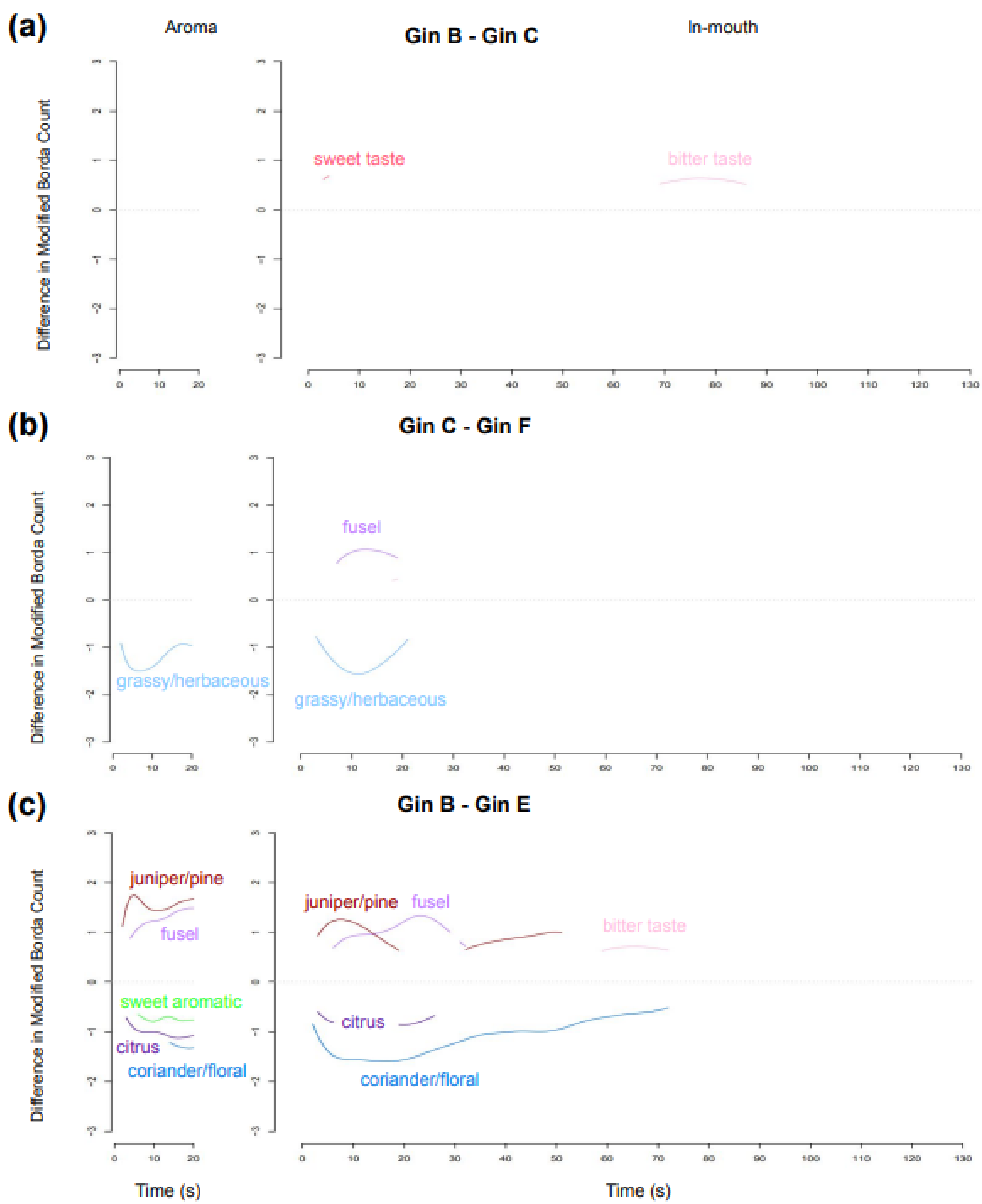


Figure 4.8 TR difference curves for gins that were (a) only slightly different, (b) moderately different, and (c) very different. Aroma was evaluated for 20 s before a 15-s pause, followed by 135 s total of in-mouth evaluation with expectoration after 15 s and aftertaste persisting until the end. Positive values are associated with the gin listed first, while negative values are associated with the gin listed second. Significant differences determined via McNemar's test ($\alpha = 0.05$).

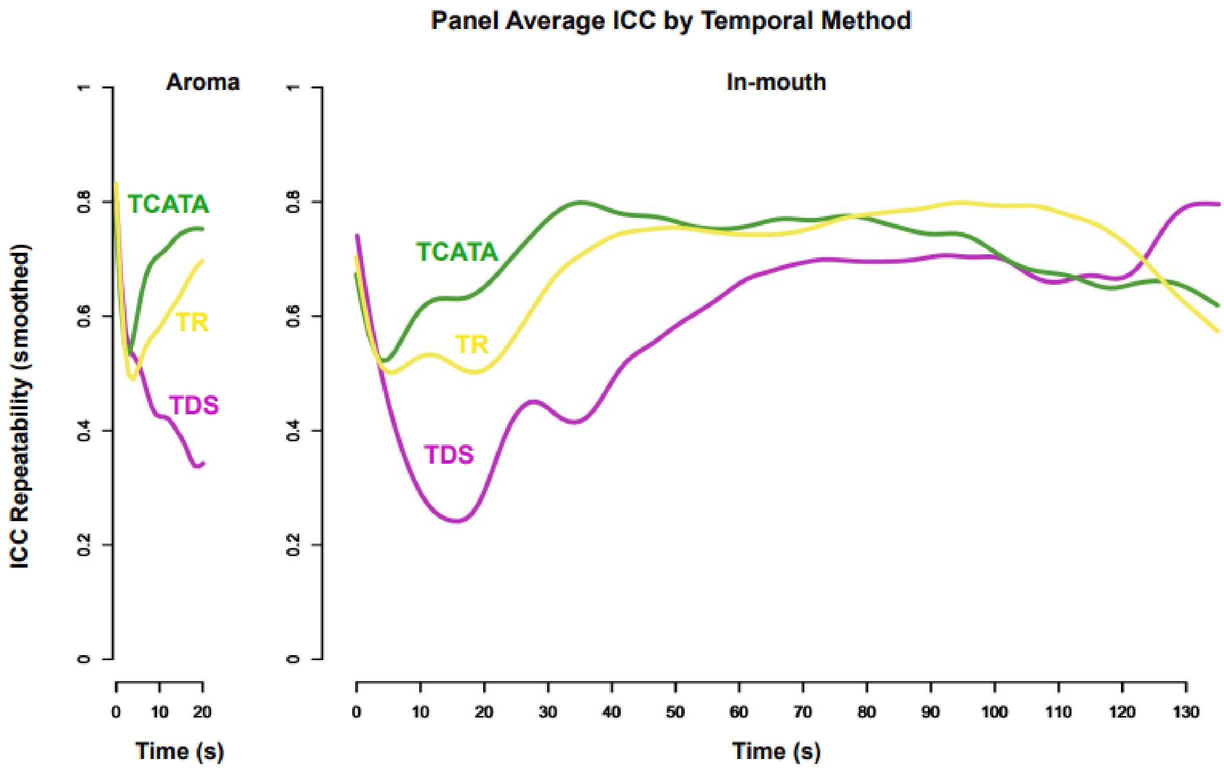


Figure 4.9 Intraclass Correlation Coefficient (ICC) repeatability for TDS, TCATA, and TR over the course of evaluation (20 s aroma, 135 s in-mouth), averaged across the panel. Higher ICC values represent better repeatability.

CHAPTER 5:

Consumer Acceptance of Protein Beverage Ingredients: Less is More

Consumer Acceptance of Protein Beverage Ingredients: Less is More

D. Rovai[1], M. E. Watson[1], D. M. Barbano[2], M. A. Drake[1]

[1] Southeast Dairy Foods Research Center, Department of Food, Bioprocessing, and Nutrition Sciences, North Carolina State University, Raleigh, NC 27606

[2] Northeast Dairy Food Research Center, Department of Food Science, Cornell University, Ithaca, NY 14853

Choice of journal/topic: Journal of Dairy Science, Accepted for publication, In Press.

Interpretive Summary

This research provides insight into consumer perception of protein beverages and the ingredients found in them. Consumers drink protein beverages as a convenient way to increase protein content. Consumers want protein and natural sweeteners in their beverages but would prefer beverages that do not contain stabilizers or thickeners. Consumers prefer natural sweeteners. Preference for protein source varies depending on the group of consumers.

Abstract

An array of ingredients are added to protein beverage formulations. These ingredients may not be desirable to consumers. Our objective was to determine consumer perception of ingredients in protein beverages. An online survey was conducted with protein beverage consumers (n=405). Maximum difference scaling and projective mapping were applied to determine the relative acceptance of ingredients based on their functional role (protein source, sweetener, stabilizer, thickener). Subsequently, four 120-min focus groups were conducted (n=25 consumers). Survey data were evaluated by univariate and multivariate statistics. Consumer sentiment from focus groups was compiled and grouped based on themes that emerged across multiple focus groups. Consumers placed the highest importance on the amount of protein followed by protein type in protein beverages. Plant protein, whey protein, and milk protein were most appealing, while soy protein, collagen, and casein/caseinates were less appealing ($p<0.05$). Natural sweeteners (agave, monk fruit, cane sugar) were the most appealing sweeteners ($p<0.05$). Fibers and starches were more appealing than gums (carrageenan, gellan gum) ($p<0.05$). Stabilizers were the least desirable class of beverage ingredients, with sodium and potassium phosphates the least desirable ($p<0.05$). In regard to the package of a protein beverage, consumers placed the greatest importance on recognizable ingredients and plain language ($p<0.05$). Consistent with survey results, consumers in focus groups expressed skepticism and feeling overwhelmed by all of the ingredients on the label of protein beverages. Protein was their primary desire and the presence of sweeteners was acceptable, but they did not desire additional ingredients. There is an opportunity to increase the acceptance and competitiveness of dairy protein beverages by reformulating beverages to include fewer and more familiar ingredients.

Functional proteins, such as those derived from dairy, may have opportunities to exclude undesirable ingredients (stabilizers, thickeners) from the label.

Key Words: ready-to-drink, protein beverage, clean label

5.1 Introduction

Within the food industry, there is a rise in consumer demand for healthier foods and clean label products (Bartelme et al., 2024). Clean labels are perceived by consumers as healthier, more socially responsible, and indicative of a product with better sensory appeal (Cao and Miao, 2023). While there is no established legal definition for the term "clean label" (Osborne, 2015), it is often used in the context of food products to describe products with minimal ingredients, only necessary ingredients, recognizable ingredients, natural ingredients, free from additives, and minimally processed (Osborne, 2015; Cao and Miao, 2023). Consumers use cues found on both the front and back of packages to form their perception of clean label (Asioli et al., 2017), thus influencing their purchase decision.

In the United States, despite little growth in fluid dairy sales (Mills, 2024a), growth of dairy products is driven by higher protein and value-added products such as yogurt with estimated sales of $1.7 billion (Mills, 2023) and cheese which represented $31.2 billion in sales in 2023 (Mills, 2024b). Sales of nutrition drinks were $8.0 billion in 2023 and are expected to grow to $9.96 billion by 2028 (Chychula, 2024). The growth and demand for protein beverages has led to a diverse product space that includes a variety of forms and types of ingredients. Protein beverages consist of ready-to-mix (RTM) powders and ready-to-drink (RTD) beverages, and can be made with either dairy protein, plant protein, or a blend of proteins (Singh et al., 2022). With this shift towards health-conscious purchasing, consumers are faced with tradeoffs as they shop in the grocery aisle. Considering ready-to-drink (RTD) protein beverages, consumers want a protein beverage that has low calories, gives them protein and energy to engage in activity, and provides satiety (Oltman et al., 2015). However, to deliver the protein content that consumers desire, current formulations for RTD protein beverages include additives.

These additives aid in beverage shelf-life stability, ensuring the products hold up during processing and maintain quality (flavor and lack of gelation/thickening) throughout their shelf life (Corbo et al., 2014). Additives for beverage texture include ingredients that enhance heat stability and/or improve/maintain viscosity such as buffering agents (e.g. phosphate salts, citric acid), chelating agents (e.g. sodium hexametaphosphate, EDTA), thickeners (e.g. carrageenan, guar gum), and emulsifiers (e.g. lecithin) (Lindsay, 2008). While consumers tend to prefer food products with minimal ingredients, it is unclear how consumers perceive protein beverages with additives, and furthermore, if different classes of additives and/or types of additives within a class affect consumer perception of these products. Understanding consumer perception towards these different classes of ingredients may help product developers target their efforts to create cleaner label RTD protein beverages, while maintaining product functionality.

Previous research has used surveys and focus groups to explore attributes of interest for consumers in a protein beverage (Childs et al., 2008), and investigated the effect of priming consumers on perceived liking of protein beverages (Oltman et al., 2015). There is minimal literature which evaluates consumer perception of specific ingredients and classes of ingredients which are currently used in commercial protein beverage manufacture. The goal of this research was to understand consumer perception of clean label as it pertained to protein beverages, explore how different types of ingredients were perceived relative to each other, and use these insights to understand how the industry may move forward towards creating clean label protein beverages.

5.2 Methods

5.2.1 Commercial Protein Beverage Ingredients

Ingredient labels were compiled for over 60 different commercial protein beverages available locally (Raleigh, NC) and online. Ingredients were sorted and categorized by their role in the protein beverage from a nutritional/functional standpoint. Four ingredient classes of interest were determined: protein, sweetener, thickener, stabilizer. Proteins and sweeteners were straightforward classes, comprising ingredients labeled as proteins, and known sweeteners. Thickeners comprised ingredients such as fibers, flours, starches, gums, and gels–ingredients added for their ability to thicken food matrices. The class "stabilizers" was designated to encompass the other ingredients added to food products to add stability in a variety of ways–heat stability, flavor stability, emulsion stability, color stability, etc. While the authors acknowledge that some of these ingredients also provide some other functions to the food matrix, such as sweetness or thickening, they were ultimately assigned to the class of stabilizers.

After categorizing ingredients by class, ingredients were consolidated to remove redundancies in ingredient terminology and create manageable lists of 13 to 21 ingredients per class (68 total ingredients). These ingredients were then evaluated in the context of a protein beverage using an online survey and moderator-guided focus groups to determine consumer perception towards each of these classes of ingredients, and the relative acceptability of the ingredients within each class.

5.2.2 Experimental Overview

The experimental is summarized by Figure 5.1. An online survey was conducted to assess consumer perceptions of the ingredients found in protein beverages. Participants who completed the survey in its entirety (n=405) were compensated with a $5 Amazon e-gift card. Following the survey, four 2-hour moderator guided focus groups were conducted with a subset of participants (n=25) to further explore the insights generated by the survey data and provide context about perception towards the ingredients in protein beverages. Participants who completed the focus groups were compensated with a $75 Amazon e-gift card. Survey and focus group methodology are discussed further in the following sections. All surveys and focus groups were conducted in compliance with North Carolina State University Institutional Review Board regulations.

5.2.3 Online Survey

An online survey was developed using Lighthouse Studio (Sawtooth Software version 9.14.2, Orem, UT) and Compusense Cloud (Compusense Inc., Guelph, ON, Canada). The survey was uploaded, and participants were recruited from a database of over 11,000 individuals, maintained by the Sensory Service Center at North Carolina State University. Participants first answered demographic questions and were screened out to include only panelists who were 18 years of age or older, responsible for at least 25% of household grocery shopping, and protein beverage consumers who read the nutrition and ingredients labels. Panelists (n = 405) proceeded to evaluate ingredient acceptability (Maximum Difference Scaling exercises, modified projective mapping exercise) and answer questions about their protein beverage consumption, purchase habits, preferences, and other psychographics (single select, CATA, agreement scales, sliding scales, and constant sum questions).

5.2.3.1 Maximum Difference Scaling (MXD)

Beverage ingredients within a class were first compared to each other using a Maximum Difference Scaling exercise (MXD). This exercise works by presenting participants with a series of tasks, where each task presented a subset of five ingredients from the master list (13 - 21 ingredients per class) to participants. During each task, participants were asked to indicate which ingredient was the most appealing, and which ingredient was the least appealing for purchase of a protein beverage. Participants performed a series of these tasks such that each ingredient from the master list was shown five times.

Following the MXD exercises for each class of ingredients, a single select question was asked for each ingredient: If I saw this ingredient on a label, I would… 1) Not buy it, 2) Still consider buying it. This question was integrated into the calculation of the utility score by acting as an anchor point or utility boundary (Lattery, 2011), where ingredients below the anchor point indicate that participants would not buy the protein beverage if they saw the ingredient on the label, and ingredients above the anchor point indicate that participants would still consider buying the product.

5.2.3.2 Modified Projective Mapping

Following the MXD exercise, a modified projective mapping exercise was used to visualize differences in ingredient perception in a 2D space. Previous research has used classical projective mapping (no axis anchoring) to group ingredients, but the results provided limited information by way of understanding the nuances in ingredient perception (Aschemann-Witzel et al., 2019). Thus, the modified projective map exercise in this current study utilized anchors with increasing familiarity on the x-axis (left to right), and increasing liking on the y-axis (bottom to top). These anchors were chosen because research has shown that familiar foods are associated with higher liking (Hong et al., 2014), and ingredient familiarity/recognizability is a key factor to the ingredients consumers like seeing on their labels (Oltman et al., 2015). Participants were presented the ingredients for projective mapping using a balanced incomplete design where each participant saw a random presentation of 17 out of 68 total ingredients. Each ingredient was seen an average of 101 times. The average placement of each ingredient was determined by plotting the average x and y-coordinates.

5.2.4 Focus Groups

Four moderator-guided focus groups (120 minutes each) were conducted with a subset of ready-to-drink (RTD) protein beverage consumers who completed the online survey (n = 25). The focus groups were conducted within 60 days after the survey concluded and were used as a tool to probe the insights from the survey results and supplement them with qualitative data to capture consumer sentiment and emotions that would otherwise not be captured through quantitative survey data collection. The focus groups were structured to first discuss current protein beverage purchase habits and characteristics of each individual's favorite protein beverage(s). Next, the discussion transitioned to protein beverage usage occasion to understand when and why these beverages are consumed. Finally, the discussion focused on the ingredient decks and nutrition labels, providing participants with commercial products to examine, discuss, and compare. This discussion guide was designed to start broad by gathering information about typical usage and behavior and allowing for natural consumer thoughts and behavior to emerge, before narrowing the focus to discuss ingredients. This approach was chosen to avoid bias that may result from hyper focusing participants on ingredients too early in the discussion. Similarly, the first half of the focus group focused on all protein beverages (including powders), while the second half of the discussion was narrowed down to RTD protein beverages, as these products were the focus of this research.

Qualitative data were collected through in-person note taking during each focus group, as well as video and audio recording. After the conclusion of all focus groups, notes were compiled, and key themes were identified as they pertained to usage, purchase behavior, and consumer perception of ingredients. Key themes were determined by assessment of important topics that recurred across all four focus group sessions. These key themes and participant quotes that support each theme were summarized graphically (Figure 5.2).

5.2.5 Statistical Analysis

Utility scores for MXD exercises were determined via hierarchical Bayesian (HB) estimation (Orme, 2009) using the anchored approach (Lattery, 2011). HB estimates the utilities for the entire sample (upper level) and how much individual respondents differ from the sample average (lower level). The HB estimation performs a series of iterations starting with random seeding. As more and more iterations are run, the model begins to converge. Once the model has converged (utility variance is consistent), the subsequent runs are used to calculate the final utility score (Howell, 2009). The matrix output contains utility scores for each attribute (column) for each respondent (row). ANOVA with Fisher's least significant difference (α=0.05) is used to compare utility scores across attributes (columns). Segmentation of MXD utility scores was performed using Latent Class analysis. Latent Class analysis simultaneously estimates part worth utilities for each segment and the probability of each respondent belonging to each segment. It does this by first selecting random estimates for each segment's utility scores, and then performing a series of iterations where the group utility scores are used to fit each respondent's data and estimate the probability that they belong to each group. Then these probabilities are used as weights to re-estimate the probability of belonging to each group expressed as log-likelihood of belonging to each group. This process is repeated until the log-likelihood fails to improve, meaning the estimation has converged, and panelists are assigned to the cluster they have the highest log-likelihood of belonging to (Sawtooth Software, 2021). Agreement and scaled questions (5-point scales) were analyzed via Kruskal-Wallis with Dunn's post hoc test (α=0.05). Correspondence analysis symmetric row plots were generated for CATA terms describing ingredients from different clusters.

To ensure reliability of the data collected, a few methods were implemented. Two attention check questions were included in the survey which asked participants to select a specific response, and individuals who did not read the directions and selected the wrong answer were disqualified from the survey. In addition, to identify individuals who were just clicking through the MXD exercises without putting meaningful thought into their responses, data (n = 300) were randomly generated through Sawtooth for each of the MXD exercises. This randomly generated data were then run using hierarchical Bayesian (HB) estimation and the 95th percentile root likelihood (RLH) score for these responses was used as a cutoff value. Data from participants with RLH scores at or below this value were removed from the data set (Orme, 2019). Of the 422 survey participants who completed the survey in its entirety, data from 17 participants were discarded based on the criteria above. Data from n=405 participants were used for all survey analyses.

5.3 Results and Discussion

5.3.1 Demographics

The majority of participants (n=405) were white/Caucasian (62.5%) females (71.6%) who belonged to Gen Z (35.1%, 18-27 years old) or Millennials (33.3%, 28-42 years old), resided in North Carolina (84.7% - the state where the affiliated university is located), were well educated (78.3% with 4+ years of college), employed full time (52.6%), with a household income of $100,000 or more (27.4%) in a two person household (35.1%) without children (72.6%). Over 60% of respondents were responsible for all of the grocery shopping and were weekly or daily protein beverage consumers (56.3%) who consumed ready-to-mix (RTM)

(67.9%), refrigerated ready-to-drink (RTD) (64.4%), and shelf-stable RTD (57.3%) protein beverages. These participants also reported reading the nutrition and ingredient labels often (35.8%) or always (45.9%) (Table 5.1).

5.3.2 Purchase Factors for Protein Beverages

Consumers purchased protein beverages for a variety of usage occasions, but they were primarily consumed as a convenient means to boost protein consumption when time and effort were limited (Figure 5.2). When consumers rated the importance of various package messaging attributes on a 5-point importance scale, the most important attributes were recognizable ingredients (4.1) and plain language and packaging (4.1) ($p>0.05$) (Table 5.2). Consumers showed neutral importance to messaging about products being local, or without genetic modification/engineering or GMOs. This result suggests that consumers are most concerned with the general composition of the products (the actual ingredients listed on the product). When asked about the importance of attributes related to the composition of ingredients, the most important attribute was the amount of protein (4.3) ($p<0.05$), followed by the type of protein (4.0), type of sweetener (3.9), and complete protein (3.9) which were not different from each other in perceived importance (Table 5.3, Table 5.4). This result is consistent with other research which reported protein content as a top factor that was important to all consumers when evaluating protein beverages and high protein products (Childs et al., 2008; Oltman et al., 2015). Consumers in focus groups echoed this sentiment, associating protein beverages with a high protein content that they can get quicker than other sources (Figure 5.2). These results suggest that protein is the most important ingredient, which is logical since these beverages are consumed with the primary motivation of adding protein to the diet.

For most consumers, these beverages should have 20-30g of protein per serving, and "if it has less than that I don't even think about [buying] it" (Focus group participant). This result is consistent with research on protein fortified products, where consumers desired 20 - 29g protein per serving (Harwood and Drake, 2019). Consumers associated protein with muscle building and satiety (Oltman et al., 2015), which motivates their decision to consume high protein beverages. Consumers in focus groups agreed that they look towards protein beverages as a healthy and convenient means to keep them full throughout the day (Figure 5.2). The importance of protein type and complete protein mentioned in this study (Table 5.3) has been explored further in literature. Research by Keefer et al. (2024) evaluated the factors that influenced consumer motivations for protein choice in an online survey and found that consumers identified complete protein as a positive influence for protein choice. However, consumers also showed a preference for plant-based proteins over dairy-based proteins (Keefer et al., 2024), even though dairy proteins are complete proteins while most plant proteins are not (Gorissen et al., 2018). Consumers may like the perception of complete protein but may not fully understand what makes a protein a complete protein. Consumer preferences for specific types of protein will be discussed further later in the study.

While added protein is expected in these beverages, some consumers were confused as to why there were so many additional ingredients in many protein beverages. Consumers in focus groups questioned what the rationale was for adding all of these ingredients, but suspected there was a reason, otherwise companies wouldn't go to the trouble of doing so (Figure 5.2). Because the intent of this research was not to educate or persuade consumers on the reasons ingredients are used, no information was provided to consumers about any ingredients. Some consumers hypothesized that all of the non-protein ingredients were added for a reason and were necessary to deliver the high protein content in a palatable way: "You are artificially adding more protein than the normal amount in that food, so you have to add something. You have to balance it out to make it more like you want it to be" (Focus group participant), "I just don't see a protein drink as something that can be minimally processed because these are things that are just not naturally found" (Focus group participant) (Figure 5.2). However, even though some consumers justified the added ingredients, all consumers viewed the prospect of a clean label as a positive characteristic. The sentiment across focus groups was that "if they can accomplish fewer ingredients with the same standards of a protein beverage, I would love that." (Focus group participant). "I want to feel good about what I put into my body…If there was an option that gave me the same effect as [protein beverages with long ingredient lists] but made me feel better about what it is…I'd buy it" (Focus group participant).

When consumers were asked to look at a protein beverage label and assess the ingredient list, consumers responded with "I don't know because I can't interpret this… if I don't have to Google as many things it's better" (Focus group participant) (Figure 5.2). This result was consistent with previous research which suggested that consumers favored products with understandable, short, known, and simple ingredient lists because all of these things reduce the cognitive effort needed for consumers to assess a product (Asioli et al., 2017). Consumers in the current study expressed that the type of ingredients (protein, sweetener, etc.) were important when reading the label for a protein beverage (Table 5.3, Table 5.4), and the following sections will discuss consumer perception of the ingredients *within* classes, as well as *between* classes.

5.3.3 Comparing Ingredients Within Classes of Ingredients (MXD Scaling)

Results from the MXD exercises show the relative acceptability of the protein beverage ingredients *within* each class. Figure 5.3 shows the MXD results for protein ingredients. As a whole, consumers placed the highest utility on plant proteins, followed by whey protein and milk protein. All these ingredients were above the anchor point (“If I saw this ingredient on a label, I would still consider buying it”), suggesting that none of these ingredients were completely unacceptable to consumers. While consumers liked the idea of “plant protein”, consistent with Keefer et al. (2024), individual plant proteins (soy, pea, etc.) received lower utility scores than the general category “plant protein”. This is worth noting because “plant protein” is not an acceptable ingredient designation on a label because it does not identify the food source of the protein in sufficient detail (21CFR102.22). In addition, some protein ingredients differed only by their prefixes, and the utility scores assigned to these different ingredients suggests consumers are wary of ingredients they are not familiar with. For example, hydrolyzed proteins received lower utility scores than their non-hydrolyzed counterparts (whey protein = 49.0, hydrolyzed whey protein = 26.2; soy protein = 33.9, hydrolyzed soy protein = 16.6, $p<0.05$). In addition, sodium had a negative connotation. This is evident by the fact that sodium caseinate received a utility score of 6.4, while calcium caseinate (17.0) and casein protein (19.8) received higher utility scores ($p<0.05$). This result is likely due to the association of sodium with undesirable health conditions such as high blood pressure, hypertension, etc. (Grillo et al., 2019), while calcium is associated with building strong bones (Vannucci et al., 2018).

Casein is a dairy protein that is found in concentrations 4x that of whey protein in bovine milk (Roy et al., 2020), but it received lower utility scores ($p<0.05$) than whey protein, dairy protein or milk protein (Figure 5.3). Keefer et al. (2024) also reported that consumers had a more positive perception (healthier, more sustainable, more ethical, more natural, easier to digest, better taste) of whey protein compared to casein, possibly due to greater familiarity with whey protein over casein. Research by Schiano and Drake (2021) explored consumer knowledge about fluid milk and cheese processing and composition and found that consumers were largely uneducated about the composition of dairy foods. Thirty five percent of surveyed consumers indicated that whey came from dairy products (25% thought it was a plant-based product), 90% of consumers indicated that they knew what whey protein was and 33% associated it with a fitness or muscle building supplement (Schiano and Drake, 2021). Thus, while consumers may be uneducated about dairy proteins, they are at most familiar with whey protein which may contribute to the high acceptability of whey protein as a source of protein.

Among sweeteners, consumers assigned the highest utility to naturally sweetened, followed by agave nectar and then monk fruit (Figure 5.4). The lowest utility was assigned to acesulfame potassium (Ace K), which was the only ingredient that fell below the anchor point, suggesting that consumers have a strong negative perception towards this ingredient. These results were similar to previous research which found that consumers preferred natural sweeteners to artificial sweeteners (Bearth et al., 2014; Oltman et al., 2015; Bearth et al., 2016). As a whole, consumers who drink protein beverages want their beverages to have minimal to zero sugar, but they also want natural sweeteners that they recognize and don't have off flavors (Olsen, 2022). Previous research has observed a similar trend, where consumers preferred their protein fortified products to be naturally sweetened–with a preference for sugar over non nutritive/low calorie options–while at the same time showing a preference for low carb/low sugar (Harwood and Drake, 2019). These aspirations from consumers are often at odds with one another, as many natural and recognizable sweeteners (agave nectar, cane sugar, coconut nectar, etc.) contain high calories, while natural zero calorie sweeteners (stevia, monk fruit) have been shown to elicit metallic flavor and/or bitter taste in dairy matrices (Parker et al., 2018). A similar result was found when sweeteners were evaluated in yogurt conceptually through a survey, and through blind consumer tasting (Crown et al., 2024). Despite conceptual liking of stevia sweetened yogurt, consumers disliked the taste of stevia sweetened yogurt, and a trained panel noted bitter and metallic taste associated with stevia sweetened samples (Crown et al., 2024).

Despite these conflicting aspirations, high conceptual appeal suggests that sweetening with natural low calorie sweeteners may be feasible if the off flavors and aftertastes can be mitigated. Literature has suggested that using blends of these sweeteners in protein beverages can be effective in reducing off flavors and increasing acceptability (Parker et al., 2018). In addition, across all focus groups there were consumers who felt that all commercial RTD protein beverages were overly sweet and flavored, and they were interested in a beverage that was less sweet/flavored (Figure 5.2). From a sweetener perspective, this begs the question of whether or not there is opportunity for protein beverages that are more neutral in flavor and/or are lower in sweetness intensity to appeal to this demographic.

Among thickeners, vegetable fiber received the highest utility score. In addition, a general trend was observed where fibers received higher utility scores than flours or starches, and gums/gels received the lowest utility scores (Figure 5.5). Previous research observed a similar trend where starches were perceived as healthier than gums (Varela and Fiszman, 2013). This may be due to the positive association of fiber with gut health and being a part of a healthy diet (Rani et al., 2018), while gums and gels tend to have a negative perception and were described by focus group participants as something that might "clog you up" (Focus group participant) (Figure 5.2).

Among stabilizers, sunflower lecithin received the highest utility score by a large margin, and sodium hexametaphosphate received the lowest utility score (Figure 5.6). Trends were observed based on the types of ingredients and their prefixes. In general, consumers tended to assign higher utility to citrates compared to phosphates. This may be due to a more positive association with citrates (citrus, citric acid, etc.) compared to phosphates. In addition, ingredients with potassium and magnesium received higher utility than ingredients with sodium. This trend was also observed among proteins (Figure 5.3) and is likely due to the negative association of sodium with diseases like hypertension and high blood pressure (Grillo et al., 2019). Utility scores also decreased as higher order prefixes (tri-, di-, etc.) were assigned to ingredients (utility score for potassium phosphate > dipotassium phosphate > tripotassium phosphate). Consumers may perceive these ingredients as more "complex", "scientific", etc. and thus less desirable (Varela and Fiszman, 2013). Consumers in focus groups agreed that they wanted to know what they were ingesting, and it was "scary" not knowing what was in their foods/beverages (Figure 5.2).

5.3.4 Sustainable, Natural, Healthy, Clean Label and Familiar Ingredients

Following MXD scaling exercises, a CATA question was asked which presented a grid of ingredients and the terms: sustainable, natural, healthy, clean label, familiar. A correspondence analysis (CA) plot was used to visualize these variables along with the frequency response from the anchor question (If I saw this ingredient on a label, I would not buy it) in a two-dimensional space. Figure 5.7 shows that the terms sustainable, natural, clean label, and healthy were not well differentiated from each other by consumers. Previous literature also reported cognitive overlap in the way consumers defined the terms "sustainable", "natural", "healthy" as they pertained to dried dairy ingredients (Schiano et al., 2021). These terms loaded negatively on the x-axis, while "would not buy" loaded positively on the x-axis. This suggests that these terms are perceived positively by consumers. Familiarity loads positively on the y-axis suggesting that familiarity is not strongly correlated with any of the other terms–just because consumers are familiar with an ingredient does not mean that they necessarily perceive it as sustainable, natural, clean label, or healthy.

Looking at the placement of each class of ingredients on the CA biplot (Figure 5.7), proteins and sweeteners consisted of polarizing ingredients including many that consumers did not want to buy, and others that they did want to buy and perceived as sustainable, healthy, natural, and clean label (Figure 5.7). Stabilizers and thickeners also exhibited a range in perception, but were more closely grouped than proteins or sweeteners. Previous research found that stabilizers and thickeners were perceived as less natural than other classes of ingredients (Maruyama et al., 2021), and these findings are consistent with the current study, though some thickeners and stabilizers, especially those including plant terms (vegetable, sunflower, pea, etc.), loaded in close proximity to positive terms (sustainable, natural, healthy, clean label) (Figure 5.7).

Ingredients that received low utility scores in MXD exercises were more closely associated with the "would not buy" attribute. For example, Ace K and allulose were correlated with "would not buy" in Figure 5.7 and received low utility scores through MXD (Figure 5.4). The same could be said for bovine collagen peptides and sodium caseinate (Figure 5.3), gellan gum, inulin, carrageenan gum (Figure 5.5), datem, and sodium hexametaphosphate (Figure 5.6). Overall, the grouping of individual ingredients on this biplot (Figure 5.7) validates the utility scores from the MXD exercises (Figure 5.3, Figure 5.4, Figure 5.5, Figure 5.6) and suggests that there are differences in consumer preference for different ingredients both *between* and *within* classes.

5.3.5 Segmenting Consumers by Preference for Protein Source

As discussed earlier, protein was the most important factor for protein beverages, and the other ingredients were viewed as a means to make the protein palatable. This notion was used as the basis for segmenting consumers by their preference for protein via the protein MXD exercise (Figure 5.3). Latent class cluster analysis was performed on these data, and a three-cluster solution was chosen based on leveling off of the Bayesian Information Criterion (Weller et al., 2020) and meaningful interpretation of the clusters. The three clusters were named according to consumer preferences for different types of protein, as indicated by the distribution of utility scores for proteins (Figure 5.8), dairy-based (DB, n=132), plant-based (PB, n=137), and protein is protein (PiP, n=136). The dairy-based (DB) cluster assigned the highest utility to whey protein (65.1), milk protein (56.1), and dairy protein (54.4) (Figure 5.8a). Non-dairy proteins (faba bean, pea, soy, brown rice, etc.) were assigned low utility scores. Similar to the total population, plant protein received a higher utility score than any of the individual plant proteins (Figure 5.3). DB consumers preferred whey proteins to casein proteins, but casein protein still received higher utility scores than any of the individual plant proteins (other than "plant protein"). The exception to this was sodium caseinate, which received a very low utility score, exemplifying consumer aversion towards ingredients with the word sodium.

In contrast to the DB cluster, the plant-based (PB) cluster assigned the highest utility to plant protein (66.2), pea protein (57.8), and pumpkin seed protein (51.6) (Figure 5.8b). All plant proteins were assigned higher utility scores than dairy proteins, apart from whey protein (30.2), which scored higher than hydrolyzed soy protein (25.3). This result further emphasizes the negative perception towards the term "hydrolyzed", and may stem from consumer lack of familiarity with the term hydrolyzed. The protein is protein (PiP) cluster did not have a strict preference for dairy vs plant proteins, assigning the highest utility to milk protein (65.8), dairy protein (64.3), and plant protein (61.6) (Figure 5.8c).

After identifying these three clusters based on their preference for protein, all data were re-analyzed for each cluster. However, interpretation of other survey results by cluster yielded similar interpretation of the results regarding differences in ingredient perception between and within classes of ingredients. Thus, results are only displayed for the total population, not for each individual cluster.

5.3.6 Visualizing Differences Between Classes of Ingredients (Modified Projective Mapping)

While the MXD results capture differences in consumer perception of ingredients *within* each class, the modified projective mapping exercise is useful in visualizing the relative perception of ingredients *between* classes. Figure 5.9 shows the average placement of each ingredient from the modified projective mapping exercise. These ingredients are color coded by class for visualization purposes, but were randomly presented to participants regardless of ingredient class. In general, the less familiar consumers were with an ingredient, the less likely they were to want to see it on the ingredient list (Figure 5.9). Proteins and sweeteners had a wide range of acceptability, with some ingredients in the class undesirable, and others very desirable, as indicated by the large range on the y-axis (Figure 5.9). Thickeners and stabilizers had a narrower range of acceptability, suggesting that these classes of ingredients were less liked as a whole, partly because they are less familiar to consumers. Some consumers in focus groups acknowledged their acceptance of some of these ingredients by hypothesizing that they were probably necessary to help deliver the taste and texture that was important to them in their protein beverages, such as settling out which they have experienced with some beverages (Figure 5.2). However, the majority of consumers expressed that they did not want stabilizers and/or thickeners in their protein beverages because “besides the protein and sugar, everything in it is manufactured” (Focus group participant). Previous studies have found similar perception towards stabilizers and thickeners, with consumers describing them as less natural than other classes of ingredients (Maruyama et al., 2021), and associating them with “industrially processed foods” (Varela and Fiszman, 2013).

As a whole, consumers want to see proteins and sweeteners on the label, they do not want to see stabilizers, and they have a neutral attitude towards thickeners (Figure 5.9). However, within each of these ingredient classes there is large variability in terms of ingredient acceptability (Figure 5.3, Figure 5.4, Figure 5.5, Figure 5.6). When considering consumer perception of protein beverage ingredients, it is important to understand both *within* and *between* class perception to make the most informed decision when it comes to ingredient substitution or reformulation.

5.3.7 Opportunities and Challenges to Formulating a Clean Label Protein Beverage

While consumers are unlikely to purchase a clean label beverage if it compromises on taste or quality, the findings from this research suggest that a clean label offers a competitive advantage over other beverages that do not have a clean label. This begs the question: what options do processors and product developers have when it comes to formulating clean label protein beverages? Two logical solutions come to mind: 1) substituting less-desirable ingredients for ingredients that are more desirable, or 2) removing undesirable ingredients completely. Both solutions pose challenges, and the pros and cons of both approaches will be discussed briefly.

Substituting less-desirable ingredients with ingredients that are more desirable is based on consumer perceptions of ingredients within a class. For example, with this approach, a product developer may choose to utilize vegetable fiber as a thickener rather than carrageenan, or sunflower lecithin as a stabilizer instead of sodium hexametaphosphate based on the results from the MXD exercises (Figure 5.5, Figure 5.6). However, while this approach is logical from a consumer insight standpoint, it has no basis when it comes to the functional role of these additives in the protein beverages. For example, while all the stabilizers in this study were chosen because they are present in commercial protein beverages for their role in stabilizing the protein beverage matrix, the mechanism for doing so is vastly different for different ingredients and they cannot necessarily be substituted for each other 1 for 1. Sunflower lecithin is added to food products because it acts as an emulsifier, stabilizing the protein beverage matrix by emulsifying lipids in an aqueous matrix to prevent separation and ensure a smooth mouthfeel (Pan et al., 2002). On the other hand, sodium hexametaphosphate is added to protein beverages to function as a calcium-chelating agent which helps stabilize the beverage by preventing protein denaturation and aggregation during thermal processing (Rasouli et al., 2019). Thus, while both ingredients serve to help stabilize the protein beverage matrix in some way, they could not be easily substituted for one another with the same sensory and functional results. In addition, it is important to remember that both ingredients belong to an ingredient class that consumers dislike altogether. Thus, even adding the most preferred ingredient from a class sometimes still means adding an undesirable ingredient. Furthermore, consumer perception of food additives is a moving target that is heavily influenced by trends and information that may or may not be based on science. While research has shown that providing information to consumers about the safety of additives and their role in the product increases acceptability (Shim et al., 2011; Bearth et al., 2016), consumers are more sensitive to negative information about food additives than positive information (Zhong et al., 2018) which means that consumers are more likely to be swayed *against* food additives, rather than *for* them. Carrageenan is one such example of this, being labeled by consumers as an unsafe ingredient, despite scientific support for its safety (Bixler, 2017). This issue highlights the power of consumer perception and the notion that additives like thickeners and stabilizers will always pose a risk of negative consumer perception. For these reasons, ingredient substitution may be a difficult task to accomplish with a small margin for actual improvement in consumer perception.

Complete removal of undesirable ingredients–stabilizers, thickeners (Figure 5.9)–provides opportunity for significant improvements in consumer perception but is considerably more difficult to implement. In high protein beverages, one of the key concerns is product stability during processing (Corbo et al., 2014), such as the prevention of sedimentation of insoluble proteins and heat induced protein sedimentation (Özer and Kirmaci, 2010). By attempting to remove most/all the ingredients that help address these issues, processors will need to find another way to prevent these issues, starting by understanding the role that each ingredient plays in beverage stability. With this approach, the source of protein is a critical factor. While plant-based protein sources are gaining popularity in functional beverages, they face processing and sensory challenges (Sethi et al., 2016). Dairy proteins offer better functionality than plant proteins, such as solubility and heat stability (Singh et al., 2022), that may be leveraged in clean label protein beverages.

Considering dairy proteins as an ingredient for high protein beverages, the type of dairy protein is also an important factor. Whey proteins were the most preferred by consumers (Figure 5.3) and common in commercial protein beverages, but whey protein lacks the heat stability for sole source of protein in neutral pH shelf stable or refrigerated beverages and heat denaturation of whey protein in milk protein ingredients results in undesirable sulfur flavors in beverages (Lee et al., 2017; Vogel et al., 2021; Singh et al., 2022). While conceptually less desirable by consumers (Figure 5.3) and likely due to lack of familiarity, research has shown that casein protein in the form of micellar casein concentrate (MCC) may be superior to whey or milk protein in shelf-stable low-acid protein beverages because of increased flavor and functional stability at neutral pH (Vogel et al., 2021; Carter et al., 2021; Whitt et al., 2022; Singh et al., 2022, Hoyt et al., 2023, Pranata et al., 2024). While dairy proteins appear to be a more likely candidate than plant proteins for complete removal of additives due to their superior functionality, more research is needed to better understand protein stability and the factors that affect it when processing high protein beverages.

5.3.8 Relevance for the Dairy Industry

The findings from this research are relevant to the dairy industry because they highlight consumer perception towards dairy ingredients in protein beverages and opportunities for the industry to highlight advantages of dairy proteins over plant proteins. Consumers have positive perceptions of dairy in protein beverages, but many consumers are unsure which proteins come from dairy. Thus, there is opportunity for the dairy industry to educate and familiarize consumers with individual dairy proteins (whey vs casein). In addition, this research showed that consumers desire protein beverages with fewer ingredients on the label. This presents an opportunity for the dairy industry to leverage the superior functionality of dairy proteins compared to plant proteins. Research efforts should seek to optimize processing and formulation of clean label dairy protein beverages.

5.3.9 Limitations and Future Work

The demographic of participants surveyed in this research was skewed heavily towards the population residing near the affiliated university (North Carolina State University) and surrounding areas. In addition, this research is limited in that consumers were asked about their perception of ingredients in isolation, not as they would appear on a commercial food/beverage label. Future research should explore consumer perception of complete ingredients labels as they pertain to protein beverages. Future research should also look to understand the effect of ingredient removal on protein beverage stability during processing and storage, as well as the interaction between ingredient removal, protein type, and sweetener. Consumer acceptance of protein beverage concepts with ingredient substitution versus ingredient removal is also an area of future research. Future research should also consider the effect of processing on the sensory properties of clean label protein beverages, as different heating methods have been shown to have an effect on sensory quality (Hoyt et al., 2023, Rocha et al., 2023). Ultimately, prototyping and acceptance of clean label protein beverages should be evaluated by consumers both with and without information about the label/ingredients.

5.4 Conclusion

The findings from this research exemplify a few key points about protein beverages. For protein beverages, it's all about the protein. While clusters exist with preferences for dairy protein (DB), plant protein (PB), or protein in general (PiP), consumers agree that these beverages are consumed to increase their protein intake. Sweeteners are acceptable ingredients to make the beverages more palatable, and consumers prefer natural sweeteners and minimal calories. When it comes to thickeners and stabilizers, there are ingredients within these classes that are more preferred than others, but consumers would prefer if these ingredients were removed entirely. Future research should aim to address the effect of ingredient substitution and removal on beverage stability and consumer acceptance.

Acknowledgements

Funding provided in part by Dairy West (Meridian, ID) and Dairy Management, Inc. (Rosemont, IL).

References

Aschemann-Witzel, J., P. Varela, and A. O. Peschel. 2019. Consumers' categorization of food ingredients: Do consumers perceive them as 'clean label' producers expect? An exploration with projective mapping. Food Qual Prefer. 71:117–128. https://doi.org/10.1016/j.foodqual.2018.06.003

Asioli, D., J. Aschemann-Witzel, V. Caputo, R. Vecchio, A. Annunziata, T. Næs, and P. Varela. 2017. Making sense of the "clean label" trends: A review of consumer food choice behavior and discussion of industry implications. Food Res Int. 99:58–71. https://doi.org/10.1016/j.foodres.2017.07.022

Bartelme, M. Z., A. Beckett, D. Faulkner, J. Ng, and J. Zegler. 2024. Global Food & Drink Trends. Mintel. https://www.mintel.com/

Bearth, A., M. E. Cousin, and M. Siegrist. 2014. The consumer's perception of artificial food additives: Influences on acceptance, risk and benefit perceptions. Food Qual Prefer. 38:14–23. https://doi.org/10.1016/j.foodqual.2014.05.008

Bearth, A., M. E. Cousin, and M. Siegrist. 2016. "The Dose Makes the Poison": Informing Consumers About the Scientific Risk Assessment of Food Additives. Risk Analysis. 36(1):130–144. https://doi.org/10.1111/risa.12410

Bixler, H. J. 2017. The carrageenan controversy. J. Appl Phycol. 29(5):2201–2207. https://doi.org/10.1007/s10811-017-1132-4

Cao, Y., and L. Miao. 2023. Consumer perception of clean food labels. British Food Journal. 125(2):433–448. https://doi.org/10.1108/BFJ-03-2021-0246

Carter, B.G., N. Cheng, R. Kapoor, G. H. Meletharayil, and M. A. Drake. 2021. Invited review: Microfiltration-derived casein and whey proteins from milk. J. Dairy Sci. 104(3):2465-2479. https://doi.org/10.3168/jds.2020-18811

Childs, J. L., J. L. Thompson, J. S. Lillard, T. K. Berry, and M. A. Drake. 2008. Consumer perception of whey and soy protein in meal replacement products. J. Sens Stud. 23(3):320–339. https://doi.org/10.1111/j.1745-459X.2008.00158.x

Chychula, A. 2024. Nutrition Drinks – US. Mintel. https://www.mintel.com/

Corbo, M. R., A. Bevilacqua, L. Petruzzi, F. P. Casanova, and M. Sinigaglia. 2014. Functional Beverages: The Emerging Side of Functional Foods: Commercial Trends, Research, and Health Implications. Compr Rev Food Sci Food Saf. 13(6):1192–1206. https://doi.org/10.1111/1541-4337.12109

Crown, E., D. Rovai, C. M. Racette, D. M. Barbano, and M. A. Drake. 2024. Consumer perception of sweeteners in yogurt. J. Dairy Sci. In press.

Gorissen, S., J. Crombag, J. Senden, W. Waterval, J. Bierau, L. Verdijk, and L. van Loon. 2018. Protein content and amino acid composition of commercially available plant-based protein isolates. Amino Acids. 50(12):1685-1695. https://doi.org/10.1007/s00726-018-2640-5

Grillo, A., L. Salvi, P. Coruzzi, P. Salvi, and G. Parati. 2019. Sodium intake and hypertension. Nutrients. 11(9). https://doi.org/10.3390/nu11091970

Harwood, W. S., and M. A. Drake. 2019. Understanding implicit and explicit consumer desires for protein bars, powders, and beverages. J. Sens Stud. 34(3). https://doi.org/10.1111/joss.12493

Howell, J. 2009. CBC/HB for beginners. Sawtooth Software Research Paper Series. Sawtooth

Software, Sequim, WA. https://sawtoothsoftware.com/resources/technical-papers/cbc-hb-for-beginners

Hong, J. H., H. S. Park, S. J. Chung, L. Chung, S. M. Cha, S. Lê, and K. O. Kim. 2014. Effect of familiarity on a cross-cultural acceptance of a sweet ethnic food: A case study with Korean traditional cookie (Yackwa). J. Sens Stud. 29(2):110–125. https://doi.org/10.1111/joss.12087

Hoyt, H., J. Pranata, D. Barbano, and M. A. Drake. 2023. Effect of dipotassium phosphate and heat on milk protein beverage viscosity and color. J. Dairy Sci. 106(6):3884-3899. https://doi.org/10.3168/jds.2022-22887

Keefer, H., C.M. Racette, and M. A. Drake. 2024. Factors influencing consumer motivations for protein choice. J Food Sci. 89(1):596-613. https://doi.org/10.1111/1750-3841.16805

Lattery, K. 2011. Anchoring Maximum Difference Scaling Against a Threshold-Dual Response and Direct Binary Responses. www.sawtoothsoftware.com

Lee, J. S., J. M. Park, S. H. Wi, Y. B. Ahn, N. K. Kim, K. W. Moon, C. Y. Yang, and J. M. Kim. 2014. Improving consumer recognition and awareness of food additives through consumer education in South Korea. Food Sci Biotechnol. 23(2):653–660. https://doi.org/10.1007/s10068-014-0089-1

Lindsay, R. C. 2008. Food Additives. Pages 689-749 in Fennema's Food Chemistry (4th ed.). CRC Press. ISBN-13: 978-0-8493-9272-6.

Maruyama, S., N. A. Streletskaya, and J. Lim. 2021. Clean label: Why this ingredient but not that one? Food Qual Prefer. 87. https://doi.org/10.1016/j.foodqual.2020.104062

Mills, J. 2023. Yogurt and yogurt drinks – US – 2024. Mintel. https://www.mintel.com/

Mills, J. 2024a. Milk and non-dairy milk – US – 2024. Mintel. https://www.mintel.com/

Mills, J. 2024b. Cheese – US – 2024. Mintel. https://www.mintel.com/

Olsen, K. 2022. Sugars and Alternative Sweeteners – US. Mintel. https://www.mintel.com/

Oltman, A. E., K. Lopetcharat, E. Bastian, and M. A. Drake. 2015. Identifying Key Attributes for Protein Beverages. J. Food Sci. 80(6):S1383–S1390. https://doi.org/10.1111/1750-3841.12877

Orme, B. 2009. Maxdiff analysis: Simple counting, individual-level logit, and hb. Sawtooth Software Research Paper Series. Sawtooth Software, Sequim, WA.

Orme, B. 2019. Consistency Cutoffs to Identify "Bad" Respondents in CBC, ACBC, and MaxDiff. Sawtooth Software Research Paper Series. Sawtooth Software, Sequim, WA.

Osborn, S. 2015. Labelling relating to natural ingredients and additives. Pages 207-221 in Advances in Food and Beverage Labelling: Information and Regulations. Elsevier Inc. https://doi.org/10.1533/9781782420934.3.207

Özer, B. H., and H. A. Kirmaci. 2010. Functional milks and dairy beverages. Int J. Dairy Technol. 63(1):1–15. https://doi.org/10.1111/j.1471-0307.2009.00547.x

Pan, L. G., M. C. Tomás, and M. C. Añón. 2002. Effect of sunflower lecithins on the stability of water-in-oil and oil-in-water emulsions. J. Surfactants Deterg. 5(2):135–143. http://dx.doi.org/10.1007/s11743-002-0213-1

Parker, M. N., K. Lopetcharat, and M. A. Drake. 2018. Consumer acceptance of natural sweeteners in protein beverages. J. Dairy Sci. 101(10):8875–8889. https://doi.org/10.3168/jds.2018-14707

Pranata, J., H. Hoyt, M.A. Drake, and D. Barbano. 2024. Effect of dipotassium phosphate addition and heat on proteins and minerals in milk protein beverages. J. Dairy Sci. 107(2):695-719. https://doi.org/10.3168/jds.2023-23768

Rani, V., and U. C. S. Yadav. 2018. Functional food and human health. In Functional Food and Human Health. Springer Singapore. https://doi.org/10.1007/978-981-13-1123-9

Rasouli, M., S. Abbasi, F. Azarikia, and R. Ettelaie. 2020. On the heat stability of whey protein: Effect of sodium hexametaphosphate. Int J. Dairy Technol. 73(1):46–56. https://doi.org/10.1111/1471-0307.12626

Rocha, R. S., B. Mahieu, E. R. Tavares Filho, P. B. Zacarchenco, M. Q. Freitas, E. T. Mársico, T. C. Pimentel, E. A. Esmerino, and A. G. Cruz. 2023. Free comment as a valuable approach to characterize and identify the drivers of liking of high-protein flavored milk drink submitted to ohmic heating. Food Res Int. 165:112517. https://doi.org/10.1016/j.foodres.2023.112517

Roy, D., A. Ye, P. J. Moughan, and H. Singh. 2020. Composition, Structure, and Digestive Dynamics of Milk From Different Species—A Review. Front Nutr. 7:577759. https://doi.org/10.3389/fnut.2020.577759

Sawtooth Software. 2021. Sawtooth Software technical paper series: The latent class technical paper. Version 4.8. https://sawtoothsoftware.com/resources/technical-papers/latent-class-technical-paper

Schiano, A. N., P. D. Gerard, and M. A. Drake. 2021. Consumer perception of dried dairy ingredients: Healthy, natural, and sustainable?. J. Dairy Sci. 104(12):12427–12442. https://doi.org/10.3168/jds.2021-20589

Schiano, A. N., and M. A. Drake. 2021. Consumer understanding of fluid milk and cheese processing and composition. J. Dairy Sci. 104:86644-8660. https://doi.org/10.3168/jds.2020-20057

Sethi, S., S. K. Tyagi, and R. K. Anurag. 2016. Plant-based milk alternatives an emerging segment of functional beverages: a review. J. Food Sci Technol. 53(9):3408–3423. https://doi.org/10.1007/s13197-016-2328-3

Shim, S. M., S. H. Seo, Y. Lee, G. I. Moon, M. S. Kim, and J. H. Park. 2011. Consumers' knowledge and safety perceptions of food additives: Evaluation on the effectiveness of transmitting information on preservatives. Food Control. 22(7):1054–1060. https://doi.org/10.1016/j.foodcont.2011.01.001

Singh, R., G. Rathod, G. H. Meletharayil, R. Kapoor, V. M. Sankarlal, and J. K. Amamcharla. 2022. Invited review: Shelf-stable dairy protein beverages—Scientific and technological aspects. J. Dairy Sci. 105(12) :9327–9346. https://doi.org/10.3168/jds.2022-22208

Vannucci, L., C. Fossi, S. Quattrini, L. Guasti, B. Pampaloni, G. Gronchi, F. Giusti, C. Romagnoli, L. Cianferotti, G. Marcucci, and M. L. Brandi. 2018. Calcium Intake in bone health: A focus on calcium-rich mineral waters. Nutrients. 10(12):1930. https://doi.org/10.3390/nu10121930

Varela, P., and S. M. Fiszman. 2013. Exploring consumers' knowledge and perceptions of hydrocolloids used as food additives and ingredients. Food Hydrocoll. 30(1):477–484. https://doi.org/10.1016/j.foodhyd.2012.07.001

Vogel, K.G., B. G. Carter, N. Cheng, D. M. Barbano, and M. A. Drake. 2021. Ready-to-drink protein beverages: Effects of milk protein concentration and type on flavor. J. Dairy Sci. 104(10):10640-10653. https://doi.org/10.3168/jds.2021-20522

Weller, B. E., N. K. Bowen, and S. J. Faubert. 2020. Latent Class Analysis: A Guide to Best Practice. J Black Psychol. 46(4):287–311. https://doi.org/10.1177/0095798420930932

Whitt, D., J. Pranata, B. Carter, D. Barbano, and M. A. Drake. 2022. Effects of micellar casein concentrate purity and milk fat on sulfur/eggy flavor in ultrapasteurized milk-based beverages. J. Dairy Sci. 105(7):5700-5713. https://doi.org/10.3168/jds.2021-21621

Zhong, Y., L. Wu, X. Chen, Z. Huang, and W. Hu. 2018. Effects of food-additive-information on consumers' willingness to accept food with additives. Int J Environ Res Public Health. 15(11):2394. https://doi.org/10.3390/ijerph15112394

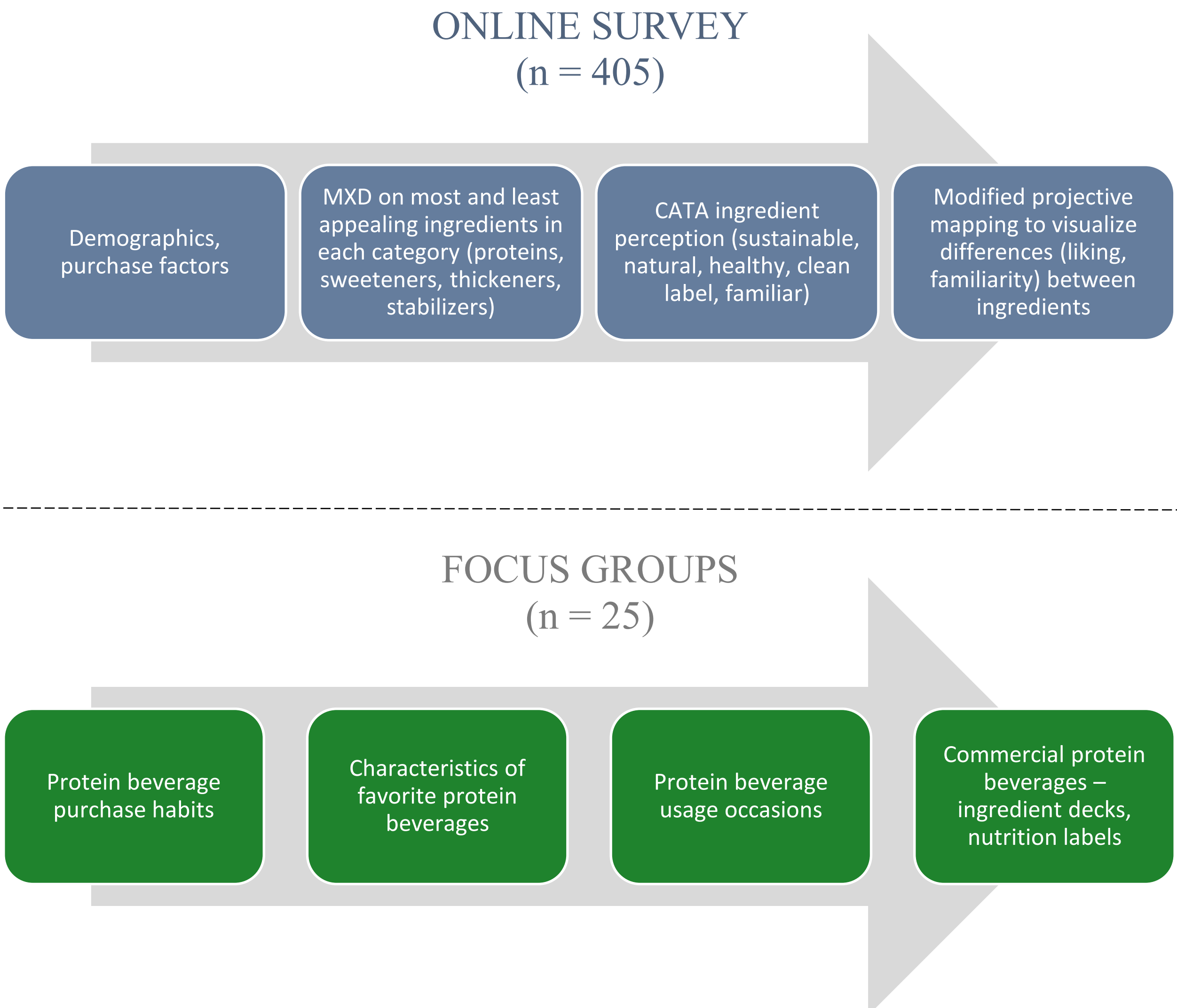


Figure 5.1 Experimental overview

Protein content, convenience, and health drive purchase

"They [protein beverages] offer higher content of protein that I can't get quickly from any other source."

"It is convenient. I can drink it in the morning as breakfast."

"They are very healthy and [the protein content] keeps me full during the day."

Consumers are skeptical about the ingredients on the label

"When they have to have more and more and more [ingredients] in there, I'm trying to figure out what the rationale is… it makes no sense to me."

"I don't know because I can't interpret [the label]… if I don't have to Google as many things it's better."

Protein is good, sugar is okay, everything else is bad

"Besides the protein and sugar, everything in it is manufactured."

"Gums sound like it is going to clog you up."

***Some* consumers think they need to sacrifice a clean label for quality**

"You are artificially adding more protein than the normal amount in that food, so you have to add something. You have to balance it out to make it more like you want it to be."

"I just don't see a protein drink as something that can be minimally processed because these are things that are just not naturally found"

"If you don't like [the beverage settling out], that is because it has less stabilizers. You can't have both."

***All* consumers see a cleaner label as a positive characteristic**

"If they can accomplish fewer ingredients with the same standards of a protein beverage, I would love that."

"I want to feel good about what I put into my body…If there was an option that gave me the same effect as [protein beverages with long ingredient lists] but made me feel better about what it is…I'd buy it."

"I look for ingredients that I know what they are… I want to know what I am ingesting…If I can't recognize exactly what it is, then it is scary."

Interest in protein beverages that aren't overly sweet/flavored

"Almost all the protein beverages are sugary and you can't get away from that, but I try to pick the ones that are less sweet and less flavored."

Figure 5.2 Themes that emerged from four 120-minute focus groups (n=25). Themes in gray represent consumer sentiment that was consistent across multiple focus groups. Select quotes from consumers support the identified themes in the consumer voice.

Table 5.1 Demographics of protein beverage consumers surveyed (n=405).

Question	Attribute	Mean
Age	Gen Z (18 - 27)	35.1%
	Millennial (28 - 42)	33.3%
	Gen X (43 - 57)	18.0%
	Baby Boomers (58 - 76)	13.3%
	Silent Generation (77 - 97)	0.2%
Gender	Male	26.2%
	Female	71.6%
	Other	2.2%
Ethnicity	White/Caucasian	62.5%
	Black/African American	11.6%
	Hispanic/Latino	7.9%
	East Asian	4.2%
	South Asian or Indian	8.1%
	Middle Eastern	2.0%
	Native American/Pacific Islander	0.2%
	Other (please specify)	2.2%
	Prefer not to answer	1.2%
State	North Carolina	84.7%
	California	3.2%
	New Jersey	1.5%
	Florida, Illinois, New York (1.0% each)	3.0%
	Colorado, Minnesota, Texas (0.7% each)	2.1%
	Hawaii, Idaho, Michigan, Pennsylvania, Washington, Wisconsin (0.5% each)	3.0%
	Arizona, Connecticut, Delaware, Georgia, Louisiana, Nevada, Ohio, South Carolina, Utah, Virginia (0.2% each)	2.5%
Education	Some high school	0.0%
	High school diploma or equivalent	8.9%
	2 years of college (i.e. Associate's degree)	12.1%
	4 years of college (i.e. Bachelor's degree)	37.8%
	5 or more years of college (i.e. Master's or PhD)	40.5%
	Technical/Trade/Vocational school	0.7%
Employment Status	Employed full-time	52.6%
	Employed part-time	13.3%
	Student	22.0%
	Homemaker	3.7%
	Retired	5.2%
	Other (please specify)	3.2%
Income	$24,999 or less	14.6%
	$25,000 - $49,999	15.3%
	$50,000 - 74,999	18.3%
	$75,000 - $99,999	17.8%
	$100,000 or more	27.4%
	Prefer not to answer	6.7%
Household Size	1 person	24.4%
	2 person	35.1%
	3 person	16.0%
	4 person	16.0%
	5+ person	8.4%
Children in Household	Yes	27.4%
	No	72.6%
Grocery Shopping	100% - I do all of the grocery shopping	60.7%
	75% - I do most of the grocery shopping	20.7%
	50% - I share equally the grocery shopping	14.8%
	25% - I do some of the grocery shopping, but I am not the primary shopper for my household	3.7%
	0% - I do none of the grocery shopping	0.0%
Protein Beverage Frequency	Daily	15.8%
	Weekly	40.5%
	Monthly	32.1%
	Yearly	11.6%
	Never	0.0%
Types of Protein Beverages (Check-all-that-apply)	Ready to mix protein beverages	67.9%
	Refrigerated ready to drink protein beverages	64.4%
	Shelf-stable ready to drink protein beverages	57.3%
	Other (please specify)	2.0%
	None of the above	0.0%
Reading Nutrition and Ingredient Labels	I ALWAYS read nutrition and ingredient labels	45.9%
	I OFTEN read nutrition and ingredient labels	35.8%
	I SOMETIMES read nutrition and ingredient labels	15.3%
	I RARELY read nutrition and ingredient labels	2.7%
	I NEVER read nutrition and ingredient labels	0.0%

Table 5.2 Package messaging importance for protein beverages

Attribute	Mean
Recognizable ingredients	4.1a
Plain language and packaging	4.1a
No antibiotics or hormones	3.7b
Fewer ingredients	3.5bc
Recognizable brand	3.4c
Organic	3.2d
Non-scientific ingredients	3.1de
Not genetically modified	3.0def
GMO-free	3.0ef
Not genetically engineered	3.0ef
Not bioengineered	2.9f
Not genetically modified through the use of modern biotechnology	2.9f
Local	2.9f

Data represents n=405 participants. Attributes were scored on a 5-point scale where 1 = not at all important, 5 = very important when reading a package for a protein beverage. Different letters following means signify significant differences ($p<0.05$). Statistical lettering was determined using Kruskal-Wallis with Dunn's post hoc test.

Table 5.3 Ingredient messaging importance for protein beverages.

Attribute	Mean
Amount of protein	4.3a
Type of protein	4.0b
Type of sweetener	3.9b
Complete protein	3.9b
All natural ingredients	3.6c
No added sugar	3.6c
Plant protein	3.5d
No artificial flavorings	3.4d
No food additives	3.4d
Sugar-free or reduced sugar	3.3d
Dairy protein	3.1e
Reduced allergens or allergen-free	3.0e
Lactose free	2.7f

Data represents 405 participants. Attributes were scored on a 5-point scale where 1 = not at all important, 5 = very important when reading a package for a protein beverage. Different letters following means signify significant differences ($p<0.05$). Statistical lettering was determined using Kruskal-Wallis with Dunn's post hoc test.

Table 5.4 Attributes influencing protein beverage purchase intent

Attribute	Mean
Amount of protein	36.2
Type of protein	21.6
Type of sweetener	18.5
Length of ingredient list	12.9
Presence of additives	10.8

Data represents 405 participants. Attributes were evaluated using a constant sum (chip allocation) exercise where participants allocated a total of 100 points across the five attributes specified. Mean point allocation for each attribute is reported.

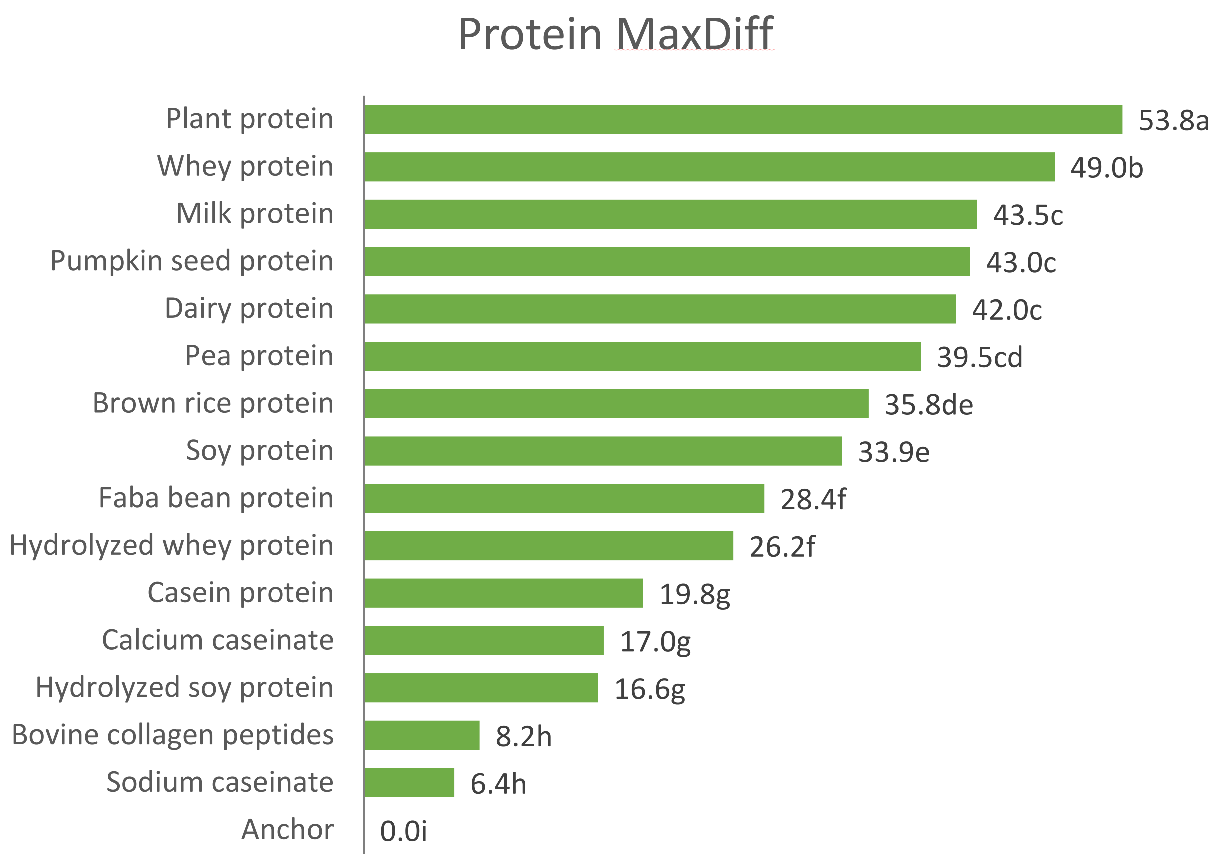


Figure 5.3 MaxDiff results for protein types in protein beverages. Data represents n=405 participants. Values presented are zero-anchored interval scaled responses. Ingredients below the anchor point indicate that consumers would not buy a protein beverage if this ingredient was present. Different letters following means signify significant differences ($p<0.05$). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

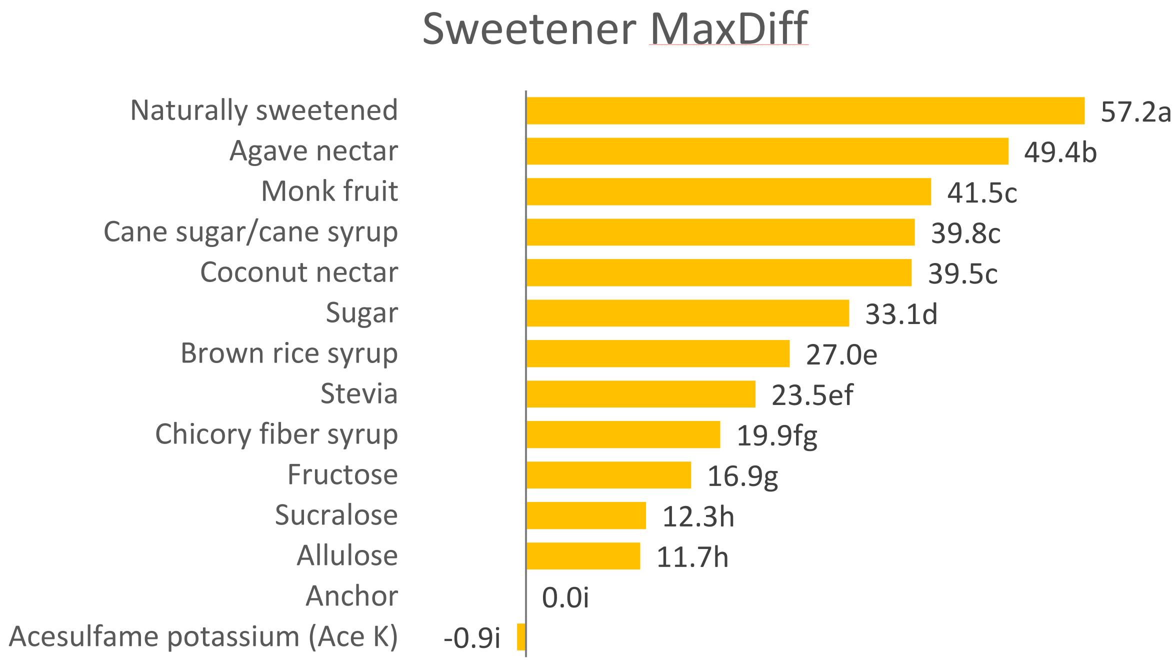


Figure 5.4 MaxDiff results for sweeteners in protein beverages. Data represents n=405 participants. Values presented are zero-anchored interval scaled responses. Ingredients below the anchor point indicate that consumers would not buy a protein beverage if this ingredient was present. Different letters following means signify significant differences ($p<0.05$). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

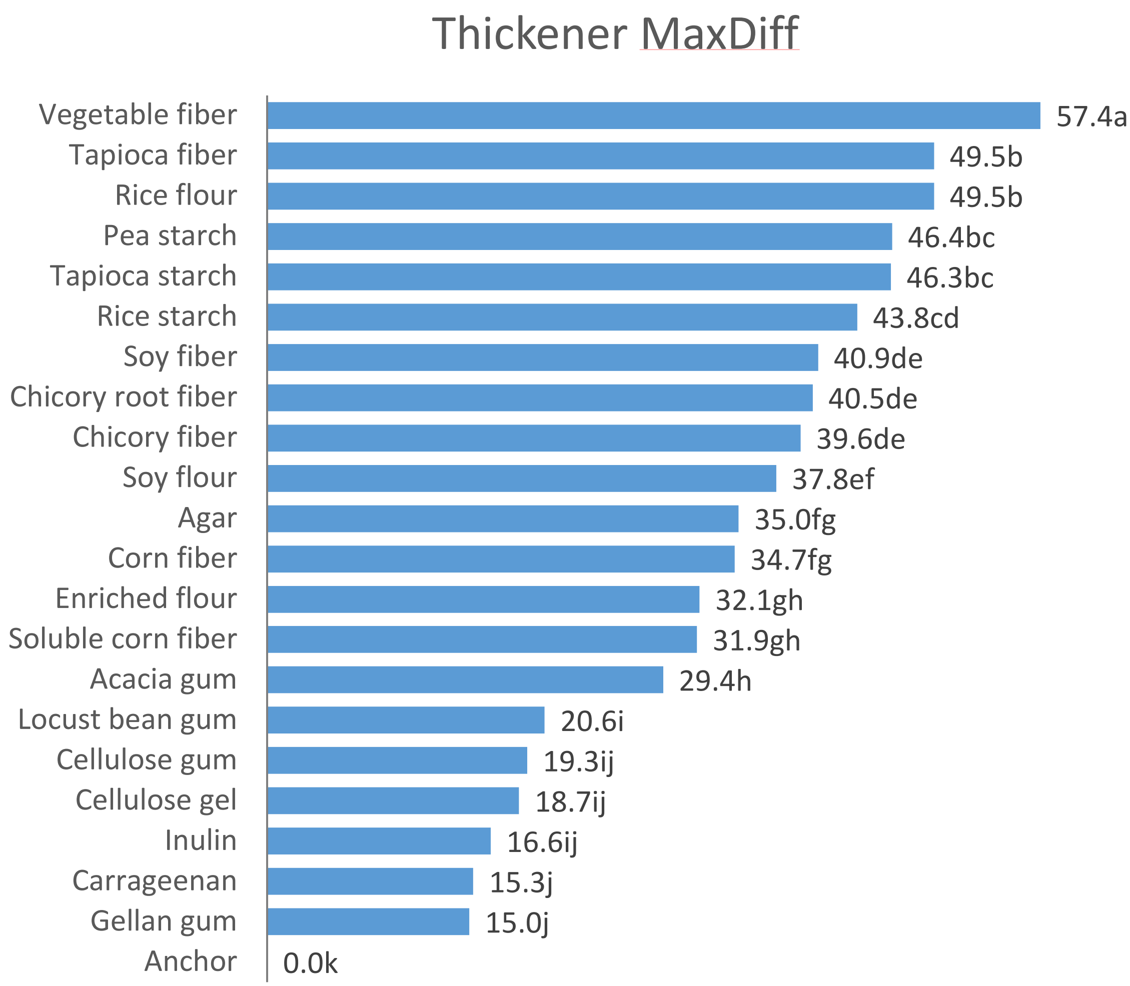


Figure 5.5 MaxDiff results for thickeners in protein beverages. Data represents n=405 participants. Values presented are zero-anchored interval scaled responses. Ingredients below the anchor point indicate that consumers would not buy a protein beverage if this ingredient was present. Different letters following means signify significant differences ($p<0.05$). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

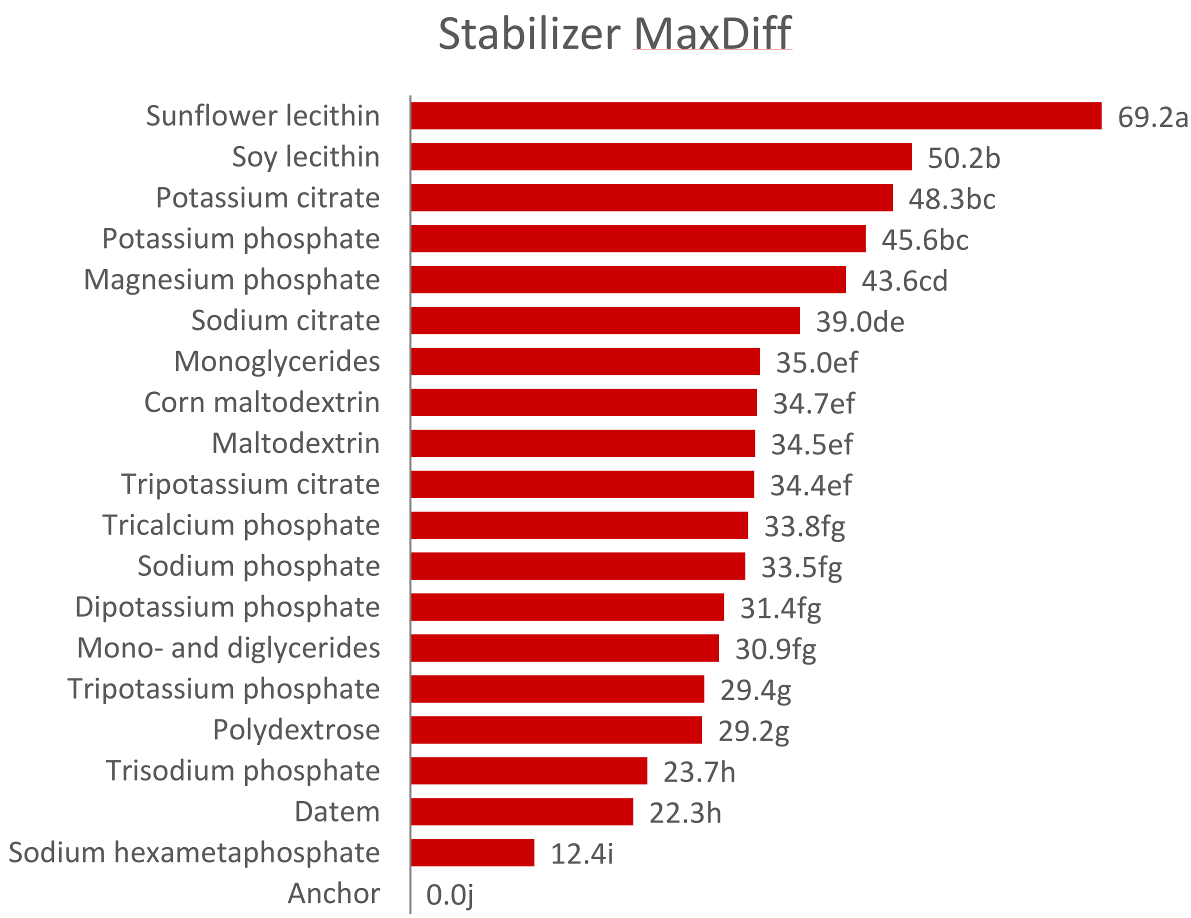


Figure 5.6 MaxDiff results for stabilizers in protein beverages. Data represents n=405 participants. Values presented are zero-anchored interval scaled responses. Ingredients below the anchor point indicate that consumers would not buy a protein beverage if this ingredient was present. Different letters following means signify significant differences ($p<0.05$). Statistical lettering was determined using ANOVA with Fisher's LSD post hoc test.

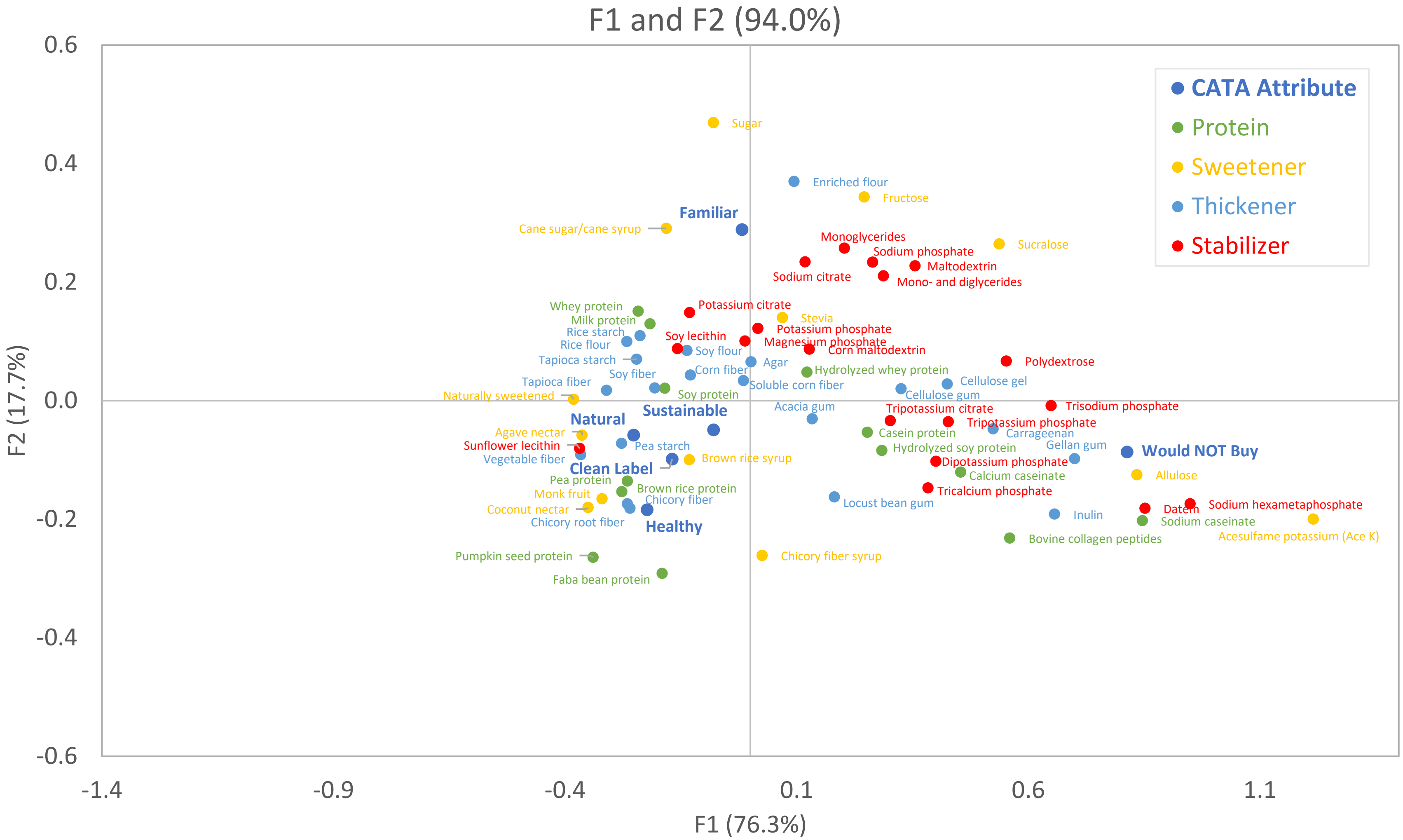


Figure 5.7 Correspondence analysis biplot of CATA terms used to characterize protein beverage ingredients. Colors were added to visually differentiate classes of ingredients.

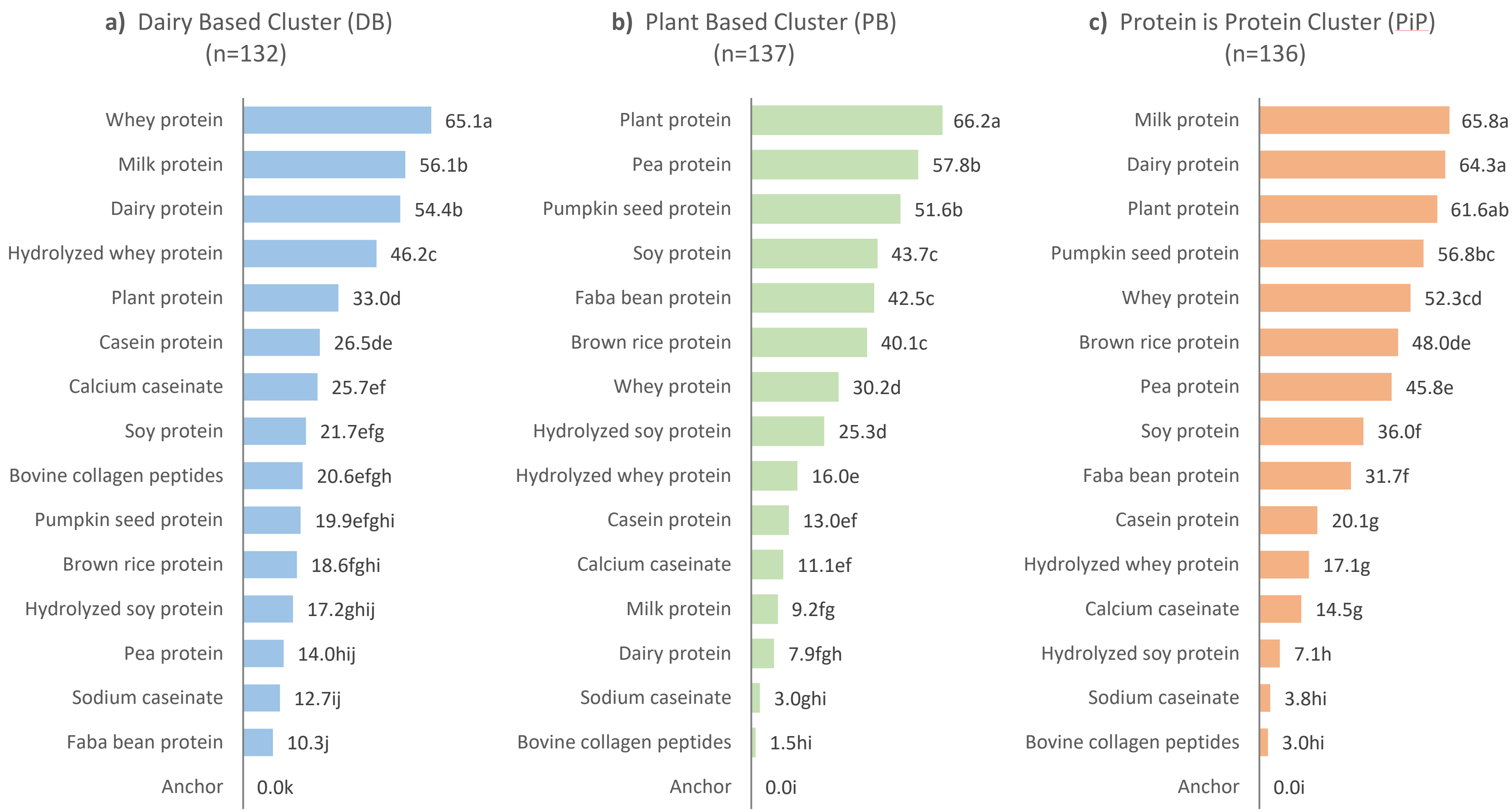


Figure 5.8 Clusters identified by preference for different protein sources in protein beverages. Clusters were identified via Latent Class analysis. Values presented are zero-anchored interval scaled responses. Ingredients below the anchor point indicate that consumers would not buy a protein beverage if this ingredient was present. Mean utility scores for ingredients were compared within each cluster. Different letters following means signify significant differences ($p<0.05$).

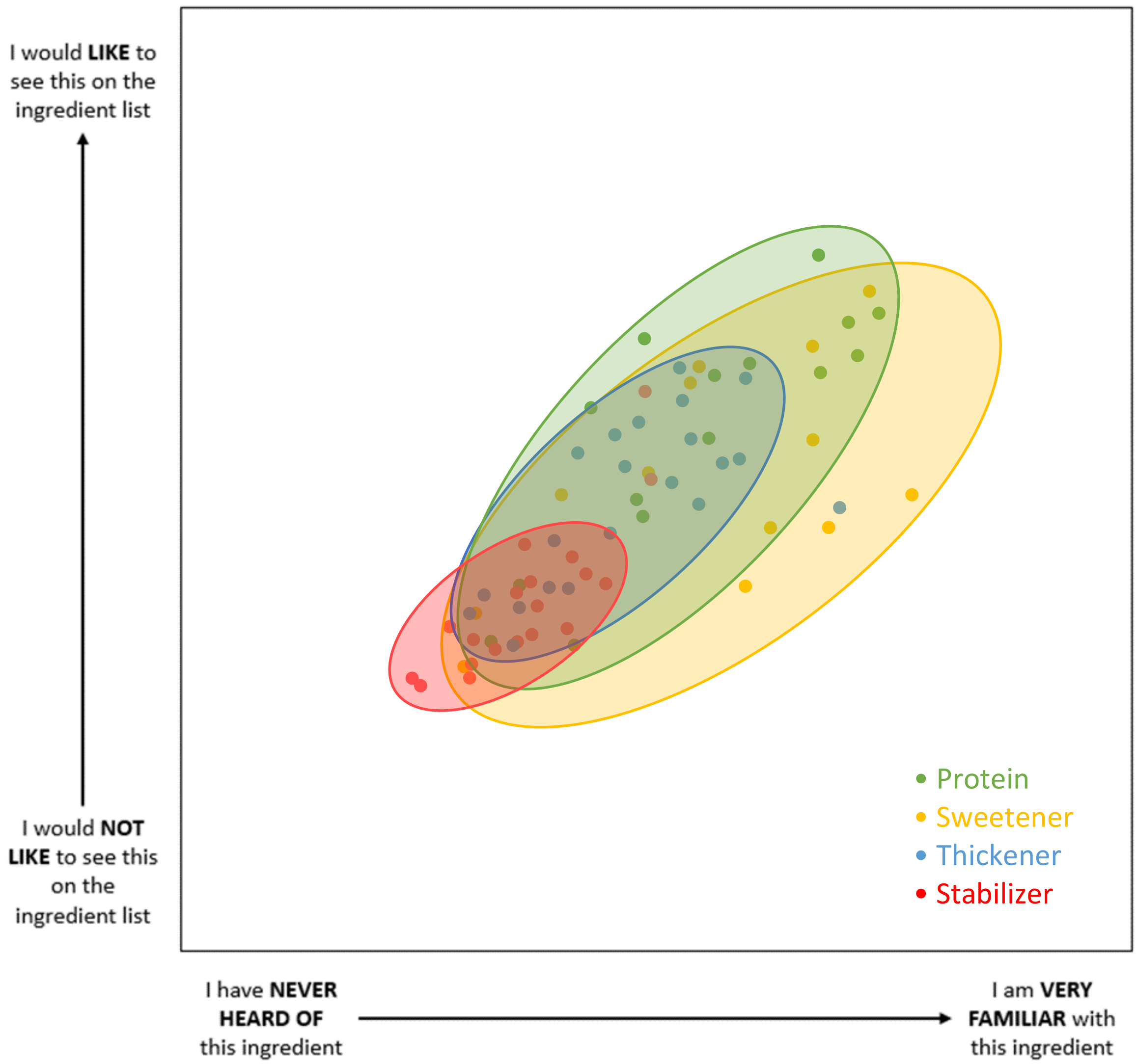


Figure 5.9 Modified projective mapping results for protein beverage ingredients by class. Participants (n=405) were randomly assigned 17 out of 68 ingredients using a balanced incomplete design and asked to place each ingredient on the map in relation to the given anchors. Each point represents the average placement of an ingredient. Colors were added to visually differentiate classes of ingredients.

APPENDICES

Appendix 2.1 Ballot for Oat Milk Survey

EXCLUSION CRITERIA

1. How old are you (years)
   - [Numerical Response] [DISQUALIFY if <18]

2. Please indicate which, if any, of the following products you have consumed in the past 3 months (select all that apply)
   - Coffee
   - Dairy milk
   - Plant-based milk (almond milk, oat milk, soy milk, etc.) [MUST SELECT]
   - Protein-fortified products (protein powder, protein bars, protein beverages)
   - Packaged snacks (chips, crackers, cookies, etc.)
   - Cheese (dairy and/or plant-based)
   - Meat (chicken, beef, etc.)
   - Alcohol
   - Juice
   - Tea
   - None of the above [fixed, exclusive]

3. How often do you consume each of the following types of milk? (frequency grid - never, occasionally, often, very often)

   - Milk (dairy)
   - Oat milk [MUST SELECT Occasionally+]
   - Almond milk
   - Coconut milk
   - Soy milk
   - Lentil milk [DISQUALIFY if Often/Very Often]
   - Other milk (cashew, hazelnut, rice, pea, etc.)
     i. Never
     ii. Occasionally
     iii. Often
     iv. Very Often

4. [Connection - If > never for more than one type] Which of the following describe why you use multiple types of milk? (select all that apply)
   - Some members of my household have dietary restrictions
   - I use them for different uses (coffee, cereal, etc.)
   - I like the flavor variety
   - I like the texture variety
   - I like the nutritional variety

- I like the price variety
- Other (specify)

[Attention Check 1] This is an attention check question to make sure you are reading the questions. Select the option below that says “Neither Like Nor Dislike”. Selecting the wrong response will disqualify you from the survey and you will not receive compensation

- Dislike Very Much
- Dislike Slightly
- Neither Like Nor Dislike
- Like Slightly
- Like Very Much

MILKS COMPARISONS

5. [Sliding scale] Below is a list of different types of milks. Please use a sliding scale for each of the following 5 milks based on the attributes provided.
    - Milks:
        i. Milk (dairy)
        ii. Oat milk
        iii. Almond milk
        iv. Soy milk
        v. Coconut milk
    - Attributes:
        i. Very Sustainable - Not At All Sustainable
        ii. Nutrient Rich (vitamins & minerals) – Not At All Nutrient Rich (vitamins & minerals)
        iii. Lots of Sugar – Hardly any Sugar
        iv. Heart Healthy - Not At All Heart Healthy
        v. Tastes Great – Does Not Taste Great
        vi. Great Texture - Bad Texture
        vii. Very Expensive - Not At All Expensive
        viii. Very Clean Label - Not At All Clean Label
        ix. Short Ingredient List - Long Ingredient List
        x. Made with Familiar Ingredients – Made with Unfamiliar Ingredients
        xi. Allergen Free - Lots of Allergens
        xii. Lots of Protein - Hardly Any Protein

6. Select which of the following contain **COMPLETE PROTEINS**.
    - Milks:
        i. Milk (dairy)
        ii. Oat milk

iii. Almond milk
iv. Soy milk
v. Coconut milk
vi. Egg

- Attributes:
  i. This is NOT a complete protein
  ii. This might be a complete protein (I'm not sure)
  iii. I'm sure this is a complete protein
  iv. I don't know what you mean by complete protein

[Text Screen] Thank you for your feedback. For the rest of the survey, we will be focusing specifically on **OAT MILK**.

OAT MILK CONSUMPTION BEHAVIOR

7. How often do you consume **OAT MILK**?
   - Multiple times per day
   - Once per day
   - A few times per week
   - Once per week
   - A few times per month
   - Once per month
   - Less than once per month

8. Select whether or not you use oat milk in the following ways (Yes/No)
   - Drinking on its own
   - Adding to cereal/oatmeal
   - Adding to smoothie
   - Adding to coffee/tea
   - Using as an ingredient in a recipe
     i. Yes
     ii. No
9. [Piped based on responses for previous Q] For each of the following applications, how often do you typically use **OAT MILK**? (Frequency)
   - Multiple times per day
   - Once per day
   - A few times per week
   - Once per week
   - A few times per month
   - Once per month
   - Less than once per month

10. [Q7 Connection: for each {Attribute} = Yes] When **{Attribute}**, why would you use **OAT MILK OVER OTHER MILK OPTIONS** (other plant-based milks, dairy milk)? (select all that apply)
   - Better taste
   - Better texture
   - Better color
   - More sustainable
   - Healthier
   - Availability
   - Easier to digest
   - Other (specify) [fixed, require explanation]

11. What **FLAVOR(S)** of **OAT MILK** do you typically consume? (select all that apply)
   - Original/plain
   - Vanilla
   - Chocolate/other flavor
   - Other (specify) [fixed, require explanation]

12. What **BRAND(S)** of **OAT MILK** do you typically consume? (select all that apply)
   - Oatly
   - Chobani
   - Califia Farms
   - Planet Oat
   - Pacific Foods
   - Silk
   - Oat Malk
   - So Delicious
   - Store brand
   - Other (specify) [fixed, require explanation]
   - Don’t know/unsure

13. What **TYPE(S)** of **OAT MILK** do you typically consume? (select all that apply)
   - Original
   - Extra creamy
   - Barista blend
   - Low fat
   - Unsweetened
   - Zero sugar
   - Full fat
   - Other (specify) [fixed, require explanation]

14. What **TYPE** of **OAT MILK** do you consume **MOST OFTEN**?
   - Original
   - Extra creamy
   - Barista blend

- Low fat
- Unsweetened
- Zero sugar
- Full fat
- Other (specify) [fixed, require explanation]

15. [Open End] Why do you consume this type most often?

16. How do you consume this type of oat milk?
    - Drinking on its own
    - Adding to cereal/oatmeal
    - Adding to smoothie
    - Adding to coffee/tea
    - Using as an ingredient in a recipe
    - Other (specify)

17. What **PACKAGE SIZE** of **OAT MILK** do you purchase **MOST OFTEN**? Select one response
    - Quart (32oz)
    - Half gallon (64oz)
    - Gallon (128oz)
    - Other (specify)
    - Don't know

18. Where do you typically buy **OAT MILK** at the store? Select one response
    - In the refrigerated section
    - On the shelf (NOT refrigerated)

19. [Free Response] How would you describe your **IDEAL OAT MILK**? Please describe the flavor, texture, and/or any other important attributes.

20. Check the attributes below that you **LIKE**, when thinking about your **IDEAL OAT MILK**. (select all that apply)
    - Mild flavor
    - Tastes like real (dairy) milk
    - Nutty flavor
    - Toasted flavor
    - Oatmeal flavor
    - Cereal flavor
    - Grassy flavor
    - Oil flavor
    - Sweet taste
    - Salty taste

- Creamy texture
- Thin texture
- Smooth texture
- Gritty texture
- Chalky texture
- Oily mouthfeel
- White color
- Cream color
- Brown color
- Other (specify)

OAT MILK PRODUCT ATTRIBUTES (ACBC)

21. **CONJOINT BYO:** Below are features of **OAT MILKS** that you may find in the grocery store. For each of the features, select the level which is most appealing to you when shopping for an oat milk. Make a selection for each feature to BUILD YOUR IDEAL OAT MILK.
22. **CONJOINT SCREENER:** Here are a few **OAT MILKS** you may see at the store. For each one, indicate whether or not you would consider purchasing it. (anchors: would consider purchasing, would NOT consider purchasing)
23. **CONJOINT TOURNAMENT:** Among these three **OAT MILKS**, which is the BEST OPTION? (I've grayed out any features that are the same, so you can just focus on the differences.)

- Factors:
  - Brand
  - Flavor
  - Type
  - Label claim
  - Free-from claim
  - Price

- Factor: Price per half gallon (64oz)
  - Base: $4.35
  - See price additions in factors below (price varies +/- 30% summed price)
- Factor: Brand (Show as pictures of logos)
  - Oatly ($1.25)
  - Chobani ($1.15)
  - Califia Farms ($1.70)
  - Planet Oat ($1.00)
  - Silk ($0.30)
  - Store brand

- Factor: Flavor
    - Original/Plain
    - Vanilla
    - Chocolate/Other flavor
- Factor: Type
    - Extra creamy
    - Unsweetened
    - Zero sugar
    - ~~Barista blend ($3.50)~~
    - Low fat
    - Original
- Factor: Label claim
    - Plant based
    - Excellent source of Calcium, vitamins A and D
    - Good source of Calcium, vitamins A and D
    - Good/excellent source of protein ($2.60)
    - Organic ($2.60)
    - Non GMO
    - Vegan
- Factor: Free-from claim
    - Gluten free
    - Dairy free
    - Lactose free
    - Free from artificial flavors and colors
    - Nut free
    - Soy free
    - No added oil

24. [Chip Allocation] The following is a chip allocation exercise. You can allocate a total of 100 chips across each of the following attributes as you see fit, putting more chips in the attributes of OAT MILKS that are most important to you when **PURCHASING OAT MILKS**.

    Please note that the TOTAL must equal 100. If you do not place any value on a specific attribute, you may leave the box blank or type “0”.
    - Brand
    - Flavor
    - Nutritional profile
    - Other attributes (creamy, sugar free, low fat, etc.)
    - Label claim (good source, organic, gluten free, etc.)
    - Price

25. **MAXDIFF - Which of the following attributes is most important/least important when you purchase an oat milk**
    - Plant based

- Excellent source of calcium, vitamins A and D
- Good source of calcium, vitamins A and D
- Good source of protein
- Excellent source of protein
- Organic
- Non GMO
- Vegan
- Gluten free
- Dairy free
- Lactose free
- Cholesterol free
- Free from artificial flavors and colors
- Nut free
- Soy free
- Extra creamy
- Tastes like real (dairy) milk
- Has the same nutrition as dairy milk (protein, vitamins, mineral)
- Is a complete protein source
- Has a short ingredient list
- Has a clean label
- No added sugar
- No added oil
- Heart healthy
- Ideal calorie content
- Sustainable

26. **Kano -** If you have the following choices for **OAT MILKS**, how will you feel? Please read each choice carefully.
    - FUNCTIONAL:
        - An oat milk that has a MILD NUTTY FLAVOR
        - An oat milk that is a BRAND I AM FAMILIAR WITH
        - An oat milk that is my PREFERRED FLAVOR
        - An oat milk that has a CREAMY TEXTURE
        - An oat milk that is LIGHTLY SWEETENED
        - An oat milk that is UNSWEETENED
        - An oat milk that is SWEETENED WITH SUGAR
        - An oat milk that is SWEETENED WITH NON-NUTRITIVE SWEETENERS
        - An oat milk that has VITAMINS AND MINERALS
        - An oat milk that is FREE FROM ALLERGENS
        - An oat milk that IS PACKAGED IN A PLASTIC CONTAINER
        - An oat milk that is PACKAGED IN A PAPERBOARD CARTON
        - An oat milk that is CLEAN LABEL
        - An oat milk that is made with RECOGNIZABLE INGREDIENTS
        - An oat milk that has a SHORT INGREDIENT LIST

- An oat milk that TASTES GREAT
- An oat milk that has GREAT TEXTURE
- An oat milk that is SUSTAINABLY MADE
- An oat milk that TASTES LIKE DAIRY MILK
- An oat milk that is a GOOD SOURCE OF PROTEIN
- An oat milk that has NO ADDED SUGAR
- An oat milk that is NON GMO
- An oat milk that has NO ADDED OIL
- An oat milk that is in the REFRIGERATED SECTION of the grocery store
- An oat milk that has the SAME NUTRITION AS DAIRY MILK
- An oat milk that has the SAME PROTEIN AND CALCIUM CONTENT AS MILK

- DYSFUNCTIONAL:
    - An oat milk that does <u>NOT</u> have a MILD NUTTY FLAVOR
    - An oat milk that is a BRAND I AM <u>NOT</u> FAMILIAR WITH
    - An oat milk that is <u>NOT</u> my PREFERRED FLAVOR
    - An oat milk that does <u>NOT</u> have a CREAMY TEXTURE
    - An oat milk that is <u>NOT</u> LIGHTLY SWEETENED
    - An oat milk that is <u>NOT</u> UNSWEETENED
    - An oat milk that is <u>NOT</u> SWEETENED WITH SUGAR
    - An oat milk that is <u>NOT</u> SWEETENED WITH NON-NUTRITIVE SWEETENERS
    - An oat milk that does <u>NOT</u> have VITAMINS AND MINERALS
    - An oat milk that is <u>NOT</u> FREE FROM ALLERGENS
    - An oat milk that IS <u>NOT</u> PACKAGED IN A PLASTIC CONTAINER
    - An oat milk that is <u>NOT</u> PACKAGED IN A PAPERBOARD CARTON
    - An oat milk that is <u>NOT</u> CLEAN LABEL
    - An oat milk that is <u>NOT</u> made with RECOGNIZABLE INGREDIENTS
    - An oat milk that does <u>NOT</u> have SHORT INGREDIENT LIST
    - An oat milk that does <u>NOT</u> TASTE GREAT
    - An oat milk that does <u>NOT</u> have a GREAT TEXTURE
    - An oat milk that is <u>NOT</u> SUSTAINABLY MADE
    - An oat milk that does <u>NOT</u> TASTE LIKE DAIRY MILK
    - An oat milk that is <u>NOT</u> a GOOD SOURCE OF PROTEIN
    - An oat milk that HAS ADDED SUGAR
    - An oat milk that is <u>NOT</u> NON GMO
    - An oat milk that HAS ADDED OIL
    - An oat milk that is <u>NOT</u> in the REFRIGERATED SECTION of the grocery store
    - An oat milk that DOES NOT HAVE the SAME NUTRITION AS DAIRY MILK
    - An oat milk that DOES NOT HAVE the SAME PROTEIN AND CALCIUM CONTENT AS MILK
    -

- Scale:
  - I will dislike this oat milk
  - I can live with this oat milk
  - I do not care
  - I expect oat milk to have this
  - I will like an oat milk more if it has this

[Attention Check 2] This is an attention check question to make sure you are reading the questions. Select the option below that says "Never". Selecting the wrong response will disqualify you from the survey and you will not receive compensation.

- Never
- Occasionally
- Often
- Most of the time
- Always

OAT MILK KNOWLEDGE

27. [Chip Allocation] In this exercise, we want to know what you think the **COMPOSITION OF OAT MILK** is. Please type in the percentage that you believe corresponds to each category.

    Please note that the TOTAL must equal 100%. If you do not believe anything is present for a category, you may leave the box blank or type "0".

    - Water
    - Protein
    - Carbs
    - Fat

DEMOGRAPHICS

28. Please indicate your gender
    - Male
    - Female
    - Other
    - Prefer not to answer

29. Which of the following best describes your ethnicity?
    - White/Caucasian
    - Black/African American
    - Hispanic/Latino

- East Asian
- South Asian or Indian
- Middle Eastern
- Native American/Pacific Islander
- Other (specify) [require explanation]
- Prefer not to answer [fixed]

30. Which state do you live in?
    - [Drop down menu]

31. What is the highest level of education you have completed?
    - Some high school
    - High school diploma or equivalent
    - 2 years of college (i.e. Associate's degree)
    - 4 years of college (i.e. Bachelor's degree)
    - 5 or more years of college (i.e. Master's or PhD)
    - Technical/Trade/Vocational school

32. Which best describes your employment status?
    - Employed full-time
    - Employed part-time
    - Student
    - Homemaker
    - Retired
    - Other (specify) [fixed, require explanation]

33. What is your annual household income?
    - $24,999 or less
    - $25,000 - $49,999
    - $50,000 - $74,999
    - $75,000 - $99,999
    - $100,000 or more
    - Prefer not to answer

34. What is your household size?
    - 1 person
    - 2 person
    - 3 person
    - 4 person
    - 5+ person

35. Are there children (less than 18 years old) in your household?
    - Yes
    - No

36. How much of your household's grocery shopping are you responsible for?

- 100% - I do all of the grocery shopping
- 75% - I do most of the grocery shopping
- 50% - I share equally the grocery shopping
- 25% - I do some of the grocery shopping, but I am not the primary shopper for my household
- 0% - I do none of the grocery shopping

Appendix 2.2 Moderator Guide for Oat Milk Focus Groups

Pre-Focus Group Homework:

- Take a picture of all of the milks you currently have in your fridge/cabinet at home. This may include dairy milks, plant-based milks, etc.
- Once a day, update a diary that documents all milk usage (plant-based and dairy)
    - What type of milk did you use? (type, brand, etc.)
    - When did you consume it?
    - How did you consume it? Did you drink it on its own or add to something?
    - What did you like about it? Did it let you down in any way?

Introduction & Procedure

- Name, grad student
- Participation is voluntary
- If you don't want to answer question/share, that is okay
- There are no wrong answers
- Video footage collected for data analysis, short quotes → will not be shared outside of lab
- Only first names → Proceed to introductions

Oat Milk vs Other Milks – Occasions & Motivations

Today we are going to be talking about a lot of different types of milk and milk alternatives, but I want to start off thinking about oat milk. I know all of you consume oat milk, so I want you to take a minute to think about the first time you tried it. Feel free to write down notes of anything that you recall, and then I would like to go around and have everyone share their experience.

1. [1 @ time share] FIRST EXPERIENCE W/ OAT MILK

   What made you consider/want to try it in the first place? What was your impression?

   a. Probe – what was the occasion?

      i. How did you hear about it? Social media, friends/family, at a restaurant/café, etc.?

      ii. Was it during a meal/what time of day?

      iii. Did you try it plain? In something (coffee, smoothie, etc.)

   b. Probe - what type/brand?
      i. Flavored? Original? Barista? Extra creamy?
   c. Probe – how did the taste/texture compare to alternatives you had been using previously?

2. [1 @ time share] PRIOR TO OAT MILK

   What were you drinking before oat milk?

   a. Probe – what made you decide to try a different type of milk?
      i. Probe to understand if driven by taste, texture, health, sustainability, price, etc.
      ii. Probe to understand if different milks are being used for different purposes

3. [1 @ time share] CURRENT LINE UP

   Thinking about all of the different types of milk out there, what do you currently use?

   a. Probe to understand different types (dairy, oat, almond, soy, etc.), flavors (plain, chocolate, other flavors), varieties (full fat, reduced fat, sweeteners, etc.)
   b. Probe – how did you end up with these milks? Was there trial and error in determining your favorite(s)? Explain this for me.
   c. Probe – If multiple types of milk are being used, probe to understand why one milk is being used over another – what is the usefulness/limitations of oat milk compared to other milks?

4. [Write on paper] IDEAL OAT MILK

   Using the paper in front of you, write out a few things that come to mind when thinking about your ideal oat milk → [Share with the group]

[Transition to discuss homework activity – participants and moderator will have printed copy of individual diary responses]

5. [Show of hands] DIF TYPES OF MILK?

   Now, I want each of you to recall your diary where you documented when you used milk. How many of you used only one type of milk the entire week? How many of you used multiple different types of milk?

6. [1 @ time share] WHY MULTIPLE TYPES OF MILK

   Explain for us why you used multiple different types of milk?

a. Probe to understand the relationship between types of milk, occasions they are using them, and attributes (texture, flavor, etc.) of milk that seem to be important for different occasions.
b. Probe to understand how much (volume estimate) is being used for each occasion → is there a link between occasion and volume used?

7. [On board sorting activity] MILKS COMPARISON

We have talked a lot about different uses and why you are using different products for different occasions. I want to use the board to compare different milk types to each other.

c. Show all milks: oat milk, dairy milk, soy milk, almond milk, [insert any other applicable milk that came up during discussion]
d. Have group rank milks based on one attribute at a time: tasty, healthy, sustainable, ethical, affordable
    i. Probe – how do you define healthy? → probe to see if heart healthy comes up/is associated with oat milk

---

[Transition to discuss oat milk specifically]

Oat Milk Attributes

Let's shift our discussion to talk about oat milk specifically.

8. [1 @ time share] MOST IMPORTANT PURCHASING OAT MILK

When thinking about purchasing oat milk, what is most important to you?

a. Probe – type, flavor, price, nutrition, ingredients, health, etc.

9. [1 @ time share] INGREDIENTS YOU WANT/DON'T WANT

Are there ingredients in oat milk that you like to see? Are there ingredients that you do not want to see? Please explain.

10. [Group sorting exercise] I WANT THIS, I DON'T WANT THIS, UNSURE

b. Probe to understand motivation for wanting/not wanting certain ingredients – do they have personal experience, outside influence? Are any of these ingredients must have/dealbreakers?

11. [1 @ time share] WHAT MAKES OAT MILK…

Thinking about this sorting exercise, what makes oat milk… healthy, sustainable, taste good, etc.?

[BRING OUT COMMERCIAL PRODUCTS]

12. [Group Discussion] COMMERCIAL OAT MILKS

In front of us are various oat milks that are currently available in the market. Take a minute to get up and look through these products

- **c.** Do you usually read the label when buying food products? Oat milk?
- **d.** What do you look for on the label?
    - **i.** Probe to understand what they care about, what they don't – protein content, calories, ingredient list/length, calories, vitamins/minerals, health claims, sustainability claims, heart healthy

13. [Individual written activity] FINAL THOUGHTS

After everything we have talked about today, I would like each of you to write down your responses to a few questions:

- **e.** How does oat milk fit into your current line-up of milks?
- **f.** Can oat milk be used for all milk purposes?
- **g.** What makes oat milk better than other options?
- **h.** What does oat milk lack compared to other options?
- **i.** If you were trying to convince a friend to switch to oat milk, what would you tell them?
- **j.** → [Share responses with group]

14. [Group Discussion] FINAL DISCUSSION

Is there anything we have not discussed today that you want to bring up before we conclude this focus group?

Appendix 3.1 Ballot for Coffee Creamer CLT

****Today you will be evaluating 4 Vanilla Coffee Creamers - Plant & Dairy****

Please LOOK at SAMPLE ### and answer the following questions regarding appearance. DO NOT TASTE the sample until instructed to do so.

1. Which statement best describes your impression of the **APPEARANCE** of this creamer?
    - 9-pt Hedonic
2. How do you feel about the **COLOR** of this creamer?
    - Much Too Dark
    - Too Dark
    - Just About Right
    - Too Light
    - Much Too Light
3. Which statement best describes your impression of the **AROMA** of this creamer?
    - 9-pt Hedonic
4. How do you feel about the **AROMA** of this creamer?
    - Much Too Strong
    - Too Strong
    - Just About Right
    - Too Weak
    - Much Too Weak

Please ADD SAMPLE ### CREAMER TO THE COFFEE and answer the following questions.
DO NOT TASTE the sample until instructed to do so.

5. How well did the **CREAMER MIX** into the coffee?
    - 9-pt Intensity scale (Extremely Poorly → Extremely Well)
6. Which statement best describes your impression of the **APPEARANCE** of this creamer in coffee?
    - 9-pt Hedonic
7. Which **COLOR** best represents the color of the coffee with creamer?
    - Coffee color intensity scale
8. How do you feel about the **COLOR** of this creamer in coffee?
    - Much Too Dark
    - Too Dark
    - Just About Right
    - Too Light
    - Much Too Light

9. Which statement best describes your impression of the **AROMA** of this creamer in coffee?
    - 9-pt Hedonic
10. How do you feel about the **AROMA** of this creamer in coffee?
    - Much Too Strong
    - Too Strong
    - Just About Right
    - Too Weak
    - Much Too Weak

Now please TASTE SAMPLE ### and answer the following questions. Remember, you are evaluating the CREAMER and NOT THE COFFEE.

11. Which statement best describes your impression of the **OVERALL LIKING** of this creamer in coffee?
    - 9-pt Hedonic
12. How likely would you be to **PURCHASE** this creamer if it was available at a market you normally shop and at a reasonable price?
    - Definitely Would NOT Buy
    - Probably Would NOT Buy
    - Might or Might Not Buy
    - Probably Would Buy
    - Definitely Would Buy
13. How do you feel about the **CREAMER TO COFFEE FLAVOR BALANCE** of this coffee with creamer?
    - Much Too Much Creamer/Not Nearly Enough Coffee
    - Too Much Creamer/Not Enough Coffee
    - Just About Right
    - Not Enough Creamer/Too Much Coffee
    - Not Nearly Enough Creamer/Much Too Much Coffee
14. Which statement best describes your impression of the **SWEETNESS** of this creamer in coffee?
    - 9-pt Hedonic
15. How do you feel about the **SWEETNESS** of this creamer in coffee?
    - Much Too Sweet
    - Too Sweet
    - Just About Right
    - Not Sweet Enough
    - Not Nearly Sweet Enough
16. How would you rate the **SWEETNESS INTENSITY** of this creamer in coffee?
    - 9-pt Intensity Scale (Not At All Sweet → Extremely Sweet)

17. Which statement best describes your impression of the **BITTERNESS** of this creamer in coffee?
    - 9-pt Hedonic
18. How would you rate the **BITTERNESS INTENSITY** of this creamer in coffee?
    - 9-pt Intensity Scale (Not At All Sweet → Extremely Sweet)
19. Which statement best describes your impression of the **OVERALL FLAVOR** of this creamer in coffee?
    - 9-pt Hedonic
20. How do you feel about the **OVERALL FLAVOR** of this creamer in coffee?
    - Much Too Strong
    - Too Strong
    - Just About Right
    - Too Weak
    - Much Too Weak
21. Which statement best describes your impression of the **VANILLA FLAVOR** of this creamer in coffee?
    - 9-pt Hedonic
22. How do you feel about the **VANILLA FLAVOR** of this creamer in coffee?
    - Much Too Strong
    - Too Strong
    - Just About Right
    - Too Weak
    - Much Too Weak
23. Which statement best describes your impression of the **MOUTHFEEL** of this creamer in coffee?
    - 9-pt Hedonic
24. How do you feel about the **MOUTHFEEL** of this creamer in coffee?
    - Much Too Thick
    - Too Thick
    - Just About Right
    - Too Thin
    - Much Too Thin
25. Which statement best describes your impression of the **CREAMINESS** of this creamer in coffee?
    - 9-pt Hedonic
26. How do you feel about the **CREAMINESS** of this creamer in coffee?
    - Much Too Creamy
    - Too Creamy
    - Just About Right
    - Not Creamy Enough

- Not Nearly Creamy Enough

27. How would you rate the **CREAMINESS INTENSITY** of this creamer in coffee?
    - 9-pt Intensity Scale (Not At All Creamy → Extremely Creamy)
28. Do you detect an AFTERTASTE in this creamer in coffee?
    - Yes
    - No [skip AFTERTASTE Liking]
29. [Connection] Which statement best describes your impression of the **AFTERTASTE** of this creamer in coffee?
    - 9-pt Hedonic
30. How well does this creamer meet your **EXPECTATIONS** for a VANILLA COFFEE CREAMER?
    - Much Worse Than Expected
    - Slightly Worse Than Expected
    - About the Same As Expected
    - Slightly Better Than Expected
    - Much Better Than Expected
31. How well does this coffee with creamer **COMPARE TO THE COFFEE WITH CREAMER YOU USUALLY CONSUME**?
    - Much Worse Than My Usual Coffee With Creamer
    - Slightly Worse Than My Usual Coffee With Creamer
    - About The Same As My Usual Coffee With Creamer
    - Slightly Better Than My Usual Coffee With Creamer
    - Much Better Than My Usual Coffee With Creamer
32. [FREE RESPONSE TO COMPARE W/USUAL Q, bottom 2 box] Describe why you rated this coffee with creamer worse than your usual coffee with creamer.
33. [After all samples tasted] Please rank the samples you just tasted in order of your preference (1 = most liked, 4 = least liked)

---

34. Do you always add the same amount of creamer to your coffee?
    - Yes
    - No
35. How do you add creamer to your coffee?
    - Based on color
    - Based on taste
    - Measure
    - Other (specify) [fixed, require explanation]
36. Thinking about the amount of creamer you typically use, how would you describe the **AMOUNT OF CREAMER PROVIDED FOR EACH SAMPLE**?
    - Much Too Much Creamer
    - Too Much Creamer

- Just About Right
- Not Enough Creamer
- Not Nearly Enough Creamer

37. What **TYPE(S)** of **CREAMER** do you use in your coffee? (select all that apply)
    - Dairy-based
    - Non-dairy based
    - Almond milk based
    - Oat milk based
    - Soy milk based
    - Coconut milk based
    - Sugar free
    - Fat free
    - Liquid
    - Powder
38. What **BRAND(S)** of **CREAMER** do you typically purchase?
    - Starbucks
    - International Delight
    - Coffee Mate
    - Dunkin
    - Coffee Mate Natural Bliss
    - Chobani
    - Silk
    - Califia
    - Nut Pods
    - Store brands
    - Other (specify) [fixed, require explanation]

Appendix 5.1 Ballot for Protein Beverage Survey

DEMOGRAPHICS

1. How old are you (years)
    - [Numerical Response] [DISQUALIFY if <18]
2. Please indicate your gender
    - Male
    - Female
    - Other
    - Prefer not to answer
3. Which of the following best describes your ethnicity?
    - White/Caucasian
    - Black/African American
    - Hispanic/Latino
    - East Asian
    - South Asian or Indian
    - Middle Eastern
    - Native American/Pacific Islander
    - Other (please specify) [require explanation]
    - Prefer not to answer [exclusive]
4. Which state do you live in?
    - [Drop down menu, all states]
        - Alabama
        - Alaska
        - Arizona
        - Arkansas
        - California
        - Colorado
        - Connecticut
        - Delaware
        - Florida
        - Georgia
        - Hawaii
        - Idaho
        - Illinois
        - Indiana
        - Iowa
        - Kansas
        - Kentucky
        - Louisiana
        - Maine
        - Maryland
        - Massachusetts
        - Michigan
        - Minnesota
        - Mississippi
        - Missouri
        - Montana
        - Nebraska
        - Nevada
        - New Hampshire
        - New Jersey
        - New Mexico
        - New York
        - North Carolina
        - North Dakota
        - Ohio
        - Oklahoma
        - Oregon
        - Pennsylvania
        - Rhode Island
        - South Carolina
        - South Dakota
        - Tennessee
        - Texas
        - Utah
        - Vermont
        - Virginia
        - Washington
        - West Virginia
        - Wisconsin
        - Wyoming
5. What is the highest level of education you have completed?

- Some high school
- High school diploma or equivalent
- 2 years of college (i.e. Associate's degree)
- 4 years of college (i.e. Bachelor's degree)
- 5 or more years of college (i.e. Master's or PhD)
- Technical/Trade/Vocational school

6. Which best describes your employment status?
    - Employed full-time
    - Employed part-time
    - Student
    - Homemaker
    - Retired
    - Other (please specify) [require explanation]
7. What is your annual household income?
    - $24,999 or less
    - $25,000 - $49,999
    - $50,000 - $74,999
    - $75,000 - $99,999
    - $100,000 or more
    - Prefer not to answer
8. What is your household size?
    - 1 person
    - 2 person
    - 3 person
    - 4 person
    - 5+ person
9. Are there children (less than 18 years old) in your household?
    - Yes
    - No
10. How much of your household's grocery shopping are you responsible for?
    - 100% - I do all of the grocery shopping
    - 75% - I do most of the grocery shopping
    - 50% - I share equally the grocery shopping
    - 25% - I do some of the grocery shopping, but I am not the primary shopper for my household
    - 0% - I do none of the grocery shopping [DISQUALIFY]
11. Please indicate which, if any, of the following products you have purchased or consumed in the **past 6 months** (please check all that apply)
    - Protein-fortified products (i.e. protein powder, protein bars, protein beverages) [MUST SELECT]

- Plant-based products (meat or dairy alternatives)
- Meat (chicken, beef, etc.)
- Alcohol
- Juice
- Coffee
- Tea
- None of the above [fixed, exclusive]

12. How often do you purchase and consume **PROTEIN BEVERAGES**?
    - Daily
    - Weekly
    - Monthly
    - Yearly
    - Never [DISQUALIFY]
13. What type(s) of **PROTEIN BEVERAGES** do you typically purchase and consume? (please check all that apply)
    - **Ready to mix protein beverages** - powders that you rehydrate to make your own beverages/shakes/smoothies. Protein powder can be flavored or unflavored (e.g. ON Gold Standard Whey Protein Powder, Muscle Milk Protein Powder, Orgain Organic Protein Powder, Premier Protein Powder, OWYN Protein Powder, Vega Sport Protein Powder, etc.)
    - **Refrigerated ready to drink protein beverages** (e.g. Fairlife Chocolate Milk, Naked Protein Smoothie, Bolthouse Farms Protein Plus Protein Shake, etc.)
    - **Shelf-stable ready to drink protein beverages** (e.g. Atkins Shake, GNC Lean Shake, Muscle Milk Protein Shake, Fairlife Core Power, Ensure Nutritional Shake, OWYN Protein Shake, Evolve Plant-based Protein Shake, Iconic Grass Fed Protein Drink, etc.)
    - Other (please specify)
    - None of these [DISQUALIFY]

[TEXT SCREEN] For the remainder of the survey, when referring to "**PROTEIN BEVERAGES,**" we are referring to all types of protein beverages, including:

- Ready to mix protein beverages - protein powder (flavored or unflavored) to make your own beverages/shakes/smoothies
- Refrigerated ready to drink protein beverages
- Shelf-stable ready to drink protein beverages

Please click "Next" to continue

---

14. Thinking about the nutrition and ingredient labels on **PROTEIN BEVERAGES**, which of the following statements best describes you?

- I ALWAYS read nutrition and ingredient labels
- I OFTEN read nutrition and ingredient labels
- I SOMETIMES read nutrition and ingredient labels
- I RARELY read nutrition and ingredient labels
- I NEVER read nutrition and ingredient labels [DISQUALIFY]

15. Thinking about shopping for a **PROTEIN BEVERAGE**, describe what you look for when reading the label.
    - [Free response]

Attention Check
This is an attention check question to make sure you are not just clicking through the survey without answering honestly. If you answer attention check questions incorrectly, your data cannot be used and you will be ineligible to receive a gift card. Please select the option below that says **Dislike Very Much**.

- Dislike Very Much
- Dislike Slightly
- Neither Like Nor Dislike
- Like Slightly
- Like Very Much

ANCHORED MAXDIFF EXERCISES

[TEXT SCREEN] The following exercises will ask you about your impression of ingredients found in PROTEIN BEVERAGES.
These exercises will present you with 5 ingredients at a time and ask you to select which ingredient is the LEAST APPEALING and MOST APPEALING if you were to encounter it on the label of a protein beverage.
The ingredients you will see are ingredients that may be present in commercial protein beverages that you purchase or may consider purchasing.
Your feedback will help us with our research on protein beverages.
Thank you in advance for your input and contribution to this research!

16. PROTEIN SOURCES: Thinking about the ingredients in **PROTEIN BEVERAGES** you may consider purchasing, which **protein sources** are **LEAST APPEALING** and **MOST APPEALING** to you?
    - Bovine collagen peptides
    - Brown rice protein
    - Calcium caseinate
    - Casein protein
    - Faba bean protein
    - Hydrolyzed soy protein
    - Hydrolyzed whey protein
    - Milk protein
    - Pea protein
    - Pumpkin seed protein
    - Sodium caseinate
    - Soy protein
    - Whey protein
    - Dairy protein

- Plant protein
    - We are interested in knowing if any of these ingredients are deal breakers (you would NOT buy a product if you saw the ingredient on the label of a protein beverage).
      Please indicate your willingness to buy a **PROTEIN BEVERAGE** with each of the **protein sources** listed below on the ingredient statement:
        - If I saw this ingredient on a label, I would **NOT BUY** it
        - If I saw this ingredient on a label, I would **STILL CONSIDER BUYING** it
    - [Grid Q] For each ingredient, please select which statements you believe to be true (please check all that apply)
        - Clean label
        - Natural
        - Healthy
        - Sustainable
        - Familiar
        - None of these [exclusive]

17. SWEETENERS: Thinking about the ingredients in **PROTEIN BEVERAGES** you may consider purchasing, which **sweeteners** are **LEAST APPEALING** and **MOST APPEALING** to you?
    - Acesulfame potassium (Ace K)
    - Allulose
    - Agave nectar
    - Brown rice syrup
    - Cane sugar/cane syrup
    - Chicory fiber syrup
    - Coconut nectar
    - Fructose
    - Monk fruit
    - Naturally sweetened
    - Stevia
    - Sucralose
    - Sugar
        - We are interested in knowing if any of these ingredients are deal breakers (you would NOT buy a product if you saw the ingredient on the label of a protein beverage).
          Please indicate your willingness to buy a **PROTEIN BEVERAGE** with each of the **sweeteners** listed below on the ingredient statement:
            - If I saw this ingredient on a label, I would **NOT BUY** it
            - If I saw this ingredient on a label, I would **STILL CONSIDER BUYING** it
        - [Grid Q] For each ingredient, please select which statements you believe to be true (please check all that apply)
            - Clean label
            - Natural
            - Healthy
            - Sustainable

        - Familiar
        - None of these [exclusive]

Attention Check
This is an attention check question to make sure you are not just clicking through the survey without answering honestly. If you answer attention check questions incorrectly, your data cannot be used and you will be ineligible to receive a gift card. Please select the option below that says **Like Very Much**.

- Dislike Very Much
- Dislike Slightly
- Neither Like Nor Dislike
- Like Slightly
- Like Very Much

18. GUMS/THICKENERS: Thinking about the ingredients in **PROTEIN BEVERAGES** you may consider purchasing, which **thickeners** are **LEAST APPEALING** and **MOST APPEALING** to you?
    - Acacia gum
    - Agar
    - Carrageenan
    - Cellulose gel
    - Cellulose gum
    - Chicory fiber
    - Enriched flour
    - Gellan gum
    - Inulin
    - Locust bean gum
    - Pea starch
    - Chicory root fiber
    - Soluble corn fiber
    - Rice flour
    - Rice starch
    - Corn fiber
    - Soy fiber
    - Soy flour
    - Tapioca fiber
    - Tapioca starch
    - Vegetable fiber
        - We are interested in knowing if any of these ingredients are deal breakers (you would NOT buy a product if you saw the ingredient on the label of a protein beverage).
          Please indicate your willingness to buy a **PROTEIN BEVERAGE** with each of the **thickeners** listed below on the ingredient statement:
            - If I saw this ingredient on a label, I would **NOT BUY** it
            - If I saw this ingredient on a label, I would **STILL CONSIDER BUYING** it
        - [Grid Q] For each ingredient, please select which statements you believe to be true (please check all that apply)
            - Clean label
            - Natural
            - Healthy

- Sustainable
- Familiar
- None of these [exclusive]

19. STABILIZERS/EMULSIFIERS: Thinking about the ingredients in **PROTEIN BEVERAGES** you may consider purchasing, which **stabilizers** are **LEAST APPEALING** and **MOST APPEALING** to you?
    - Corn maltodextrin
    - Datem
    - Dipotassium phosphate
    - Sunflower lecithin
    - Magnesium phosphate
    - Maltodextrin
    - Mono- and diglycerides
    - Monoglycerides
    - Polydextrose
    - Potassium citrate
    - Potassium phosphate
    - Sodium citrate
    - Sodium hexametaphosphate
    - Sodium phosphate
    - Soy lecithin
    - Tricalcium phosphate
    - Tripotassium citrate
    - Tripotassium phosphate
    - Trisodium phosphate
        - We are interested in knowing if any of these ingredients are deal breakers (you would NOT buy a product if you saw the ingredient on the label of a protein beverage).
          Please indicate your willingness to buy a **PROTEIN BEVERAGE** with each of the **stabilizers** listed below on the ingredient statement:
            - If I saw this ingredient on a label, I would **NOT BUY** it
            - If I saw this ingredient on a label, I would **STILL CONSIDER BUYING** it
        - [Grid Q] For each ingredient, please select which statements you believe to be true (please check all that apply)
            - Clean label
            - Natural
            - Healthy
            - Sustainable
            - Familiar
            - None of these [exclusive]

[End of Sawtooth portion of the survey → link to Compusense]

PROJECTIVE MAPPING EXERCISE

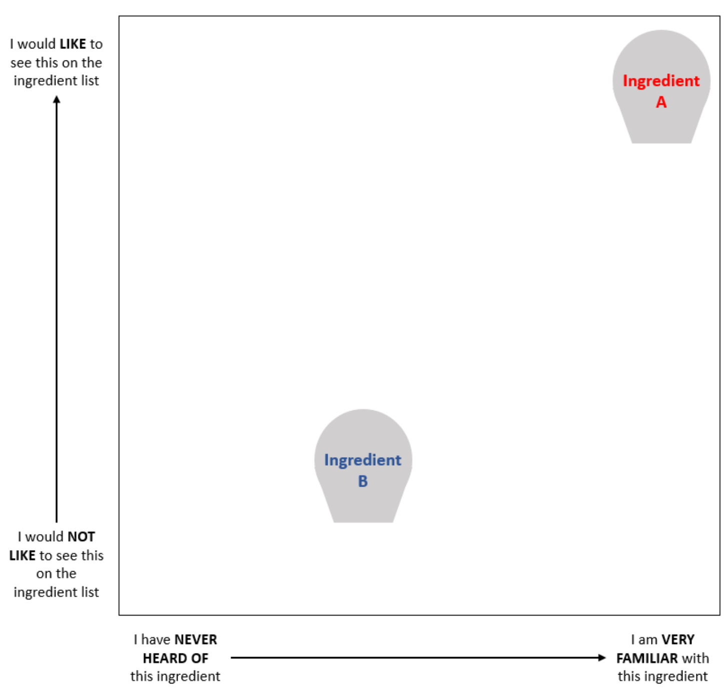
For this portion of the survey, you will be doing a **MAPPING EXERCISE** to show how you view different ingredients that may be present in **PROTEIN BEVERAGES**.

You will be presented with a grid that has anchors as shown in the example below, with **FAMILIARITY on the bottom** and **LIKING on the left axis**. During this exercise, you will place ingredients on the page one at a time until all ingredients are on the page. Feel free to move ingredients around as you see fit.

Please make sure to take into consideration the anchors on both axes. For example, consider the following ingredients and how they are placed:

- Ingredient A - you are **very familiar** with it and **like to see it on the ingredient list** for protein beverages.
- Ingredient B - you are **somewhat familiar** with it and you would **prefer not to see it on an ingredient list** for protein beverages.

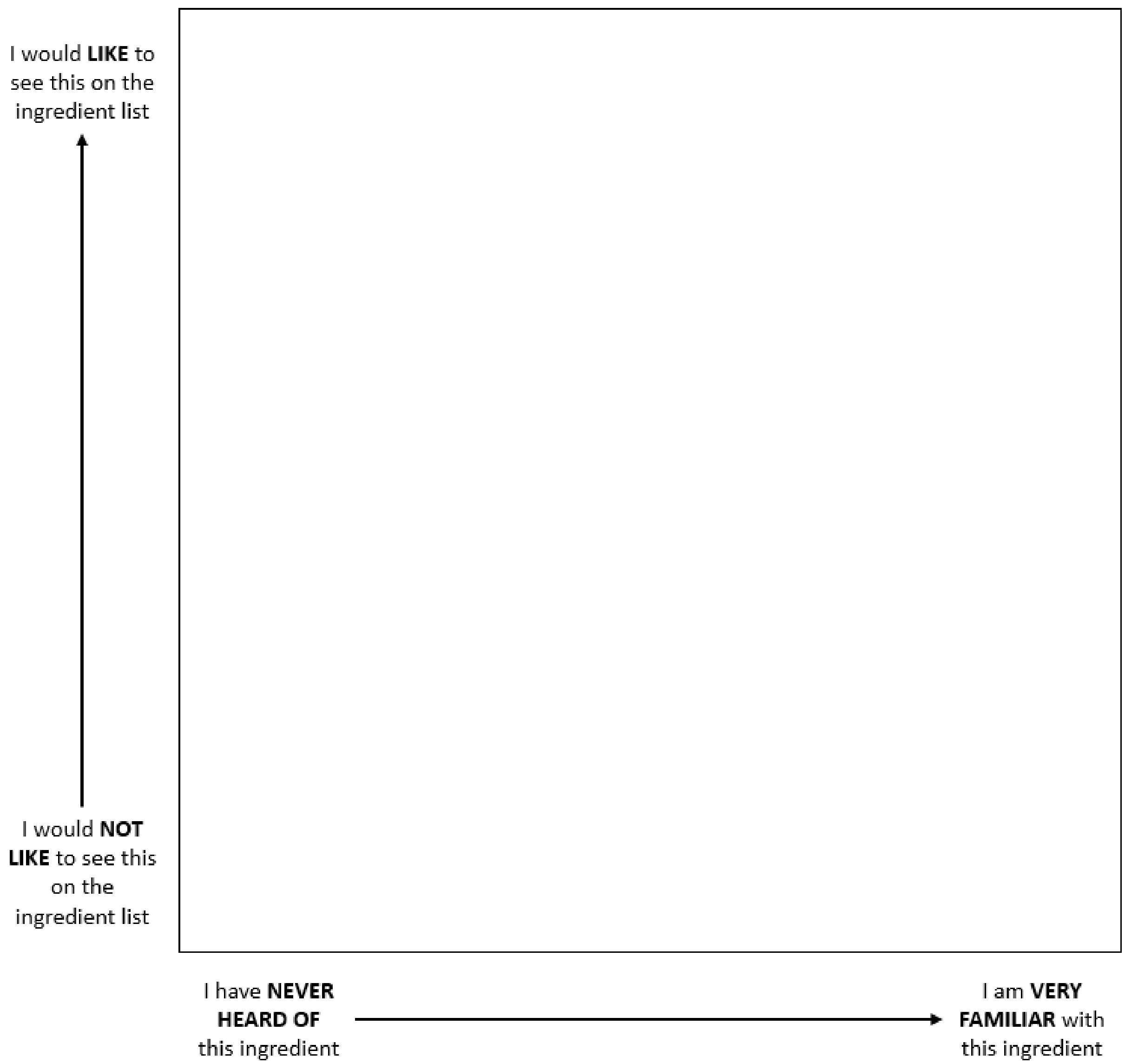
Please take as much time as you need to read through this information and use the examples to understand how to place items.



Once you are ready to start the exercise on your own, please click “Next” at the bottom of your screen.

20. Thinking about each of these **INGREDIENTS** in a **PROTEIN BEVERAGE**, please **DRAG** each ingredient onto the graph area as you see fit. You may rearrange the ingredients as necessary. Once you are happy with the placement of all ingredients, please click “Next” at the bottom of the page.

Use the “+” button on each attribute to view the ingredient name.



[Present 17 of 68]

1. Bovine collagen peptides
2. Brown rice protein
3. Calcium caseinate
4. Casein protein
5. Faba bean protein
6. Hydrolyzed soy protein
7. Hydrolyzed whey protein
8. Milk protein
9. Pea protein
10. Pumpkin seed protein
11. Sodium caseinate
12. Soy protein
13. Whey protein
14. Dairy protein
15. Plant protein
16. Acesulfame potassium (Ace K)
17. Allulose
18. Agave nectar
19. Brown rice syrup
20. Cane sugar/cane syrup
21. Chicory fiber syrup
22. Coconut nectar
23. Fructose
24. Monk fruit
25. Naturally sweetened
26. Stevia
27. Sucralose
28. Sugar
29. Acacia gum
30. Agar
31. Carrageenan
32. Cellulose gel
33. Cellulose gum
34. Chicory fiber
35. Enriched flour
36. Gellan gum
37. Inulin
38. Locust bean gum
39. Pea starch
40. Chicory root fiber
41. Soluble corn fiber
42. Rice flour
43. Rice starch
44. Corn fiber
45. Soy fiber
46. Soy flour
47. Tapioca fiber
48. Tapioca starch
49. Vegetable fiber
50. Corn maltodextrin
51. Datem
52. Dipotassium phosphate
53. Sunflower lecithin
54. Magnesium phosphate
55. Maltodextrin
56. Mono- and diglycerides
57. Monoglycerides
58. Polydextrose
59. Potassium citrate
60. Potassium phosphate
61. Sodium citrate
62. Sodium hexametaphosphate
63. Sodium phosphate
64. Soy lecithin
65. Tricalcium phosphate
66. Tripotassium citrate
67. Tripotassium phosphate
68. Trisodium phosphate

PSYCHOGRAPHICS

21. [Grid Q - 5-pt scale] Please indicate how important each of the attributes below are when reading a package for a **PROTEIN BEVERAGE**?
    - Plain language and packaging
    - Recognizable ingredients
    - Fewer ingredients [Connection]
    - No antibiotics or hormones
    - Organic
    - GMO-free
    - Not genetically engineered
    - Not genetically modified
    - Non-scientific ingredients
    - Not genetically modified through the use of modern biotechnology
    - Not bioengineered
    - Local
    - Recognizable brand
        - Not At All Important
        - Not Important
        - Somewhat Important
        - Important
        - Very Important
22. [Grid Q - 5-pt scale] Please indicate how important each of the attributes below are when reading a package for a **PROTEIN BEVERAGE**?
    - No artificial flavorings
    - No food additives
    - All natural ingredients
    - No added sugar
    - Sugar-free or reduced sugar
    - Reduced allergens or allergen-free
    - Complete protein
    - Plant protein
    - Dairy protein
    - Lactose free
    - Amount of protein
    - Type of sweetener
    - Type of protein
        - Not At All Important
        - Not Important
        - Somewhat Important
        - Important

        - Very Important
23. [Connection] You indicated that you prefer a label that does not have too many ingredients. Please describe how you decide how many ingredients you consider to be "too many" for a **PROTEIN BEVERAGE**? Does this number change for other types of food products?
    - [Free response]
24. When reading the ingredient statement for a **PROTEIN BEVERAGE** you are considering purchasing, how do you respond when you come across an ingredient that you are unfamiliar with? Do you choose not to purchase it? Compare it with other similar products? Look up what the ingredient is?
    - [Free response]
25. [Chip Allocation] The following is a chip allocation exercise. You can allocate a total of 100 chips across each of the following attributes as you see fit, putting more chips on the attributes that are more important to you when **PURCHASING A PROTEIN BEVERAGE.**

    Please note that the TOTAL must equal 100. If you do not place any value on a specific attribute, you may leave the box blank or type "0".

    - Amount of protein
    - Type of protein
    - Type of sweetener
    - Presence of additives
    - Length of ingredient list
26. [Kano] If you have the following choices for **PROTEIN BEVERAGES**, how will you feel? Please read each choice carefully.
    - Functional:
        i. A protein beverage that is made with PLANT PROTEIN
        ii. A protein beverage that is made with DAIRY PROTEIN
        iii. A protein beverage that is made with a BLEND OF PLANT AND DAIRY PROTEIN
        iv. A protein beverage that is made with ORGANIC INGREDIENTS
        v. A protein beverage that is made with SUSTAINABLE INGREDIENTS/is SUSTAINABLY PRODUCED
        vi. A protein beverage that is made with ingredients that are RECOGNIZABLE
        vii. A protein beverage that has a SHORT INGREDIENTS LIST
        viii. A protein beverage that TASTES GREAT
        ix. A protein beverage that is LACTOSE FREE
        x. A protein beverage from a RECOGNIZABLE BRAND
    - Dysfunctional:

i. A protein beverage that is NOT made with PLANT PROTEIN
ii. A protein beverage that is NOT made with DAIRY PROTEIN
iii. A protein beverage that is NOT made with a BLEND OF PLANT AND DAIRY PROTEIN
iv. A protein beverage that is NOT made with ORGANIC INGREDIENTS
v. A protein beverage that is NOT made with SUSTAINABLE INGREDIENTS/is NOT SUSTAINABLY PRODUCED
vi. A protein beverage that is NOT made with ingredients that are RECOGNIZABLE
vii. A protein beverage that does NOT have a SHORT INGREDIENTS LIST
viii. A protein beverage that does NOT TASTE GREAT
ix. A protein beverage that is NOT LACTOSE FREE
x. A protein beverage NOT from a RECOGNIZABLE BRAND

- ○ I will dislike it
- ○ I can live with it
- ○ I do not care
- ○ I must have it
- ○ I will like it

27. [Sliding Scale] Using the anchored scales below, please drag the markers where you see fit.
    How important is SUSTAINABILITY OF FOOD to you?
    - Not At All Important (1)
    - Extremely Important (100)
28. [Sliding Scale] How KNOWLEDGEABLE would you say you are about FOOD SUSTAINABILITY?
    - Not At All Knowledgeable (1)
    - Extremely Knowledgeable (100)
29. [Grid Q - 5-pt agree scale] How much do you agree or disagree with each of the following statements?
    - I am trying to eat less meat. (Mintel)
    - I am trying to eat less dairy. (Mintel)
    - I am trying to eat less animal products. (Protein Loyalty Survey and survey specific)
    - I am trying to consume more plant-based foods. (Mintel)
    - I am interested in trying animal free dairy and meat products (Mintel)
        - ○ Strongly Disagree
        - ○ Disagree
        - ○ Neither Agree Nor Disagree
        - ○ Agree
        - ○ Strongly Agree
30. Where do you shop for **PROTEIN BEVERAGES**? (please check all that apply)

- Premium Grocery Stores (e.g. Fresh Market, Whole Foods, Southern Season, etc.)
- Bulk Suppliers (e.g. Sam's Club, Costco, BJ's, etc.)
- Discount Grocery Stores (e.g. Aldi, Lidl, etc.)
- Grocery Sections within Department Stores (e.g. Walmart, Target, etc.)
- Standard Grocery Stores (e.g. Harris Teeter, Kroger, Food Lion, etc.)
- Farmer's Markets (including community sponsored agriculture and produce delivery services)
- Online (e.g. Amazon, manufacturer websites, etc.)
- Other (please specify) [fixed, require explanation]

31. When do you typically consume **PROTEIN BEVERAGES**? (please check all that apply)
    - As a replacement for a meal
    - As a snack
    - As a treat/dessert
    - Before a workout/exercise
    - After a workout/exercise
    - Other (please specify) [fixed, require explanation]
32. Which, if any, of the following diets do you follow? (please check all that apply)
    - Vegan
    - Vegetarian
    - Pescatarian
    - Whole30
    - Flexitarian
    - Gluten-free
    - Dairy-free
    - Low sugar
    - Low carb
    - Low fat
    - Keto
    - Paleo
    - Diabetic
    - Kosher
    - Halal
    - Other (please specify) [fixed, require explanation]
    - None, I do not follow any specific diet [fixed, exclusive]

33. What are the reasons you choose to follow your current diet? (please check all that apply)

- Good nutrition
- Manage weight
- Reduce risk of chronic disease
- Low cost
- Delay the effects of aging
- Maintain good immune system
- Provide energy
- Increase focus
- Sustainable
- Animal-welfare
- Environmentally friendly
- I don’t know [fixed, exclusive]

34. Are you lactose intolerant or dairy intolerant?

- No - I can eat all the dairy products I want with no intestinal discomfort
- I am mildly lactose intolerant - I can eat some dairy products
- I am moderately lactose intolerant - I have to be careful of how much dairy I consume
- I am very lactose intolerant - I cannot consume dairy products

Appendix 5.2 Moderator Guide for Protein Beverage Focus Groups

HOMEWORK

Before coming to campus for the in-person focus group, there is a short homework assignment that you need to complete. This will involve submitting two pictures:

1. Your FAVORITE protein beverage/the one you buy MOST OFTEN (this picture can be taken from home or from the store)
2. One picture showing all the protein beverages, protein supplements (bars, snacks, etc.), and functional beverages you currently have at home

When we are referring to protein beverages, we are referring to all of the following forms:

- **Ready to mix protein beverages** - powders that you rehydrate to make your own beverages/shakes/smoothies. Protein powder can be flavored or unflavored (e.g. ON Gold Standard Whey Protein Powder, Muscle Milk Protein Powder, Orgain Organic Protein Powder, Premier Protein Powder, OWYN Protein Powder, Vega Sport Protein Powder, etc.)
- **Refrigerated ready to drink protein beverages** (e.g. Fairlife Chocolate Milk, Naked Protein Smoothie, Bolthouse Farms Protein Plus Protein Shake, etc.)
- **Shelf-stable ready to drink protein beverages** (e.g. Atkins Shake, GNC Lean Shake, Muscle Milk Protein Shake, Fairlife Core Power, Ensure Nutritional Shake, OWYN Protein Shake, Evolve Plant-based Protein Shake, Iconic Grass Fed Protein Drink, etc.)

MODERATOR GUIDE
Protein Beverage Profile

- [Writing exercise] To get us started, I want to get everyone thinking about how they define a protein beverage. Using the paper in front of you, write down a bulleted list of characteristics of protein beverages. What makes something a protein beverage? What is or isn't in it?
    - [Group discussion] Share attributes that were written down.
- Now I want to ask each of you to tell us about the protein beverage you consume MOST OFTEN (identified during the homework)
    - [One at a time discussion] Describe your protein beverage for me. What do you like about it?
    - [Probe if not brought up]
        - Is it flavored (what kind)?
        - What form? (ready to mix, ready to drink, refrigerated, clear acidic?). Why do you like this form?

        - If using multiple forms, why multiple forms? What are the benefits of one form over another?
    - Where do you buy this protein beverage?
    - [Probe if not brought up]
        - What store? Online?
        - Where in the store - nutrition/supplements aisle, refrigerated case, other?
    - How long have you been using this beverage? When did you start using it/how did you find out about it?
    - When you saw the homework asking you to take a picture of your favorite/most often protein beverage, did this one immediately come to mind, or did you have to think about it?
    - [Probe about what considerations you made]
        - Is there anything that you do NOT like about this product? What would you change about it?
        - What other products have you tried? Likes/dislikes?
        - Where do you get information (packages, online, doctor, etc.)

Protein Beverage Usage Occasion

- [Writing exercise] Next, I want to talk about *when* you are drinking these beverages. Please use the paper in front of you to write down the top three occasions when you drink protein beverages. This could be a time of day, an activity, or a certain setting.
    - [One at a time discussion] Tell me about your top occasions
    - Do you use your MOST OFTEN beverage for all of these occasions? If no, → Tell me more about this
        - [Probe] Do you use multiple different forms? When do you use each form? What makes these forms good for your occasion? When RTDs?
    - Do you purchase in bulk? On the go, individual, at the store?
    - Do you consume anything else with it? (bars, supplements, etc.)
    - If you don't have a protein beverage during these occasions, what do you do?
        - [Probe] Do you always use this product for this occasion? If not, what is the substitute? → another similar product, entirely different, go without?

Ingredients and Labeling on RTD Protein Beverages

Preface - ONLY RTDs
Up until now, we have been talking about all types of protein beverages, but for the remainder of this discussion we will only be talking about ready to drink protein beverages, NOT ready to mix protein powders.

Top of mind shopping

- When you are looking to buy a protein beverage (online, at the store, etc.) what do you usually look for when you are shopping?
- Do you read the label? Why/why not? When you do read the label, what do you look for?
    - [Probe if not brought up] I have heard other consumers say they look for…
        - "The amount of protein"
        - "The type of protein" (plant vs dairy?)
        - "The type of sweetener"
        - "Whether or not there are additives"
        - "The macros - calories, protein, fat, sugar"
        - "The number of ingredients"
        - "Whether or not the ingredients are recognizable"
    - How do you feel about these statements?

Commercial product discussion/Clean Label

- [Label reading exercise with COMMERCIAL PRODUCTS] Now I am going to bring out some protein beverages for everyone to look at. Let's pretend that you were at the store and these were the options for protein beverages. I want each of you to take a moment to look at these beverages and then we will discuss what we think.
    - Just looking at the packaging for these beverages, which of these do you like? What do you like about it?
        - [Probe to see what assumptions they are making about…]
            - How it will make you feel
            - Taste
            - Health
            - Sustainability
            - Natural
            - Function
            - Cost
    - Which of these do you NOT like? What do you not like?
        - Any specific ingredients? Nutrition information? Marketing? Etc.?
        - [Probe to see what assumptions they are making about… same as above]
            - Which attributes are a red flag to you - would NOT BUY a protein beverage if it had ____? Does anyone else agree/disagree?
- I've heard some consumers say they want a "clean label". Do any of you feel the same way when it comes to protein beverages? Why/why not?
- Would you describe any of these products as clean label? Why/why not?
- How important is a clean label to you for protein beverages? Is it more or less important than for other food products? Why?
- What does a clean label look like for protein beverages?

- As we keep talking, I'm going to make notes on the board…

	Clean Label for a Protein Beverage
What does it have in/on it	
What DOESN'T it have in/on it	
Advantages	
Disadvantages	

- [Probe if not brought up]
  - Flavor
  - Amount/Type of protein
  - Type of sweetener
  - Zero sugar
  - Additives
  - Ingredients list (length, recognizability)
  - Macros
  - Vitamins/micronutrients
  - Messaging/claims (complete protein, "burn fat", vegan, organic, non-GMO, grass fed, low fat, probiotics, electrolytes, performance, etc.)
- When you see a protein beverage with a clean label, how does that make you feel?
  - [Probe to see if they make any assumptions about health, price, taste, function, etc.]
- How many protein beverages have clean labels?

- Have you ever thought about why these ingredients are in the product? Does knowing what the ingredients are/do make them more acceptable?
  - What do you think are the motivations of companies putting these ingredients in products? Does that make it okay? Does it make you feel any different?
- I've heard consumers say that "If I see an ingredient I don't know, I will Google it." Does anyone here do that? Why/Why not?
  - What are you looking for when you Google an ingredient? Do you do this in the store or at home?
  - Does Googling the ingredient change how you feel about it? Tell me more?
    - [Probe to understand if they usually find good or bad information]
      - When you find [good/bad] information, does that change whether or not you buy the product? Tell me more about this.

Product Ideation/Creation

- [Individual activity - product creation] I want to do one final activity before we wrap up today. Each of you has a piece of paper in front of you, and I want you to design your perfect RTD protein beverage. I have provided each of you with a template, but I don't want you to feel limited by the examples or by anything we have discussed today. For this exercise I don't want you to be thinking about what is possible or not, just tell me about what you would like to see in a protein beverage

Flavor	
Size	
Nutritional information	
Ingredients you WANT	
Ingredients you DON'T WANT	
Function	
Any other features	

  - [Follow-up group discussion] Now, we don't have time for each of you to share everything, but I want to ask if your creation looks like anything you have seen on the shelf? How is it different?
  - Would you pay more for it?

Wrap-up

- Closing thoughts/questions to wrap up the group…