Improving Non-nutritive Sweetener Study Design Methodology

Emily Anne Myers, BA

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Human Nutrition, Foods, and Exercise

Valisa E. Hedrick, chair
Brenda M. Davy
Kevin P. Davy
Susan E. Duncan

September 17, 2018
Wallace 328
Blacksburg, VA

Keywords: Dietary assessment, food frequency questionnaires, dietary biomarkers, sensory evaluation, non-nutritive sweeteners
Improving Non-nutritive Sweetener Study Design Methodology

Emily Anne Myers

ABSTRACT (Academic)

Non-nutritive sweeteners (NNS) are frequently used as substitutes for added sugars. NNS are difficult to study due to the inability to accurately measure the amounts individuals consume, as well as limitations in study design and methods, including reliance on observational study designs, the use of diet soda as a proxy of total NNS intake, and the grouping of NNS into a single category rather than studying NNS as individual products. New dietary assessment methods and improved study design and methods are needed to advance researchers’ abilities to study NNS and their impact on consumers’ health. The objectives of this dissertation were to 1) determine validity and reproducibility of a novel NNS food frequency questionnaire (NNS-FFQ), 2) develop methodology for an objective NNS urinary biomarker, 3) identify an appropriate carrier for NNS intake in studies, and 4) examine the literature on the relationship between NNS and weight-related outcomes based on study design and methods.

Objectives 1 and 2: participants (n=125) completed three 24-hr dietary recalls, the NNS-FFQ, and 2 24-hr urine samples. NNS intake via NNS-FFQ and recalls were compared using Bland-Altman analyses, with agreement levels ranging from 92.7-99.2% for individual NNS types and total intake. The NNS biomarker methodology was developed using ultra performance liquid chromatography (UPLC-MS/MS), which analyzes each sample for the presence of NNS and related metabolites. This method observed a range from very strong presence of NNS to not detectable, indicating that this biomarker could identify specific NNS consumption (n=9). Objective 3: a sensory evaluation (n=67) was conducted to identify if applesauce or water was a more appropriate carrier for NNS for future interventions. Applesauce was preferred (sucralose=83.6%; aspartame=79.1%; stevia=74.6%) significantly more than water (p≤0.001), indicating that applesauce could be used as an acceptable carrier of NNS in research studies. Objective 4: a systematic literature review focusing on study design and methods used in investigations on NNS and weight-related outcomes found that 81% of RCT had improved weight outcomes, while 76% of observational studies had higher weight outcomes. Improving NNS study design and methods will increase the quality of research conducted on NNS and related health outcomes.
Improving Non-nutritive Sweetener Study Design Methodology

Emily Anne Myers

ABSTRACT (Public)

Artificial sweeteners [non-nutritive sweeteners (NNS)] are often used to replace added sugars. NNS are difficult to study due to the inability to accurately measure the amounts individuals consume, as well as limitations in study design and methods, including reliance on observational study designs, the use of diet soda to represent total NNS intake, and the grouping of NNS into a single category rather than studying NNS as individual products. New dietary assessment tools and improved study design and methods are needed to allow researchers better to study NNS and their impact on health. The objectives of this dissertation were to 1) to determine the ability of a NNS food frequency questionnaire (NNS-FFQ) to measure typical NNS intake, 2) develop methodology for an objective NNS urinary biomarker, 3) identify an appropriate carrier for NNS intake in studies, and 4) to review the currently available research on the relationship between NNS and weight status.

Objectives 1 and 2: 125 participants completed three 24-hr dietary recalls, the NNS-FFQ to measure usual NNS intake, and 24-hour urine samples for a NNS biomarker. Amounts of NNS that consumers reported in recalls were compared to amounts reported in the NNS-FFQ, with the tools finding similar NNS amounts in participants’ diets. The NNS biomarker methodology was developed using ultra performance liquid chromatography (UPLC-MS/MS). This rapid method measures the presence of NNS and related products (saccharin, acesulfame potassium, sucralose, steviol glucuronide, and erythritol) in urine. Among 9 participants, this method identified wide differences, ranging from strong presence of NNS to not measureable. Objective 3: participants (n=67) completed a sensory evaluation (taste testing) study to identify a more appropriate food or beverage (applesauce or water) for NNS (sucralose, aspartame, and stevia) to be mixed with in research studies. Applesauce was preferred over water for all sweeteners (>74%), indicating that applesauce would be an acceptable NNS carrier.

Objective 4: a systematic review on study design and research methods used in studies on NNS and weight-related outcomes. Improving NNS measurement tools and study design methods will improve the quality of research that can be conducted on NNS and related health outcomes.
Acknowledgements

I would like to thank my advisor and mentor, Dr. Valisa Hedrick, for her mentorship, advice, and patience throughout my graduate education. I cannot thank you enough for all the opportunities you’ve provided throughout this program and your endless encouragement. I would like to thank my committee members, Dr. Brenda Davy, Dr. Kevin Davy, and Dr. Susan Duncan, for their support and expertise on this project. I would also like to thank Dr. Andrew Neilson for his guidance and expertise with the biomarker analysis.

I would like to thank my lab mates in the Hedrick lab for creating such a positive and supportive place to learn. I would like to thank Mrs. Heather Cox for her mentorship throughout my dietetics coursework and for encouraging me to grow through teaching experiences. I would like to thank Mrs. Carol Papillon for her mentorship in my dietetic internship. I would also like to thank the Water INTERface Interdisciplinary Graduate Education Program for supporting my graduate education and providing opportunities for collaboration.

Finally, I would like to thank my family and friends who encouraged and supported me throughout this process. To my husband Courtney, our son Benjamin, my parents, and the rest of my family and friends – I could not have done this without your help, cooperation, and understanding.
Table of Contents

Acknowledgements ........................................................................................................ iv
List of Tables .................................................................................................................. vii
List of Figures ............................................................................................................... ix
Attributions ................................................................................................................... x

Chapter 1 – Literature Review ...................................................................................... 1
  Dietary Assessment and Study Design Challenges for NNS ................................. 3
  Dietary Assessment Methodology .......................................................................... 6
  Validation of Dietary Assessment Methods .......................................................... 12
  Sensory Evaluation .................................................................................................. 13
  Conclusion .................................................................................................................. 15
  References .................................................................................................................. 17

Chapter 2 – The Reproducibility and Comparative Validity of a Non-Nutritive Sweetener
  Food Frequency Questionnaire .............................................................................. 31
  Abstract ...................................................................................................................... 31
  Introduction ............................................................................................................... 33
  Materials and Methods ............................................................................................ 35
  Results ......................................................................................................................... 41
  Discussion ................................................................................................................... 51
  Conclusion .................................................................................................................. 56
  References .................................................................................................................. 57

Chapter 3 – Identifying an Appropriate Carrier for Non-nutritive Sweeteners in Metabolic
  and Controlled Feeding Investigations via Sensory Evaluation .......................... 62
  Abstract ...................................................................................................................... 62
  Introduction ............................................................................................................... 63
  Methods ....................................................................................................................... 65
  Results ......................................................................................................................... 70
  Discussion ................................................................................................................... 75
  Conclusions ............................................................................................................... 78
  References .................................................................................................................. 80

Chapter 4 – Non-nutritive Sweetener Urinary Biomarker Methodology .................. 83
  Abstract ...................................................................................................................... 83
  Introduction ............................................................................................................... 85
  Methods ....................................................................................................................... 88
  Results ......................................................................................................................... 92
Discussion.................................................................................................................. 93
References..................................................................................................................... 98

Chapter 5 – A systematic review of study design elements and outcomes of non-nutritive sweetener and weight status research: Where do we go from here?................................. 104

Abstract....................................................................................................................... 104
Introduction.................................................................................................................. 106
Methods....................................................................................................................... 108
Results.......................................................................................................................... 111
Conclusions.................................................................................................................. 131
Future Directions ....................................................................................................... 136
References.................................................................................................................... 138

Appendix A. Institutional Review Board Protocol Approvals ....................................... 149
Appendix B. Risk of bias assessment for observational studies .................................... 156
Appendix C. Risk of bias assessment for randomized controlled trials ......................... 161
List of Tables

Chapter 2

Table 1. Participant demographic characteristics

Table 2. Non-nutritive sweetener (NNS) consumption patterns among 123 adults

Table 3. Test-retest reproducibility of a non-nutritive sweetener food frequency questionnaire (NNS-FFQ) (n=123)

Table 4. Reproducibility statistics of total non-nutritive sweetener (NNS) mg consumption from NNS-FFQ 1 NNS-FFQ 2 recalls based on demographic characteristics

Table 5. Comparative validity of a non-nutritive sweetener food frequency questionnaire (NNS-FFQ) as compared to three 24-hour dietary recalls (n=123)

Table 6. Validity statistics of total non-nutritive sweetener (NNS) mg consumption from NNS-FFQ 2 and dietary recalls based on demographic characteristics (n=123)

Table 7. Descriptive statistics of total and individual non-nutritive sweetener (NNS) mg consumption via two administrations of a NNS food-frequency questionnaire (NNS-FFQ 1 and NNS-FFQ 2) and dietary recalls

Chapter 3

Table 1. Carrier preference for each non-nutritive sweetener type (n=67)

Table 2. Mean hedonic scores for sucralose (a), aspartame (b), and stevia (c) with water or applesauce (n=67)

Table 3. Mean overall affective scores among NNS consumers and non-consumers (n=64)

Chapter 4

Table 1. Optimized mass spectrometry parameters for the quantitation of non-nutritive sweeteners (NNS)

Table 2. Spiking range and mean recovery percentage of previous NNS biomarker investigations
Chapter 5

Table 1. Observational studies evaluating NNS intake and weight status

Table 2. Randomized controlled trials evaluating NNS intake and weight status

Table 3. Study design element frequencies with weight outcomes

Table 4. Weight outcome frequencies based on study design types and methods
List of Figures

Chapter 2

Figure 1. Study design and timeline

Figure 2. Non-nutritive Sweetener Food Frequency Questionnaire (NNS-FFQ)

Figure 3. Bland-Altman plots of total and individual non-nutritive sweetener (NNS) mg consumption via a food-frequency questionnaire (NNS-FFQ) and dietary recalls (n=123)

Figure 4. Bland-Altman plots of total and individual non-nutritive sweetener (NNS) mg consumption via a food-frequency questionnaire (NNS-FFQ) and dietary recalls (n=123)

Chapter 3

Figure 1. Sensory evaluation questionnaires utilized in this investigation

Figure 2. Distribution of frequently selected emotional terms among panelists (n=67)

Chapter 4

Figure 1. Study design and timeline of a non-nutritive sweetener biomarker validity study

Figure 2. Raw chromatogram outputs from a non-nutritive sweetener (NNS) consumer and a NNS non-consumer

Chapter 5

Figure 1. Study selection process flow diagram
## Attributions

### Attributions Chapter 2: The reproducibility and comparative validity of a non-nutritive sweetener food frequency questionnaire

All authors designed the study. Emily Myers and Valisa Hedrick directed its implementation, including quality assurance and control, and designed the study’s analytic strategy. Emily Myers and Valisa Hedrick developed the manuscript and analyzed the data. All authors (Emily Myers, Erin Passaro, and Valisa Hedrick) provided critical feedback and approved the final draft of the manuscript.

This manuscript was published by MDPI in their journal, Nutrients in 2018. PubMed Central link: [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5872752/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5872752/)

### Attributions Chapter 3: Identifying an appropriate carrier for non-nutritive sweeteners in metabolic and controlled feeding investigations via sensory evaluation

All authors were involved in developing the study design concept. Emily Myers and Aili Wang coordinated data collection. Emily Myers conducted statistical analyses and drafted the manuscript. All authors (Emily Myers, Susan Duncan, Aili Wang, and Valisa Hedrick) reviewed and contributed to the final version of the manuscript.

This manuscript is under review in the Journal of Sensory Studies in 2018.

### Attributions Chapter 5: A systematic review of study design elements of non-nutritive sweetener and weight status studies: Where do we go from here?

Emily Myers developed the study design concept and all authors approved of the study methods. Emily Myers conducted article screenings and drafted the manuscript. All authors (Emily Myers, Brenda Davy, Susan Duncan, and Valisa Hedrick) reviewed and contributed to the final version of the manuscript.

This manuscript is in preparation for submission to Obesity Reviews Journal in 2018.
Chapter 1 – Literature Review

Non-nutritive sweeteners (NNS), or artificial sweeteners, are frequently debated for their role among functional foods\(^1\). These low-calorie intense sweeteners have been recommended for their potential to reduce consumption of added sugars, encourage weight loss, and control blood glucose levels\(^2-4\). However, they have also been investigated for their potential negative side effects, such as cancer, insulin resistance, and compensatory appetite\(^5-9\).

NNS are substances that have a concentrated sweet taste in a very small amount of the product. These “high-intensity sweeteners” range between 160 and 1000 times sweeter than sucrose\(^4\). These sweeteners have negligible caloric content due to the very small volume needed to provide a sweet flavor. The Food and Drug Administration (FDA) regulates NNS as food additives with six NNS that are currently FDA approved: saccharin, aspartame, acesulfame potassium, neotame, advantame, and sucralose\(^10\). Two additional NNS, stevia and luo han guo, also known as monk fruit, are currently classified as “Generally Recognized As Safe” or GRAS by the FDA, but have not been evaluated for FDA approval\(^11\). When a food additive has been “adequately shown to be safe” by qualify experts, it is given the GRAS label\(^11\). GRAS labels can be determined by companies without informing the FDA and GRAS labeled products do not require premarket approval from the FDA\(^10\).

Saccharin was the first NNS, discovered in 1879, and has been regulated by the FDA since 1958 when the FDA began regulating food additives. As a sweetener, saccharin is perceived to be 200 to 700 times sweeter than sucrose\(^4\). Saccharin is commonly found in sweetener packets, commonly branded as Sweet’n Low. When consumed, saccharin is not broken down by the body and is excreted unchanged in the urine\(^4,12\).
Aspartame has been approved by the FDA for use in beverages since 1981. As a sweetener, aspartame is perceived to be between 160 and 220 times sweeter than sucrose\(^4\). Aspartame is not heat stable and loses its sweet flavor when heated; therefore, it is not intended for use in baked products. When consumed, aspartame is broken down into aspartic acid, phenylalanine, and menthol\(^{12,13}\), which can be further metabolized into formaldehyde, formic acid, and diketopiperazine\(^14\). Aspartame should not be consumed by individuals with phenylketonuria due to the presence of phenylalanine in the substance. Aspartame is commonly consumed in diet beverages, sweetener packets (commonly branded as Equal), light yogurts, and sugar-free chewing gum.

Acesulfame potassium (Ace K), a derivative of an organic acid and potassium, has been approved by the FDA for use in beverages since 1998. When consumed, Ace K is not broken down by the body and is excreted unchanged\(^12\). As a sweetener, Ace K is perceived to be 200 times sweeter than sucrose\(^4\). Ace K is typically used in combination with other sweeteners (usually aspartame or sucralose). Typical sources include diet soda and sugar-free chewing gum, with sweetener packets of Ace K being less commonly consumed.

Sucralose, which is a chlorinated sucrose molecule, has been approved by the FDA for use in beverages since 1999. Sucralose is water soluble, heat stable, and pH stable\(^15\), making it a popular choice for both food products and beverages, as well as for use in cooking and baking. As a sweetener, sucralose is perceived to be between 600 and 800 times sweeter than sucrose\(^4\). Typical sources include diet soda, sweetener packets (commonly branded as Splenda), and light yogurts. When consumed, the majority is excreted unchanged in the stool with the remainder (~11-27%) being absorbed and excreted unchanged in the urine, with only minor amounts
appearing as metabolites\textsuperscript{14-17}. There are two glucuronide conjugates of sucralose found in urine in addition to unchanged sucralose\textsuperscript{18}.

Stevia, which is derived from the \textit{Stevia rebaudiana} plant, has been approved for use under a GRAS label in beverages since 2008. Stevia is perceived to be about 300 times sweeter than sucrose\textsuperscript{4}. Typical sources include sweetener packets and diet sodas. Brand names of stevia packets include Truvia, PureVia, and SweetLeaf. The sweet flavor of stevia is derived from stevia glycosides, stevioside, rebaudioside (A to F), steviolbioside, and isosteviol\textsuperscript{19}, all of which undergo glucuronidation in the liver and are later excreted in the urine\textsuperscript{20}. As a naturally derived NNS, the popularity of stevia has risen steadily since its introduction into the market in 2008. Between 2009 and 2013, stevia’s market share among NNS food and beverages rose from 5\% to 15\%\textsuperscript{21}. Popularity of plant-derived sweeteners, included stevia, is expected to continue to rise.

When sold in sweetener packets, stevia is often combined with erythritol. Erythritol is a very low-calorie (0.2 kcal/gram) sugar alcohol that is frequently used as a bulking agent and flavor enhancer and is considered to be 60-80\% as sweet as sucrose\textsuperscript{15}.

**Dietary Assessment and Study Design Challenges for NNS**

Despite their widespread adoption by the food industry, the average consumer may be unaware of how many products contain NNS and unable to accurately report how much NNS they consume. NNS can be found in a wide variety of beverages and food products beyond diet soda and sweetener packets. Some products, such as light yogurts, sugar-free coffee creamers, and sugar-free puddings may not be easily recognizable or recallable as artificially sweetened. The number of products that utilize NNS has increased rapidly, with over 6,000 new products released between 1999 and 2004\textsuperscript{22}. A 2018 analysis determined that 4\% of food and beverage
products in the United States utilize NNS. NNS are included in the ingredients list on the Nutrition Facts label, which may be difficult for the average consumer to interpret, especially for food items with a large number of ingredients. Additionally, there are a myriad of terms used by the food industry to label products with NNS, such as “diet,” “light,” “sugar-free,” “reduced-sugar,” and “no sugar added.”

Current research study design methods also contribute to confusion surrounding NNS-related health outcomes. Investigators often rely on cross-sectional study designs rather than randomized controlled trials and this choice of study design likely contributes to research findings. In both randomized controlled trials and cross-sectional studies, diet soda is frequently used as a proxy for measuring NNS intake, excluding a variety of other common dietary sources of NNS. Furthermore, studies tend to examine NNS as a single category rather than as individual compounds. These factors create challenges for investigators when designing robust research studies on NNS to better understand their impacts on the health of consumers.

Reported consumption rates of NNS vary, making it difficult to determine how many people consume them, how much they consume, and from which sources. According to Mattes et al. (2009), 15% of Americans consumed NNS products, with ~11% reporting beverages with NNS and ~6% reporting food products with NNS. This indicates a steady rise in NNS consumption since 1965 when consumption was reported among 3% of the population. Using NHANES data, Sylvestsky et al. (2012) found NNS consumption from 1999–2000 to 2007–2008 increased from ~19% to 24% among US adults. Examining 1999–2008 NHANES data, Drewnowski and Rehm (2015) found that 30% of US adults reported consuming NNS sweetened foods and/or beverages. A 2017 observational study of a rural population in southwest Virginia found that 33% of participants reported consuming NNS, while another study in a similar
geographic region found that NNS intake varied widely (22-50%) depending on which products were included in the analysis\textsuperscript{27}. Clearly, there is a large discrepancy in NNS consumption rates based on a variety of studies, and thus further investigation is needed to provide assessment tools and study methodology to overcome this discrepancy.

The varying and increasing consumption levels reported point to the need to carefully select an accurate assessment tool in observational studies as well as an appropriate NNS carrier in RCT. Currently, there is a lack of consistency in how NNS consumption is quantified or delivered to participants in research studies. Diet beverages are the most frequently consumed NNS product, followed by sweetener packets and food products\textsuperscript{25}. Many observational studies use diet soda as a proxy for NNS consumption\textsuperscript{28,29}, and RCT use a variety of carriers for NNS, including diet soda, beverages, and foods\textsuperscript{3,30}. Few controlled feeding studies have been conducted on NNS and these few used a combination of beverages and foods sweetened with NNS\textsuperscript{31,32}. These differences make it difficult to compare results across studies that utilize different carriers for delivering NNS. The reliance on diet sodas as a representation of NNS intake is a major limitation in the literature, due to the fact that caffeine content as well as carbonation of diet sodas may confound outcomes\textsuperscript{33-35}. The use of diet soda as a proxy for NNS intake can exclude a large percentage of NNS consumers, as one study found vast difference when diet soda alone was used to assess NNS intake (22% of participants were considered NNS consumers) compared to all possible food and beverage NNS sources (50% were considered NNS consumers)\textsuperscript{27}. Moreover, using diet soda as a representation of NNS intake contributes to the perception that all NNS types impact the body in a similar way, rather than as individual compounds with unique effects on the body. Foods and beverages are known to have differing satiating effects that may also impact the effects of introducing NNS into the diet\textsuperscript{36,37}; therefore,
NNS interventions in RCT should be thoughtfully designed with confounding factors as well as participant compliance in mind.

The impact of NNS on health outcomes has been controversial for many years. Early studies on saccharin and aspartame intake in lab animals indicated there may be a risk for certain types of cancer\(^{38,39}\); however, these results were never seen in humans\(^{39,40}\). NNS have also been studied for their impact on both weight loss and weight gain; however, most studies pointing to weight gain are observational studies\(^{28,38,41,42}\), while intervention studies tend to show weight loss\(^3,43\). NNS have also been studied for their impact on compensatory appetite\(^{4,8,44-46}\), cardiovascular disease\(^{1,47-49}\), diabetes\(^{6,45,50-52}\), and migraines\(^{53,54}\). More recently, NNS are being investigated for their impact on gut microbiota\(^{6,55,56}\). The reliance on observational studies to examine health outcomes of NNS pose a problem for understanding the true health impact of these additives. Observational studies can be especially difficult to interpret without accurate and precise methods to identify NNS consumption among participants (e.g., self-reported consumption rates range from 15–50% of the population as mentioned). Determining true consumption rates of NNS and designing more robust research studies on NNS are important steps for improving the understanding of NNS use and associated long-term health outcomes.

**Dietary Assessment Methodology**

Assessing the diets of individuals and populations is challenging with the continuous introduction of new food and beverage products to the market, the evolution of consumption habits, and the limitations of current dietary assessment methods. There are two broad categories of dietary assessment methods, subjective and objective measures. A variety of self-reported subjective assessment methods exist, including 24-hour dietary recalls, food records, and food
frequency questionnaires. Subjective methods of assessment can be influenced by personal feelings or opinion, and thus, self-reported dietary information has limitations due to high levels of error and bias\textsuperscript{57}. However, when working with human subjects, subjective measures are frequently utilized due to their availability and practicality.

Multiple 24-hr dietary recalls are considered the preferred method to measure self-reported dietary intake\textsuperscript{58}. Dietary recalls require a trained interviewer, typically a dietitian, to ask an individual detailed questions regarding their dietary intake from the previous day\textsuperscript{59}. Each dietary recall takes about 20 minutes to administer. While a single day is likely not representative of usual intake, studies have shown that three 24-hr dietary recalls are sufficient to accurately reflect energy intake in adult populations\textsuperscript{58}. Dietary recalls can be especially helpful when measuring recent dietary intake or changes in dietary intake. However, researchers must be trained on proper methods for administering a dietary recall and must build rapport with the individual whose recall is being collected. Because dietary recalls rely on the participants’ memory, it can be burdensome for participants to recall their recent diet, resulting in underreporting\textsuperscript{57}.

Food records are another self-reported assessment method where subjects keep a written record of all foods and beverages consumed over a given period of time, typically between three and nine days. Food records may require participants to either estimate or weigh their food choices when creating their record\textsuperscript{59,60}. One advantage of food records is that they do not rely on the participants’ memory as dietary recalls do. While considered similar in accuracy to dietary recalls, food records increase participant burden by asking them to keep a written record for multiple days and the records may not contain the detailed information required for analysis. When participants are required to complete more than four consecutive food records, reported
intake decreases likely due to participant burden and fatigue\textsuperscript{61}. Food records require that the participant be instructed on the amount of detail required for each food item\textsuperscript{62}. Furthermore, food records may be more susceptible to reporting bias, where knowing their dietary choices must be recorded can cause participants to alter their intake and avoid consuming socially undesirable foods\textsuperscript{57}.

Food frequency questionnaires (FFQ) are a subjective method for dietary assessment that typically focus on one category of food, such as beverages or vegetables, to measure habitual consumption within that category\textsuperscript{63}. According to Margetts and Nelson\textsuperscript{64}, FFQ are questionnaires “in which the respondent is presented with a list of foods and is required to say how often each is eaten in broad terms such as x times per day/per week/per month, etc.”. FFQ can be fully quantitative, measuring how often and how much of each item on a list is consumed, or semi-quantitative, measuring only how often each item on a list is consumed. Time may be measured using increments of days or weeks (e.g., “1 time per week”, “1 time per day”, etc.). Portions can be measured in specific amounts depending on the type of food or beverage (e.g., “Less than $\frac{1}{2}$ cup”, “$\frac{1}{2}$ cup”, “1 cup”, “2+ cups”, or generalized to relative amounts such as small, medium, and large).

Currently, many FFQ have been established to measure a variety of categories of foods and beverages. The National Institute of Cancer Division of Cancer Control and Population Sciences currently reports 162 validated FFQ\textsuperscript{65}. These validated FFQ include categories such as beverages\textsuperscript{66,67}, fruits and vegetables\textsuperscript{68,69}, and dietary fat\textsuperscript{70,71}. Additionally, some FFQ measure specific food additives or components such as sodium\textsuperscript{72,73}, calcium\textsuperscript{74,75}, and carotenoids\textsuperscript{76}. As FFQ are created to measure habitual consumption of a fairly narrow range of food or beverage items, a wide variety of FFQ must be developed.
An advantage of FFQ is that they can be used to gauge habitual dietary intake or recent intake, for example, measuring from the previous year, month, or week, whereas dietary recalls only measure recent intake (i.e., previous day). FFQ can also be self-administered, making it convenient and economical to collect large amounts of data in epidemiological studies. FFQ can be designed to be machine-readable, making them more easily scored for analysis. Although FFQ tend to collect less detailed information than dietary recalls and records, the scale at which they can be used make them useful for large population-level studies.

As subjective assessment methods are associated with increased error and burden, objective assessment methods may be an attractive alternative to researchers. One type of objective measure is dietary biomarkers, which are substances or indicators that can be used to objectively measure dietary intake. The National Institutes of Health defines a biomarker as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to therapeutic interventions.” These markers of intake or nutritional status may also be used to determine the biological outcomes of dietary intake. The National Academy of Medicine (formerly the Institute of Medicine) has encouraged the development and expansion of nutritional biomarkers for the purpose of improving dietary assessment methods, including validating other dietary assessment methods and measuring dietary intake. As objective measures are not influenced by opinion or feelings, biomarkers allow for greater validity and neutrality in dietary assessment.

Biomarkers can be used to accurately assess dietary intake of specific substances without the limitations of self-reported data. These indicators may detect primary compounds or related metabolites of the primary compounds formed in the digestive process. There are three main categories of dietary biomarkers: recovery, predictive, and concentration. Recovery biomarkers
are a type of biomarkers where the exact amount of inputs and outputs of a given substance is known and analyzed\(^83\). The amount collected via excretion is expected to be highly correlated with the amount consumed\(^80\). This type of biomarker is typically used to validate the accuracy of other methods. Examples of recovery biomarkers are doubly-labeled water, used to measure total caloric intake, and urinary nitrogen, used to measure protein consumption\(^80\).

Predictive biomarkers are also used to validate other assessment methods; however, it is understood that the substance will be proportionally represented in the sample, rather than completely reflected in the sample\(^80\). This type of biomarker is particularly helpful for dose-response studies. An example of predictive biomarkers is 24-hour urinary sucrose and fructose, used to measure sugar intake\(^84\). 24-hour urinary sucrose and fructose are expected to be found in small amounts in urine samples but proportionally to one’s sucrose and fructose consumption.

Concentration biomarkers examine the concentration of a specific substance found in the output\(^83\). Concentration biomarkers are especially helpful for measuring associations between diet and disease risk\(^80\). Examples of concentration biomarkers are cholesterol levels in the blood or sodium concentrations in the urine. For each of these biomarkers, a small sample of the specimen is collected to measure the concentration within that specific sample to represent what exists throughout the body.

A variety of specimens, including blood, urine, hair, and nails, can be collected for biomarker analysis. The type of specimen affects the proximity of intake measured as well as the subject burden and cost of the study\(^85\). Urine and blood samples are typically used to evaluate more recent dietary intake, while hair and nail samples reflect long-term intake\(^85\). An appropriate specimen should be chosen depending on the proximity of intake being measured, the accuracy required of the tissue assay, and the subject burden in collecting the tissue. Hair and nail samples
may be collected more easily, while urine and blood samples require more invasive collection methods. Additionally, urine and blood samples lead to an increased cost, with a greater need for trained staff and processing equipment.

Biomarkers can be classified as short-term, medium-term, and long-term depending on the proximity of intake measured\textsuperscript{85}. Short-term biomarkers measure intake as recent as the past hours or days, medium-term biomarkers measure intake over the past weeks or months, and long-term markers measure intake over the past months or years\textsuperscript{85}. Each type of biomarker can be helpful depending on the intended use.

The number of biomarkers being studied has expanded in recent years, as determined in a biomarker review by Hedrick, et al.\textsuperscript{82}. These include food items that may make up a larger portion of one’s diet, such as whole grains\textsuperscript{86,87}, dietary fat\textsuperscript{88,89}, protein\textsuperscript{90,91}, and sugar\textsuperscript{84,92,93}. Biomarkers have also been developed for substances consumed in smaller quantities, including caffeine\textsuperscript{94}, cocoa\textsuperscript{95}, and wine\textsuperscript{96,97}. Hydration status can be measured via a variety of methods, including urinary specific gravity, plasma osmolality, and change in body weight\textsuperscript{98}. These biomarkers vary in their specimen type as well as validity, reliability, and sensitivity.

There are many benefits to incorporating biomarkers into dietary assessment and broader nutrition studies. As mentioned, biomarkers are useful for validating other dietary assessment methods, such as FFQ and 24-hr dietary recalls. Their objective nature minimizes the errors associated with self-reported measures\textsuperscript{99}. Depending on the specimen required, biomarkers may be considered less of a burden to participants than self-reported methods\textsuperscript{85}. Because recovery biomarkers measure a known consumption quantity within a given timeframe, they are considered stronger metabolomic measures than concentration or predictive biomarkers\textsuperscript{100}; however, all three types can be useful for validating self-reported intake. On the other hand,
biomarker use is currently limited to those that have been discovered and validated thus far\textsuperscript{81}. Although specificity is required for biomarkers to be useful, the specificity of biomarkers can limit their utility. While the number of biomarkers available is increasing, current use is limited until new methods are developed\textsuperscript{81}.

**Validation of Dietary Assessment Methods**

In order for a dietary assessment method to be considered useful and accurate, it must first be tested for validity, reliability, and sensitivity to change\textsuperscript{62}. Validity studies should be conducted to demonstrate that the assessment method has both internal and external validity, showing that it measures what it intends to measure and that it can be used in a generalizable population and setting\textsuperscript{101}. Typically, comparative validity is established by comparing measurements or results from a new dietary assessment method to those from established dietary assessment methods, such as 24-hour dietary recalls or food records\textsuperscript{63,102,103}. The use of a validated reference tool allows researchers to determine if correlations in reported data between the two measures merit validity in the new tool. Comparative validity is generally established using correlation statistical analyses which allows researchers to determine how strong the relationship is between the known amount consumed and the amount detected in the specimen\textsuperscript{104}. For an instrument to be considered valid, Spearman’s correlation r values should be between 0.5 and 0.7\textsuperscript{57}.

Reliability studies should be conducted to verify that the results are reproducible in similar settings or over time\textsuperscript{57}. In dietary assessment studies, reliability is typically measured using a test-retest method, where participants are asked to complete the same questionnaire or task within a certain time frame, such as one week\textsuperscript{101}. If the instrument is a reliable measure,
outcomes from the first and second measurements should be significantly correlated. For an instrument to be considered reliable, Spearman’s correlation $r$ values should be between 0.5 and 0.7\textsuperscript{37}.

Finally, sensitivity to change, or responsiveness to change, should be assessed to show how precisely the assessment method can measure change over time as well as the specificity of the method\textsuperscript{105}. In dietary assessment studies, this may be determined through a study intervention, such as a dietary intervention. In these cases, measurements would be taken at baseline prior to the intervention and after the intervention. This pre-test/post-test measurement allows researchers to determine if the measurement tool is capable of detecting various levels of change. Dose response studies have been conducted on a variety of dietary biomarkers, including protein\textsuperscript{106}, omega-3s\textsuperscript{107}, and whole grains\textsuperscript{87}.

**Sensory Evaluation**

Another aspect of NNS study design to be explored is how to deliver NNS to participants as part of an intervention in randomized controlled trials. Sensory evaluation includes a variety of methods “used to evoke, measure, analyze, and interpret [responses] to products as perceived through the senses of sight, smell, touch, taste, and hearing”\textsuperscript{108}. These evaluations allow researchers and food manufacturers to measure and categorize responses to food and beverage products in a blinded setting that eliminates bias\textsuperscript{109}.

Previous investigations have examined the taste interactions between various stimuli. Plain water with no additives has been shown to elicit neutral sensory responses\textsuperscript{110,111}, making it a suitable carrier beverage for many dietary interventions. Wang et al., determined the influence of iron and other minerals on taste perceptions of NNS in water, finding that increased iron and
water hardness increased the perceptions of sweetness of Ace K as well as sucrose and honey\textsuperscript{112}. Leitch et al. found that NNS were considered to be acceptable as an alternative to sucrose when used in beverages\textsuperscript{113}. Water may be an appropriate carrier for NNS due to its satiety qualities as a beverage, allowing it to be compared with other beverages while eliminating other taste interactions.

Though there are many aspects of NNS consumption and related health outcomes to be investigated, researchers must have an improved understanding of how the taste of NNS is perceived depending on the carrier through which it is delivered. Without the ability to accurately understand how tastes are perceived, it is difficult to assess how likely people are to comply with consuming them as a dietary intervention in randomized controlled trials. Improved understanding of differences between and within subjects will allow researchers to better understand the appropriate carrier for these substances.

There are a number of methods used for sensory evaluation including affective attributes testing, Choose-All-That-Apply (CATA) emotional term questionnaires, and paired preference tests. Affective tests allow researchers to assess consumers’ preference or acceptance of a product based on a variety of characteristics, such as sweetness or bitterness\textsuperscript{114}. CATA emotional term questionnaires provide researchers with a method to describe and categorize products with a sensitive measure\textsuperscript{115}. Paired preference tests can be used to determine consumers’ preference for one item versus another. Preferences for three or more samples can be compared using rank preference or multiple paired preference tests\textsuperscript{114}.
Conclusion

Though there are many aspects of NNS consumption and related health outcomes to be investigated, accurate and effective methods for measuring NNS intake must be developed and stronger research studies must be designed. Without evidence showing that NNS consumption can be accurately measured and that participants are willing to be compliant with NNS interventions, it is difficult to assess their true impact on the health of consumers. Developing methods to measure NNS intake and deliver NNS in research studies will improve our understanding of how much of the population consumes NNS and allow researchers to inferentially examine potential health outcomes of NNS consumption.

Currently, the “gold standard” for measuring ad libitum subjective NNS intake is multiple dietary recalls or food records, which are then analyzed with dietary analysis software (Nutrition Data System for Research software [NDS-R])\(^6\). Dietary assessment for all foods has known challenges, which are compounded in the case of NNS. The FDA does not require specific amounts of NNS used in food products to be reported, making it difficult to quantify NNS intake, even with validated dietary analysis software. With new products being developed using different combinations and amounts of NNS, it is difficult for software to keep up with the new products being released. Furthermore, if consumers are unable to identify the products they are consuming as artificially sweetened, dietary recalls may fail to capture NNS in consumers’ diets. To address these limitations, novel measurement tools need to be developed to better assess NNS consumption, including amounts consumed of specific types.

This dissertation addresses many of these limitations via several investigations. A novel NNS food frequency questionnaire (NNS-FFQ) will allow researchers to quickly detect habitual NNS consumption in both clinical and population-level research studies. The NNS-FFQ will also
have utility in a clinical setting, where practitioners seek to better understand patients’ habitual NNS intake. A NNS biomarker will complement the use of the NNS-FFQ and would be useful in research and clinical settings and allow researchers to validate new NNS assessment tools and to design intervention studies aimed at determining both the long-term and short-term effects of NNS use. Identifying an appropriate carrier of NNS via sensory evaluation will allow researchers to improve participant compliance in dietary interventions. Developing both a NNS-FFQ and NNS biomarker as well as identifying an appropriate carrier for NNS in RTC has potential to improve the quality of research studies conducted on NNS and related health outcomes.
References


47. Gardener H, Rundek T, Markert M, Wright CB, Elkind MSV, Sacco RL. Diet soft drink consumption is associated with an increased risk of vascular events in the Northern Manhattan Study. *J Gen Intern Med.* 2012;27(9):1120-1126.

48. Fowler SP. Low-calorie sweetener use and energy balance: Results from experimental studies in animals, and large-scale prospective studies in humans. *Physiol Behav.* 2016;164(Pt B):517-523.


Chapter 2 – The Reproducibility and Comparative Validity of a Non-Nutritive Sweetener Food Frequency Questionnaire

Abstract

In order to better assess non-nutritive sweetener (NNS) consumption, measurement tools with greater utility are needed. The objective of this investigation is to determine the reproducibility and validity of a newly developed NNS food frequency questionnaire (NNS-FFQ) that measures five types of NNS (saccharin, aspartame, acesulfame potassium, sucralose and erythritol). Adult participants (n = 123, 56% female, 75% Caucasian, mean age = 36.8 ± 16.6) completed the NNS-FFQ twice and had 24-hr dietary recalls three times over a two-week study period. Reproducibility between two administrations of the NNS-FFQ was assessed via Bland–Altman plots, Spearman’s correlations (r) and paired samples t-tests. Bland–Altman plots, Cohen’s κ, Spearman’s correlations (r), and paired samples t-tests compared NNS intake between the two methods for validity. For reproducibility analyses, Bland–Altman analyses revealed agreement levels above the 95% acceptance level for total NNS (99.2%), erythritol (99.2%), and aspartame (96.7%). Agreement levels for acesulfame potassium (94.3%), saccharin (94.3%), and sucralose (94.3%) were slightly below the acceptable level. For validity analyses, Bland–Altman analyses revealed agreement levels above the 95% acceptance level for total NNS (95.1%), sucralose (95.9%), saccharin (95.9%), and erythritol (95.1%). Agreement levels for aspartame (94.3%) and acesulfame potassium (92.7%) were slightly below the acceptable level. Although less than desirable agreement was found between the methods for aspartame and acesulfame potassium, some variance was expected due to the habitual nature of the NNS-FFQ as compared to the recent intake reported by recalls. Within the context of this constraint, the NNS-FFQ demonstrates acceptable reproducibility and validity. The NNS-FFQ, the first
questionnaire of its kind, is a brief questionnaire that could be administered among diverse participants at the individual and population levels to measure habitual NNS intake.
Introduction

Non-nutritive sweeteners (NNS), or artificial sweeteners, are substances that have a concentrated sweet taste within a very small amount of the substance. NNS are frequently debated for their role among functional foods. These low-calorie intense sweeteners have been promoted for their potential to reduce added sugar consumption, facilitate weight loss, and control blood glucose levels. However, they have also been investigated for their potential negative side effects, such as cancer, insulin resistance, and compensatory appetite.

These “high-intensity sweeteners” are between 160 and 1000 times sweeter than sucrose. While some of these sweeteners do contain calories, the amount is negligible due to the very small amount needed to provide a sweet flavor. In the United States (US), the Food and Drug Administration (FDA) regulates NNS as food additives with six NNS that are currently FDA approved: saccharin, aspartame, acesulfame potassium, neotame, advantame, and sucralose.

When a food additive has been “adequately shown to be safe” by qualified experts, it is given the label “Generally Recognized As Safe” or GRAS. GRAS labels can be determined by companies without informing the FDA and GRAS products do not require premarket approval from the FDA. Stevia and luo han guo, also known as monk fruit, are currently classified as GRAS by the FDA, but have not been evaluated for FDA approval as food additives. In Europe, NNS are evaluated and regulated by the European Food Safety Authority’s Panel on Food Additives and Nutrient Sources Added to Food (ANS Panel). There are eight NNS approved for use in Europe, acesulfame potassium, aspartame, cyclamate, neohesperidin dihydrochalcone, saccharin, sucralose, thaumatin, and stevia. There are many potential associations between NNS consumption and health outcomes that are yet to be investigated; however, the lack of an accurate and rapid method for measuring NNS intake hinders the
advancement of this topic. The majority of investigations exploring the impact of NNS intake look at NNS intake as a whole or use diet soda as a proxy for NNS consumption\textsuperscript{14}. This approach has inherent issues, as it is not a valid assumption that all NNS have equal impacts on health outcomes. Without the ability to accurately and quickly measure habitual NNS consumption (total and specific types), it is difficult to assess the true impact of NNS on the health of consumers. The availability of such a method would advance knowledge related to how much NNS individuals consume, as well as amounts of specific types of NNS; and thus, allow researchers to inferentially examine potential associated health outcomes.

Food frequency questionnaires (FFQ) may potentially fill this gap. FFQ are subjective dietary assessment tools that measure habitual consumption within certain food or beverage categories over various time periods\textsuperscript{15,16}. FFQ can be fully-quantitative, measuring how often and how much of each item on a list is consumed\textsuperscript{17}, or semi-quantitative, measuring only how often each item on a list is consumed\textsuperscript{15}. The National Institute of Cancer Division of Cancer Control and Population Sciences currently reports 161 validated FFQ\textsuperscript{18}. To our knowledge, there is no validity FFQ specific to NNS intake. To address limitations related to NNS research, valid and reproducible measurement tools need to be developed to better assess NNS consumption, including amounts consumed of specific NNS types. Thus, the objective of this investigation is to determine the reproducibility and comparative validity of a newly developed fully-quantitative NNS-FFQ that can quickly assess NNS consumption in approximately 5–10 min (8th grade reading level), as compared to multiple 24-h dietary recalls, which could take 20–30 min to collect the information from the participant (~60-90 min total), plus nutrient analysis time (~1 hour).
Materials and Methods

Subjects and design

Adult participants residing in southwest Virginia (n = 125) were recruited to participate in this observational study, with 123 participants being included in the final analysis after possible outliers were removed. Eligible participants were English-speaking adults aged 18 years or older. Participants were recruited through traditional methods, including flyers and listservs. This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all study procedures involving human subjects were approved by the Virginia Tech Institutional Review Board (IRB #15-682, approved 25 September 2015). Participants provided written informed consent before enrollment.

Participants completed three visits over the course of two weeks (Figure 1). Measures included the collection of three 24-h dietary recalls and the completion of the NNS-FFQ twice. During the first visit to the laboratory, participants provided demographic information, height, without shoes, was measured in centimeters using a research-grade digital stadiometer, and weight, in light clothing without shoes, was measured to the nearest 0.1 kg using a calibrated digital Tanita scale (Model: TBF-310GS; Tokyo, Japan). During the second visit, participants completed the newly developed NNS-FFQ. A trained research assistant, supervised by a PhD level registered dietitian nutritionist, collected a 24-h dietary recall from the previous day. Between the second and third visit, participants completed a 24-h dietary recall from the previous day. Between the second and third visit, participants completed one unannounced dietary recall via phone call. At the final visit, participants completed NNS-FFQ for a second time, and a third 24-h dietary recall was collected.
Development of the Non-Nutritive Sweetener Food Frequency Questionnaire (NNS-FFQ)

The NNS-FFQ (Figure 2) is a fully-quantitative tool measuring how often a NNS-containing dietary item is consumed (i.e., never, 1 time per week, 2–3 times per week, etc.) over the past month and how much of the product is consumed each time (i.e., <6 fl oz, 1 tablespoon, 1 cookie, etc.). To create this questionnaire, common sources of NNS were identified\(^{14}\), and then Nutrition Data System for Research (NDS-R) 2015 was used to characterize each type of food or beverage by the types of NNS used. The first page of the NNS-FFQ is comprised of beverages sweetened with NNS and the second page includes sweetener packets as well as food items, including but not limited to yogurts, ice cream, candy, and chewing gum. Each line of the FFQ represents a unique combination and amount of NNS used. For example, some diet sodas are sweetened with aspartame and acesulfame potassium, while others are sweetened with sucralose and acesulfame potassium. Additionally, the NNS-FFQ identifies products by categories as well as by brand names, allowing participants to more easily identify NNS products they consume.
The questionnaire gathers data on five NNS: acesulfame potassium, aspartame, saccharin, sucralose, and stevia products that use erythritol as a bulking agent. Erythritol was measured rather than stevia as NDS-R 2015 does not yet report stevia content but does include grams of erythritol in its database. Erythritol is a non-caloric sugar alcohol frequently used as a bulking agent with stevia products. Other sugar alcohols were not included in this analysis, since they do contain nutritive content\textsuperscript{19} and are not typically categorized as NNS. While using erythritol for the analysis may present some limitations, it is currently the best option available for the validation of the NNS-FFQ until dietary data on stevia becomes more available.

Instructions on the questionnaire state that the participant should respond based on intake over the previous month, review each category, and indicate how often each item is consumed and how much is consumed each time. In addition to the categories listed, there is an option for “Other” NNS products for when participants recognize that a product they consume is artificially sweetened but not listed. The NNS-FFQ administration time is between five and ten minutes, with an additional five minutes of scoring time. Scoring instructions are freely available from the corresponding author upon request.
Figure 2. Non-nutritive Sweetener Food Frequency Questionnaire (NNS-FFQ)

Artificial Sweetener Intake Questionnaire

Instructions:
- In the past month, please indicate your response for each food or beverage item by marking an “X” in the box for “how often” and “how much each time”.
  1. Indicate how often you consumed the following items. For example, you drink diet soda 5 times per week, mark 4-6 times per week.
  2. Indicate the approximate amount of each item you purchased each time.
  3. Count packets of artificial sweeteners added to foods/beverages at the top of page 2 in the artificial sweetener packet category.
  4. Please complete both the front and back of the questionnaire.

<table>
<thead>
<tr>
<th>Beverages</th>
<th>HOW OFTEN (MARK ONE)</th>
<th>HOW MUCH EACH TIME (MARK ONE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never or &lt;1x/week</td>
<td>1x/week</td>
</tr>
<tr>
<td><strong>Flavored Water</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gatorade G2, Propel Zero</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td><strong>Sugar-Free sparkling or carbonated water</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet Tonic water</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>VitaZest, FruiDC, Sugar-Free flavored water</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Light</strong></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Tang, Crystal light (packets for water bottle), Sugar-Free Kool-Aid (dry mix), Country Time Light Lemonade (dry mix)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ocean Spray Lightstyle Juice, Orchard Light Cranberry Juice, Tropicana Twister Light, Diet Snapple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ocean Spray Light or Light cranberry juice</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Juice or Flavored Drink</strong></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td><strong>Diet Sodas Plop Soft Drink</strong></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Shasta, Rite Pure Zero, Coke with Splenda, Diet 7UP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet Pepsi, Wild Cherry, Vanilla, Lime, or Max</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet Coke, Vanilla, Lime, or Cherry, Coke Zero</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Diet Tea</strong></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Crystal Light decaffeinated, Diet Nestea</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet Lipton Instant Iced Tea Mix</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td><strong>Chocolate Drink</strong></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Sugar-Free Nestle Nesquik or Sugar-Free Swiss Miss</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Coffee</strong></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Frappuccino Light, ready to drink</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Energy Drink</strong></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Sugar-Free Versions: SoBe, No Fear, AMP, Rockstar, Monster, Venom Majave</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sugar-Free Red Bull</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td><strong>Protein Drink</strong></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Carnation Instant Breakfast No Sugar Added, Slim-fast Easy Digest, Slim-fast Optima</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slim-fast Meal Low Carb</td>
<td></td>
<td>O</td>
</tr>
</tbody>
</table>

Please continue to the next page to complete this questionnaire. Please circle each item you consumed.
Figure 2. Non-nutritive Sweetener Food Frequency Questionnaire (NNS-FFQ) (continued)

### Artificial Sweetener Intake Questionnaire

<table>
<thead>
<tr>
<th>Artificial Sweetener Packets</th>
<th>HOW OFTEN (MARK ONE)</th>
<th>HOW MUCH EACH TIME (MARK ONE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never or &lt;1x/week</td>
<td>1x/week</td>
</tr>
<tr>
<td>Sweet N Low Powder or Sugar Twin (Saccharin) - Pink Packets</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Equal or NutraSweet The Original (Aspartame) - Blue Packets</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Splenda (Sucralose) - Yellow Packets</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Sweet One (Acosulfame Potassium)</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Truvia (Stevia)</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Stevia In The Raw Packets (Stevia)</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

**Yogurt and Ice Cream**

<table>
<thead>
<tr>
<th>Yogurt and Ice Cream</th>
<th>Never or &lt;1x/week</th>
<th>1x/week</th>
<th>2-3x/week</th>
<th>4-6x/week</th>
<th>1x/day</th>
<th>2x/day</th>
<th>3x/day</th>
<th>&lt;1/2 cup</th>
<th>1 cup</th>
<th>1/2 cups</th>
<th>2 cups</th>
<th>3 cups</th>
<th>&gt;2 cups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breyers Light Yogurt</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yoplait Light Thick &amp; Creamy</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Blue Bunny Light Yogurt</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dannon Light &amp; Fit or Activia Light Yogurt</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yoplait Light Fat Free Yogurt</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Light Ice Creams: Edy's, Dreyer's, Blue Bell, Baskin-Robbins, Popsicle</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Blue Bunny Light Ice Cream</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free Creamsciple or Dreamscicle</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Cookies and Bars**

<table>
<thead>
<tr>
<th>Cookies and Bars</th>
<th>Never or &lt;1x/week</th>
<th>1x/week</th>
<th>2-3x/week</th>
<th>4-6x/week</th>
<th>1x/day</th>
<th>2x/day</th>
<th>3x/day</th>
<th>&lt;1 cookie or bar</th>
<th>1 cookie or bar</th>
<th>2 cookies or bars</th>
<th>3 cookies or bars</th>
<th>4+ cookies or bars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugar-Free Versions: Pillsbury, Nabisco, Murray Cookies, Slim-fast Snack/Meal Bar</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free Versions: Quaker Chewy, Power Bar, Snackwell, Tastylake Sensibles, Pepperidge Farm Milano, Slim-fast Meal Bar (Chocolate Peanut Caramel)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Candy**

<table>
<thead>
<tr>
<th>Candy</th>
<th>Never or &lt;1x/week</th>
<th>1x/week</th>
<th>2-3x/week</th>
<th>4-6x/week</th>
<th>1x/day</th>
<th>2x/day</th>
<th>3x/day</th>
<th>&lt;1 piece</th>
<th>1 piece</th>
<th>2 pieces</th>
<th>3 pieces</th>
<th>&gt;4 pieces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugar-Free Chewing gum</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free Versions: Twizzler, Jolly Rancher, Fifty 50, Sweet 'N Low Fruit Splash, Werther's</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free Chocolate (Crystal Light Candy, DeMent's, Dove, Fifty 50, York Peppermint Patty)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free Gum drops, Gummy worms/bears</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Jello and Pudding**

<table>
<thead>
<tr>
<th>Jello and Pudding</th>
<th>Never or &lt;1x/week</th>
<th>1x/week</th>
<th>2-3x/week</th>
<th>4-6x/week</th>
<th>1x/day</th>
<th>2x/day</th>
<th>3x/day</th>
<th>&lt;1/2 cup</th>
<th>1 cup</th>
<th>1 1/2 cups</th>
<th>2 cups</th>
<th>3 cups</th>
<th>&gt;2 cups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugar-Free or Reduced-Calories Jell-O Pudding (cooked or instant)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free or Reduced-Calories Jell-O Pudding Cup, Hunt's Snack Pack/Pudding Cup (ready to eat)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free or Reduced-Calories Jell-O Gelatin (prepared from dry mix)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free or Reduced-Calories Jell-O Gelatin (ready to eat)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free or Reduced-Calories Hunt's Snack Pack Gelatin Cup (ready to eat)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Jelly**

<table>
<thead>
<tr>
<th>Jelly</th>
<th>Never or &lt;1x/week</th>
<th>1x/week</th>
<th>2-3x/week</th>
<th>4-6x/week</th>
<th>1x/day</th>
<th>2x/day</th>
<th>3x/day</th>
<th>&lt;1 tsp</th>
<th>1 tsp</th>
<th>2 tsp</th>
<th>3 tsp</th>
<th>&gt;3 tsp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugar-Free Jelly with sucrose (Smucker's Sugar-Free with Splenda, Great Value Sugar-Free Preserves)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sugar-Free Jelly with stevia or Truvia (Smucker's Sugar-Free with Truvia)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Coffee Creamer**

<table>
<thead>
<tr>
<th>Coffee Creamer</th>
<th>Never or &lt;1x/week</th>
<th>1x/week</th>
<th>2-3x/week</th>
<th>4-6x/week</th>
<th>1x/day</th>
<th>2x/day</th>
<th>3x/day</th>
<th>&lt;1 serving</th>
<th>1 serving</th>
<th>1 1/2 servings</th>
<th>2 servings</th>
<th>3 servings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugar-Free Flavored Creamer (powder or liquid), Sugar-Free Coffee mate or International Delight</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Other Products with Artificial Sweeteners:**

<table>
<thead>
<tr>
<th>Other Products with Artificial Sweeteners</th>
<th>Never or &lt;1x/week</th>
<th>1x/week</th>
<th>2-3x/week</th>
<th>4-6x/week</th>
<th>1x/day</th>
<th>2x/day</th>
<th>3x/day</th>
<th>&lt;1/2 serving</th>
<th>1 serving</th>
<th>1 1/2 servings</th>
<th>2 servings</th>
<th>3 servings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other (List):</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Dietary recalls

Dietary recalls were collected by trained graduate-level research assistants. Three recalls were collected on non-consecutive days, including two weekdays and one weekend day, using a multi-pass method. Dietary recalls were analyzed using NDS-R 2015 nutrition analysis software (Nutrition Coordinating Center, University of Minnesota, Minneapolis, MN, USA). For consistency, each participant worked with the same research assistant for all three dietary recalls and the research assistant collecting the recalls was also responsible for the data entry. Dietary recalls were analyzed to determine NNS consumption for the days reported. NDS-R 2015 provides data on saccharin, aspartame, sucralose, acesulfame potassium, and erythritol in products, but currently not for stevia.

Consumer versus non-consumer analysis

Using a published novel method to categorize NNS consumers and non-consumers, participants were identified as NNS consumers if they reported consuming the NNS equivalent of 1 fl oz of diet soda from all foods and beverages. This intake level corresponds to 3 mg acesulfame potassium, 17 mg aspartame, 12 mg saccharin, or 6 mg sucralose. Erythritol was not included in this analysis due to the fact that it is not used in stevia-sweetened soda.

Statistical analysis

Descriptive statistics (mean ± standard deviation and frequencies) were used to assess participant demographics and NNS consumption patterns. The data were analyzed for normality using a Shapiro–Wilk test. This analysis determined that total NNS intake was not normally distributed. Thus, participants with NNS intake greater than ±3 standard deviations from the mean were removed (n = 2), giving a final analytical sample of 123. Reproducibility of the NNS-FFQ was assessed by comparing the quantities of each NNS type (and total NNS intake) reported
in the NNS-FFQ at time 1 (visit 2) and time 2 (visit 3) using Bland–Altman analyses, Spearman’s correlations (r) and paired samples t-tests. The comparative validity of the NNS-FFQ was measured by comparing quantities of each NNS type (and total NNS intake) reported in the second administration of the NNS-FFQ to the quantities reported in the participants’ three-day average of their dietary recalls via Bland–Altman analyses, Spearman’s correlations (r), and paired samples t-tests. When interpreting Bland–Altman plots, an agreement level of 95% was considered acceptable\textsuperscript{20–23}. The second administration of the NNS-FFQ was used in the validity analyses as it measures intake over the past month, and was thus representative of the same time period as the 24-h dietary recalls. Cohen’s κ was used to determine the level of agreement between the two methods for identifying NNS consumers vs. non-consumers. An a priori significance level was set at \( p \leq 0.05 \). Statistical analyses were conducted using IBM SPSS statistical analysis software (v. 24 for Mac, 2016, SPSS Inc., Chicago, IL, USA).

**Results**

**Demographic characteristics**

All enrolled adults completed the study (\( n = 125 \)); however, two outliers were removed giving an analytical sample of 123 adults. Participants were mainly Caucasian (75.6\%) with an age range of 18–86 years old. Mean body mass index (BMI) was considered slightly overweight (26.0 kg/m\(^2\)), but 55\% of the participants had a normal BMI. Income was widely varied, however the majority of participants had a college degree (79\%) (Table 1). Table 2 details NNS consumption patterns, including the number of participants who reported consuming any amount of the five types of NNS based on dietary recalls.
Table 1. Participant demographic characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total Sample (n=123), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>54 (44)</td>
</tr>
<tr>
<td>Female</td>
<td>69 (56)</td>
</tr>
<tr>
<td><strong>Mean age ± SD (years)</strong></td>
<td>36.8 ± 16.6</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>93 (75)</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>17 (14)</td>
</tr>
<tr>
<td>African American</td>
<td>7 (6)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3 (2.5)</td>
</tr>
<tr>
<td>More than 1 race</td>
<td>3 (2.5)</td>
</tr>
<tr>
<td><strong>BMI (kg/m²)</strong></td>
<td></td>
</tr>
<tr>
<td>Mean BMI ± SD</td>
<td>26.0 ± 5.7</td>
</tr>
<tr>
<td>Underweight (≤18.4)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Normal weight (18.5-24.9)</td>
<td>68 (55)</td>
</tr>
<tr>
<td>Overweight (25-29.9)</td>
<td>32 (26)</td>
</tr>
<tr>
<td>Obese (≥30)</td>
<td>22 (18)</td>
</tr>
<tr>
<td><strong>Education Level</strong></td>
<td></td>
</tr>
<tr>
<td>High School Graduate</td>
<td>6 (5)</td>
</tr>
<tr>
<td>Some College</td>
<td>20 (16)</td>
</tr>
<tr>
<td>College Graduate</td>
<td>45 (37)</td>
</tr>
<tr>
<td>Graduate School</td>
<td>52 (42)</td>
</tr>
<tr>
<td><strong>Household Income ($)</strong></td>
<td></td>
</tr>
<tr>
<td>≤14,999</td>
<td>19 (15)</td>
</tr>
<tr>
<td>15,000-29,999</td>
<td>29 (23.5)</td>
</tr>
<tr>
<td>30,000-49,999</td>
<td>12 (10)</td>
</tr>
<tr>
<td>50,000-99,999</td>
<td>29 (23.5)</td>
</tr>
<tr>
<td>≥100,000</td>
<td>22 (18)</td>
</tr>
<tr>
<td>No response</td>
<td>12 (10)</td>
</tr>
</tbody>
</table>
Table 2. Non-nutritive sweetener (NNS) consumption patterns among 123 adults

<table>
<thead>
<tr>
<th>NNS Type</th>
<th>Number of participants reporting any consumption via dietary recall n (%)</th>
<th>Number of participants reporting any consumption via FFQ 1 n (%)</th>
<th>Number of participants reporting any consumption via FFQ 2 n (%)</th>
<th>Cohen’s κ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acesulfame Potassium</td>
<td>58 (47)</td>
<td>93 (76)</td>
<td>91 (74)</td>
<td>0.681***</td>
</tr>
<tr>
<td>Aspartame</td>
<td>68 (55)</td>
<td>83 (68)</td>
<td>84 (68)</td>
<td>0.417***</td>
</tr>
<tr>
<td>Saccharin</td>
<td>7 (6)</td>
<td>21 (17)</td>
<td>17 (14)</td>
<td>0.601***</td>
</tr>
<tr>
<td>Sucralose</td>
<td>27 (22)</td>
<td>68 (55)</td>
<td>53 (43)</td>
<td>0.517***</td>
</tr>
<tr>
<td>Erythritol a</td>
<td>2 (2)</td>
<td>5 (4)</td>
<td>5 (4)</td>
<td>n/a b</td>
</tr>
</tbody>
</table>

aErythritol values were converted from grams to milligrams to compare values across NNS types.
bCohen’s κ was not included for erythritol due to the inability to classify participants as consumers or non-consumers. ***p≤0.001.

Because the study population was quite diverse, one-way analysis of variance (ANOVA) tests were run to determine differences in total NNS reported via each assessment method between demographic groups. No statistical differences were found in NNS intake reported in the first or second administration of the NNS-FFQ or dietary recalls based on based on sex, age (ages 18–64 and ages 65+), race (Caucasian and non-Caucasian), and BMI (underweight/normal weight and overweight/obese). Significant differences were detected between groups based on education (high school degree or less and some college or more) for total NNS reported via dietary recalls (F = 4.407, p<0.01).

Test-retest reproducibility

When comparing the consumption of individual NNS types between the first and second administration of the NNS-FFQ, Bland-Altman analyses revealed strong agreement for total NNS (99.2%), erythritol (99.2%), and aspartame (Figure 2). Acesulfame potassium (94.3%), saccharin (94.3%), and sucralose (94.3%) agreement levels were slightly below the acceptable 95% value.
Figure 3. Bland-Altman plots of total and individual non-nutritive sweetener (NNS) mg consumption via two administrations of a NNS food frequency questionnaire (NNS-FFQ 1 and NNS-FFQ 2) (n=123)

Note: The center line represents the mean difference and the upper and lower lines indicate the mean ± 1.96 times the standard deviation.
In reproducibility analyses examining correlations and mean differences, all NNS types, as well as total NNS intake, were found to be significantly correlated (Table 3). No significant differences were found between acesulfame potassium, aspartame, saccharin, erythritol, or total NNS intake; however, a small significant mean difference (p ≤ 0.05) was found for reported sucralose values. The range of total NNS reported at each administration varied greatly, with the first administration ranging from 0.0 to 6079.1 mg and the second administration ranging from 0.0 to 1221.0 mg.

Table 3. Test-retest reproducibility of a non-nutritive sweetener food frequency questionnaire (NNS-FFQ) (n=123)

<table>
<thead>
<tr>
<th>NNS Type</th>
<th>NNS-FFQ Time 1 Mean ± SD&lt;sup&gt;a&lt;/sup&gt; (Median, range)</th>
<th>NNS-FFQ Time 2 Mean ± SD&lt;sup&gt;a&lt;/sup&gt; (Median, range)</th>
<th>Correlation (r)</th>
<th>Mean Difference (Mean ± SE)&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acesulfame Potassium (mg)</td>
<td>18.6 ± 28.9 (5.0, 0.0-132.5)</td>
<td>18.8 ± 32.6 (5.0, 0.0-169.5)</td>
<td>0.81***</td>
<td>0.2 ± 2.1</td>
</tr>
<tr>
<td>Aspartame (mg)</td>
<td>35.3 ± 67.8 (7.0, 0.0-383.1)</td>
<td>38.7 ± 85.1 (7.1, 0.0-694.5)</td>
<td>0.81***</td>
<td>3.4 ± 6.4</td>
</tr>
<tr>
<td>Saccharin (mg)</td>
<td>3.3 ± 11.4 (0.0, 0.0-78.5)</td>
<td>2.9 ± 9.2 (0.0, 0.0-52.3)</td>
<td>0.87***</td>
<td>0.5 ± 0.5</td>
</tr>
<tr>
<td>Sucralose (mg)</td>
<td>18.2 ± 37.3 (2.7, 0.0-250.9)</td>
<td>12.1 ± 28.0 (0.0, 0.0-144.6)</td>
<td>0.77***</td>
<td>6.2 ± 2.4*</td>
</tr>
<tr>
<td>Erythritol (mg)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>65.3 ± 548.7 (0.0, 0.0-6000.0)</td>
<td>25.3 ± 137.9 (0.0, 0.0-1071.4)</td>
<td>0.81***</td>
<td>40.1 ± 49.4</td>
</tr>
<tr>
<td>Total NNS (mg)</td>
<td>140.8 ± 566.8 (24.0, 0.0-6079.1)</td>
<td>97.6 ± 194.4 (24.6, 0.0-1221.0)</td>
<td>0.92***</td>
<td>43.8 ± 48.4</td>
</tr>
</tbody>
</table>

<sup>a</sup>Reported values are means ± standard deviation. <sup>b</sup>Mean differences ± standard error according to a paired sample t test, slight differences may be noted from the preceding columns due to rounding. <sup>c</sup>Erythritol values have been converted from grams to milligrams. *P≤0.05 ***P≤0.001

Reproducibility analyses were also conducted to detect differences within demographic groups, including sex (male and female), age (18–64 years and ≥65 years), race (white and non-white), education (high school or less and some college or more), and BMI (underweight/normal
weight and overweight/obese) (Table 4). When comparing reproducibility analyses within demographic groups, there were no significant mean differences among demographic groups and Spearman’s correlations were all significant (range: 0.74–0.93; all \( p \leq 0.05 \))

**Table 4.** Reproducibility statistics of total non-nutritive sweetener (NNS) mg consumption from NNS-FFQ 1 and NNS-FFQ 2 based on demographic characteristics

<table>
<thead>
<tr>
<th></th>
<th>Total NNS Mean Difference ± SE</th>
<th>Spearman’s Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>36.8±24.5</td>
<td>0.93**</td>
</tr>
<tr>
<td>Female</td>
<td>16.2±25.8</td>
<td>0.92**</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–64 years</td>
<td>46.6±51.3</td>
<td>0.92*</td>
</tr>
<tr>
<td>65+ years</td>
<td>12.9±11.1</td>
<td>0.89**</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>50.7±63.9</td>
<td>0.93**</td>
</tr>
<tr>
<td>Non-White</td>
<td>20.0±11.4</td>
<td>0.74**</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School or Less</td>
<td>20.8±45.1</td>
<td>0.77*</td>
</tr>
<tr>
<td>Some College or More</td>
<td>46.5±50.8</td>
<td>0.92**</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight/Normal Weight</td>
<td>10.4±17.0</td>
<td>0.90**</td>
</tr>
<tr>
<td>Overweight/Obese</td>
<td>111.7±107.8</td>
<td>0.93**</td>
</tr>
</tbody>
</table>

*P≤0.05, **P≤0.01

**Comparative validity**

Bland–Altman analyses comparing reported consumption of total and each type of NNS (Figure 4) were conducted to measure agreement of mg intake between the second NNS-FFQ administration and dietary recalls. The second NNS-FFQ was used in this analysis to reflect the same time period during which the dietary recalls were collected. The Bland–Altman analyses revealed agreement levels above the acceptable 95%\(^{\text{20-23}}\) for total NNS (95.1%), sucralose (95.9%), saccharin (95.9%), erythritol (95.1%) and slightly below for acesulfame potassium (92.7%) and aspartame (94.3%).
Figure 4. Bland-Altman plots of total and individual non-nutritive sweetener (NNS) mg consumption via a food-frequency questionnaire (NNS-FFQ) and dietary recalls (n=123)\textsuperscript{ab}

\textsuperscript{a}The center line represents the mean difference and the upper and lower lines indicate the mean ± 1.96 times the standard deviation. \textsuperscript{b}Values based on natural log transformations due to the non-normal distribution of data.

When assessing the validity of the NNS-FFQ as compared to the dietary recalls, Spearman’s r values for the five sweeteners (acesulfame potassium, aspartame, saccharin, and
sucralose) and total NNS ranged from 0.51 to 0.59 (p ≤ 0.001) with the exception of erythritol (r = −0.03) (Table 5). Significant mean differences were found between the NNS-FFQ and dietary recalls for reported acesulfame potassium values (12.0 ± 27.0 mg, p ≤ 0.001). No significant differences were found between aspartame, sucralose, saccharin, erythritol, and total NNS intake.

Table 5. Comparative validity of a non-nutritive sweetener food frequency questionnaire (NNS-FFQ) as compared to three 24-hour dietary recalls (n=123)

<table>
<thead>
<tr>
<th>Non-nutritive Sweetener Type</th>
<th>NNS-FFQ Mean ± SDa (Median, range)</th>
<th>Dietary Recall Mean ± SDa (Median, range)</th>
<th>Correlation (r)</th>
<th>Mean Difference (Mean ± SE)b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acesulfame Potassium (mg)</td>
<td>18.8 ± 32.6 (6.0, 0.0-169.5)</td>
<td>6.8 ± 16.4 (0.0, 0.0-99.3)</td>
<td>0.51**</td>
<td>12.0 ± 2.4***</td>
</tr>
<tr>
<td>Aspartame (mg)</td>
<td>38.7 ± 85.1 (7.1, 0.0-694.5)</td>
<td>36.5 ± 97.2 (1.2, 0.0-526.4)</td>
<td>0.59**</td>
<td>2.2 ± 6.7</td>
</tr>
<tr>
<td>Saccharin (mg)</td>
<td>2.9 ± 9.2 (0.0, 0.0-52.3)</td>
<td>3.9 ± 19.6 (0.0, 0.0-170.9)</td>
<td>0.55**</td>
<td>1.0 ± 1.5</td>
</tr>
<tr>
<td>Sucralose (mg)</td>
<td>12.1 ± 28.0 (0.0, 0.0-144.6)</td>
<td>8.0 ± 19.7 (0.0, 0.0-103.7)</td>
<td>0.53**</td>
<td>4.0 ± 2.1</td>
</tr>
<tr>
<td>Erythritol (mg)c</td>
<td>25.3 ± 137.9 (0.0, 0.0-1071.4)</td>
<td>16.3 ± 134.0 (0.0, 0.0-1333.3)</td>
<td>-0.03</td>
<td>9.0 ± 17.5</td>
</tr>
<tr>
<td>Total NNS (mg)</td>
<td>97.6 ± 194.4 (25.9, 0.0-1221.0)</td>
<td>72.4 ± 183.6 (8.0, 0.0-1335.7)</td>
<td>0.55**</td>
<td>25.3 ± 18.0</td>
</tr>
</tbody>
</table>

aReported values are means ± standard deviation. bMean differences ± standard error according to a paired sample t test, slight differences may be noted from the preceding columns due to rounding. cErythritol values have been converted from grams to milligrams to allow values to be compared across all NNS. **P≤0.01, ***P≤0.001

Additional comparisons were made to determine if the NNS-FFQ was able to categorize NNS consumers and non-consumers in a similar way to dietary recalls. Based on dietary recall data, 80 participants, or 65.0%, were categorized as NNS consumers compared to 77 participants, or 62.6%, based on the NNS-FFQ data. Cohen’s κ demonstrated a substantial level of agreement between the two methods for identifying NNS consumers vs. non-consumers when looking at reported total NNS consumption, κ=0.669 (p≤0.001). When examining the κ value for the individual types of NNS, a substantial level of agreement was demonstrated for acesulfame
potassium consumers ($\kappa=0.618$, $p\leq0.001$), and moderate levels of agreement for aspartame, saccharin, and sucralose ($\kappa=0.417$, 0.601, and 0.517, respectively; all $p\leq0.001$).

Validity analyses were also conducted to detect differences within demographic groups, including sex (male and female), age (18–64 years and ≥65 years), race (white and non-white), education (high school or less and some college or more), and BMI (underweight/normal weight and overweight/obese) (Table 6). When comparing validity analyses within demographic groups, there were no significant mean differences based on sex, age, education or BMI, and Spearman’s correlations were all significant (range: 0.44–0.96; all $p\leq0.05$). A significant mean difference was detected among white participants (38.4 ± 18.0; $p=0.04$) and Spearman’s correlation among non-white participants was non-significant ($r=0.10$).

**Table 6.** Validity statistics of total non-nutritive sweetener (NNS) mg consumption from NNS-FFQ 2 and dietary recalls based on demographic characteristics (n=123)

<table>
<thead>
<tr>
<th>Demographic Categories</th>
<th>Total NNS Mean Difference ± SE</th>
<th>Spearman’s Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>30.9±19.2</td>
<td>0.69**</td>
</tr>
<tr>
<td>Female</td>
<td>101.2±84.5</td>
<td>0.44**</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–64 years</td>
<td>23.0±18.8</td>
<td>0.53**</td>
</tr>
<tr>
<td>65+ years</td>
<td>62.9±55.2</td>
<td>0.96**</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>38.4±18.0*</td>
<td>0.63**</td>
</tr>
<tr>
<td>Non-White</td>
<td>15.6±47.9</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School or Less</td>
<td>70.3±112.5</td>
<td>0.75*</td>
</tr>
<tr>
<td>Some College or More</td>
<td>30.2±18.0</td>
<td>0.55**</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight/Normal Weight</td>
<td>15.9±27.4</td>
<td>0.44**</td>
</tr>
<tr>
<td>Overweight/Obese</td>
<td>37.2±21.4</td>
<td>0.64**</td>
</tr>
</tbody>
</table>

*P≤0.05, **P≤0.01
<table>
<thead>
<tr>
<th></th>
<th>Mean (mg)</th>
<th>Std. Deviation</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Skewness Statistic</th>
<th>Std. Error</th>
<th>Kurtosis Statistic</th>
<th>Std. Error</th>
<th>Shapiro-Wilk Tests of Normality</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNS-FFQ 1 Erythritol (mg)</td>
<td>65.3</td>
<td>548.7</td>
<td>0.00</td>
<td>0.00</td>
<td>6000.0</td>
<td>10.6</td>
<td>0.2</td>
<td>114.8</td>
<td>0.4</td>
<td>≤0.001</td>
<td></td>
</tr>
<tr>
<td>NNS-FFQ 1 Acesulfame Potassium (mg)</td>
<td>18.6</td>
<td>28.9</td>
<td>5.0</td>
<td>0.00</td>
<td>132.5</td>
<td>2.2</td>
<td>0.2</td>
<td>4.2</td>
<td>0.4</td>
<td>0.7</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 1 Aspartame (mg)</td>
<td>35.3</td>
<td>67.8</td>
<td>7.0</td>
<td>0.00</td>
<td>383.3</td>
<td>2.9</td>
<td>0.2</td>
<td>8.7</td>
<td>0.4</td>
<td>0.6</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 1 Saccharin (mg)</td>
<td>3.3</td>
<td>11.5</td>
<td>0.0</td>
<td>0.00</td>
<td>78.5</td>
<td>4.6</td>
<td>0.2</td>
<td>22.0</td>
<td>0.4</td>
<td>0.3</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 1 Sucralose (mg)</td>
<td>18.2</td>
<td>37.6</td>
<td>2.7</td>
<td>0.00</td>
<td>250.9</td>
<td>3.7</td>
<td>0.2</td>
<td>17.4</td>
<td>0.4</td>
<td>0.5</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 1 Total (mg)</td>
<td>140.8</td>
<td>566.8</td>
<td>24.0</td>
<td>0.00</td>
<td>6079.1</td>
<td>9.7</td>
<td>0.2</td>
<td>100.9</td>
<td>0.4</td>
<td>0.2</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 2 Erythritol (mg)</td>
<td>25.3</td>
<td>137.9</td>
<td>0.0</td>
<td>0.00</td>
<td>1071.4</td>
<td>6.1</td>
<td>0.2</td>
<td>39.1</td>
<td>0.4</td>
<td>0.2</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 2 Acesulfame Potassium (mg)</td>
<td>18.8</td>
<td>32.6</td>
<td>5.0</td>
<td>0.00</td>
<td>169.5</td>
<td>2.5</td>
<td>0.2</td>
<td>6.1</td>
<td>0.4</td>
<td>0.6</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 2 Aspartame (mg)</td>
<td>38.7</td>
<td>85.1</td>
<td>7.1</td>
<td>0.00</td>
<td>694.5</td>
<td>4.7</td>
<td>0.2</td>
<td>29.7</td>
<td>0.4</td>
<td>0.5</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 2 Saccharin (mg)</td>
<td>2.9</td>
<td>9.2</td>
<td>0.0</td>
<td>0.00</td>
<td>52.3</td>
<td>3.8</td>
<td>0.2</td>
<td>15.5</td>
<td>0.4</td>
<td>0.4</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 2 Sucralose (mg)</td>
<td>12.1</td>
<td>28.0</td>
<td>0.0</td>
<td>0.00</td>
<td>144.6</td>
<td>3.3</td>
<td>0.2</td>
<td>11.4</td>
<td>0.4</td>
<td>0.5</td>
<td>≤0.001</td>
</tr>
<tr>
<td>NNS-FFQ 2 Total (mg)</td>
<td>97.6</td>
<td>194.4</td>
<td>24.6</td>
<td>0.00</td>
<td>1221.0</td>
<td>3.7</td>
<td>0.2</td>
<td>15.8</td>
<td>0.4</td>
<td>0.5</td>
<td>≤0.001</td>
</tr>
<tr>
<td>Recall Erythritol (mg)</td>
<td>16.3</td>
<td>134.0</td>
<td>0.0</td>
<td>0.00</td>
<td>1333.3</td>
<td>8.9</td>
<td>0.2</td>
<td>82.0</td>
<td>0.4</td>
<td>0.1</td>
<td>≤0.001</td>
</tr>
<tr>
<td>Recall Acesulfame Potassium (mg)</td>
<td>6.8</td>
<td>16.4</td>
<td>0.0</td>
<td>0.00</td>
<td>99.3</td>
<td>3.7</td>
<td>0.2</td>
<td>14.8</td>
<td>0.4</td>
<td>0.4</td>
<td>≤0.001</td>
</tr>
<tr>
<td>Recall Aspartame (mg)</td>
<td>36.5</td>
<td>97.2</td>
<td>1.2</td>
<td>0.00</td>
<td>526.4</td>
<td>3.7</td>
<td>0.2</td>
<td>13.6</td>
<td>0.4</td>
<td>0.2</td>
<td>≤0.001</td>
</tr>
<tr>
<td>Recall Saccharin (mg)</td>
<td>3.9</td>
<td>19.6</td>
<td>0.0</td>
<td>0.00</td>
<td>170.9</td>
<td>6.5</td>
<td>0.2</td>
<td>47.5</td>
<td>0.4</td>
<td>0.5</td>
<td>≤0.001</td>
</tr>
<tr>
<td>Recall Sucralose (mg)</td>
<td>8.0</td>
<td>19.7</td>
<td>0.0</td>
<td>0.00</td>
<td>103.7</td>
<td>3.0</td>
<td>0.2</td>
<td>9.6</td>
<td>0.4</td>
<td>0.5</td>
<td>≤0.001</td>
</tr>
<tr>
<td>Total NNS Reported</td>
<td>72.3</td>
<td>183.6</td>
<td>7.7</td>
<td>0.00</td>
<td>1335.7</td>
<td>4.5</td>
<td>0.2</td>
<td>24.1</td>
<td>0.4</td>
<td>0.4</td>
<td>≤0.001</td>
</tr>
</tbody>
</table>
Discussion

This is the first investigation of its kind to develop and evaluate a valid and reliable FFQ aimed at measuring habitual NNS intake. Based on this investigation, significant correlations were found between total NNS amounts and four individual NNS measured by the NNS-FFQ when compared to amounts measured by participants’ three-day average dietary recalls. This approach of comparing NNS-FFQ amounts to NNS amounts reported in dietary recalls is in agreement with similar dietary assessment validation studies\textsuperscript{18}. While there are limitations associated with comparing one self-reported assessment method to another, comparative validity is currently the preferred method for establishing validity before an objective reference measure, such as a dietary biomarker, has been validated. Similar validation studies on FFQ generally consider correlations between 0.5 and 0.7 to be valid\textsuperscript{16}. The correlation values for the NNS captured in the NNS-FFQ were between 0.51 and 0.59, with the exception of erythritol. Furthermore, $\kappa$ values demonstrated moderate (i.e., $\kappa=0.41–0.60$) to substantial (i.e., $\kappa=0.61–0.80$) agreement for identifying consumers and non-consumers for total and all types of NNS\textsuperscript{23}.

It is important to consider the magnitude of differences found between the measurement methods. Although there were significant differences, the small mass typically consumed may not be clinically relevant. In validity analyses, significant mean differences were found for reported acesulfame potassium values (12.0 mg). This mean difference, while statistically significant, represents about 4 fl oz of diet soda, with 3 mg aspartame in 1 fl oz of diet soda. NNS vary in their sweetness intensity, making the milligram amounts found in each packet different. For instance, a single packet of each sweetener contains 40 mg of saccharin, 50 mg of acesulfame potassium, 11 mg of sucralose, or 9 mg of stevia\textsuperscript{3}. Comparing these numbers to the mean difference found in validity analyses shows that all differences were less than half a packet.
difference, indicating that the statistical differences may not impact clinical outcomes. Similarly, in test–retest reproducibility analyses, significant mean differences were found for reported sucrase values (6.2 mg); however, the mean difference in sucrase is similar to that found in 1 fl oz diet soda, which is about half of what is in one sweetener packet of sucrase.

While it is noted that reported total NNS intake was not significantly different between the two methods, analyzing individual NNS types is likely to be more valuable in future investigations. One of the major limitations of the current body of literature on NNS is the reliance on analyzing NNS as a broad category rather than as individual compounds. Given that NNS are understood to be metabolized differently, and thus have different impacts on the human body (e.g., aspartame is metabolized into its constituent parts, while others are excreted in part or in whole in the urine or feces), it is important that research studies measure individual types of NNS. As previously mentioned, the larger mean difference detected between the NNS-FFQ and dietary recalls may be less meaningful due to the wide range of milligram amounts used based on the type of sweetener.

Erythritol was analyzed to represent stevia products, as it is frequently used as a bulking agent in stevia products. In the validity analysis, erythritol values were not significantly correlated with an $r$ value (−0.03) that was lower than the typical range of 0.5 to 0.7 for validity analyses. While this correlation is lower than the other individual sweeteners, this may be attributed to the larger volume at which erythritol is consumed, paired with only a few participants reporting erythritol consumption. Erythritol is generally reported in gram amounts, due to the large mass typically consumed. For instance, 1 stevia sweetener packet contains 3 g, or 3000 mg, of erythritol. Therefore, the mean difference of 9 mg is likely not clinically significant.
The NNS-FFQ is intended to measure habitual NNS intake over the previous month, whereas dietary recalls collect recent dietary intake. The dietary recalls were collected on non-consecutive days, including two weekdays and one weekend day, which is considered to be more representative of habitual intake than one day alone. However, as dietary recalls measure dietary intake over the previous day and reflect recent intake, it would be expected that NNS consumption would not be identical to that reported in the NNS-FFQ, which represents habitual intake. While the NNS-FFQ could be manipulated to measure a shorter or longer time frame, the one-month time frame will likely have higher utility for researchers to measure health outcomes and may also be more manageable for participants to report. It is important to consider the timeframe during which the NNS-FFQ is administered and take into consideration seasonal changes that may take place in individuals’ diets. For example, people may choose to drink more soft drinks during the summer than during the winter. However, the NNS-FFQ is intended to measure intake over the past month, and for this investigation the NNS-FFQ and food recalls were completed within two weeks of each other. FFQ have been studied for their ability to measure shorter and longer time frames; therefore, future investigations should consider the ability of the NNS-FFQ to assess a variety of time frames of NNS consumption and to account for seasonal variability.

The NNS-FFQ has a number of advantages over other methods of dietary assessment, such as 24-hr dietary recalls and records. The NNS-FFQ can be used to gauge habitual NNS intake over the previous month, whereas dietary recalls only measure recent intake. The NNS-FFQ can be self-administered, making it convenient and economical to collect large amounts of data in community or epidemiological studies and could be designed to be machine-readable,
making it more easily scored for analysis, and thus reducing researcher-associated costs\textsuperscript{37}. The scale at which the NNS-FFQ can be used will make it useful for large population-level studies\textsuperscript{38}.

While there are a number of advantages to using a FFQ, there are also challenges associated with all methods of self-reported data. Due to their subjective nature, self-reported dietary data has known limitations, such as under-reporting of caloric intake\textsuperscript{15}. Currently, the “gold standard” for measuring subjective NNS intake is multiple dietary recalls or food records, which are then analyzed with dietary analysis software, such as NDS-R\textsuperscript{39}. Dietary assessment for all foods has known challenges, which are compounded in the case of NNS. For example, the FDA does not require specific amounts of NNS used in food products to be reported, making it difficult to quantify NNS intake, even with validated dietary analysis software. Furthermore, if consumers are unable to identify the products they are consuming as artificially sweetened, dietary recalls may fail to capture NNS in consumers’ diets. To help address this limitation, graduate-level research assistants were trained to administer 24-h dietary recalls and enter data using state-of-the-art dietary analysis software. Data from three non-consecutive days including two weekdays and one weekend day was collected. Additionally, to establish rapport and improve data integrity, the same research assistant completed all three dietary recalls and completed the data entry and analysis. Previous investigations have determined the tendency of FFQ to overestimate dietary intake\textsuperscript{40}. This analysis also found that the NNS-FFQ was more likely to overestimate NNS intake compared to dietary recalls, with all NNS types being higher based on the NNS-FFQ with the exception of saccharin. This information should be taken into account when using the NNS-FFQ and interpreting results.

The continually expanding number of products sweetened with NNS is a concern. The FFQ includes an “other” category at the end of the questionnaire for participants to list products
not found in the questionnaire. These tended to be generic versions of items already included in the questionnaire and more recently developed products using stevia in this investigation. Twenty participants included information in the “other” category, among which 9 products were matched to related products already included in the FFQ (i.e., generic brands). Eleven participants included items that could not be matched. As previously stated, this analysis used erythritol as a representation for stevia products. When the information was available, NNS quantity was determined using NDS-R or matched the product with similar products using the same combination of sweeteners. In order to accommodate this challenge, this tool will need to be updated periodically to identify the most frequently consumed NNS products.

While the diversity of this study group is considered a strength as it allows for generalizability across a wide population, it did warrant additional analyses to determine if there were differences in reproducibility and validity results based on demographic characteristics. While there was a significant difference in total NNS consumption between education groups reported in the dietary recalls, no significant difference in validity or reproducibility was found between education groups. Moreover, as there were only seven participants who were 65 years or older, this tool may have limitations in an older adult population; however, the analyses suggest that the NNS-FFQ is valid and reproducible for this age group. Finally, due to a reported difference in total NNS intake between the NNS-FFQ and dietary recalls when analyzing only white participants, future research should thoroughly examine the potential impact of race/ethnicity on the validity of this tool.

Finally, this tool was developed based on NNS consumption patterns in the US. Therefore, the NNS-FFQ may have limited generalizability outside the US. Future investigations should consider additional versions of the NNS-FFQ that reflect the types of NNS commonly
consumed outside the US, as well as commonly consumed brands of NNS food and beverage products.

NNS consumption has been surrounded by controversy in part due to researchers’ inability to measure NNS intake. Researchers are frequently limited to studying NNS as a whole category or using diet soda as a proxy for NNS consumption, rather than as individual compounds, due to the lack of valid assessment tools able to distinguish between NNS types. Developing methods to measure intake of each NNS type will allow researchers to measure how much of the population consumes NNS and to examine inferentially the potential health outcomes of NNS consumption.

**Conclusion**

This investigation determined that the NNS-FFQ is a reproducible and valid dietary assessment tool that is able to gauge habitual NNS intake patterns relative to 24-h dietary recalls, with the possible exception of erythritol consumption. The NNS-FFQ is a rapid self-report questionnaire that could allow researchers to measure NNS consumption at the individual and population levels among diverse populations.
References


Chapter 3 – Identifying an Appropriate Carrier for Non-nutritive Sweeteners in Metabolic and Controlled Feeding Investigations via Sensory Evaluation

Abstract

**Introduction:** To strengthen metabolic study designs on non-nutritive sweeteners (NNS), sensory characteristics of NNS combined with different carriers should be better understood. The research objective was to determine an appropriate carrier (water or applesauce) for a high-dose of specific NNS types (aspartame, sucralose, or stevia) to inform future metabolic and controlled feeding study designs.

**Methods:** Adult participants (n=67) sampled six sweetener-carrier combinations (water and applesauce containing high concentrations of aspartame [30g/oz], sucralose [8.25g/oz], and stevia [6.75g/oz]). Participants completed Check-All-That-Apply emotional terminology questionnaires, affective attribute questionnaires, and paired preference questionnaires.

**Results:** Applesauce was preferred (sucralose=83.6%; aspartame=79.1%; stevia=74.6%) significantly more than water for all sweetener types (p<0.001) and mean acceptability scores were significantly higher for all applesauce samples.

**Conclusion:** Participants preferred NNS delivered in applesauce rather than water. Applesauce is likely a more appropriate and tolerable carrier for high-dose NNS, which may contribute to designing effective intervention studies with greater participant compliance.
Introduction

Non-nutritive sweeteners (NNS), also known as artificial sweeteners, are frequently recommended by healthcare professionals for their potential to reduce caloric intake and induce weight loss\textsuperscript{1,2}, as well as their ability to assist individuals with diabetes in managing blood glucose levels\textsuperscript{3,4}. However, there are many limitations to NNS study designs for metabolic and controlled feeding investigations that make it difficult to draw definitive conclusions regarding the impact of NNS on health outcomes. Many studies on NNS rely on observational data\textsuperscript{5-7} rather than randomized controlled trials (RCT)\textsuperscript{1,2}, making it challenging to determine causal relationships between NNS and health outcomes.

An additional challenge among NNS studies is the lack of consistency in how NNS consumption is quantified or delivered to participants. Many observational studies use diet soda as a proxy for NNS consumption\textsuperscript{5,6}, and RCT use a variety of carriers for NNS, including diet soda, beverages, and foods\textsuperscript{1,2}. Few controlled feeding studies have been conducted on NNS and these few used a combination of beverages and foods sweetened with NNS\textsuperscript{8,9}. These differences make it difficult to compare results across studies that utilize different carriers for delivering NNS.

There are numerous challenges in conducting RCT that measure the metabolic impacts of NNS. For example, there is a need to deliver high doses of NNS in metabolic studies; however, a high dose can increase potential of NNS-associated off-flavors, create solubility challenges, and flavors between carriers and sweeteners may be incompatible. Many sweeteners have bitter or unusual flavors at high doses, which may decrease compliance in consuming the entire food or beverage amount and may increase participant attrition rates. Compliance is essential in metabolic and controlled feeding studies; thus, delivering the dose without creating excess
beverage or food consumption requirements and with a tolerable, and preferably pleasant, consumption experience is critical. A well-planned RCT considers the participants’ ability to consume the carrier with the dosed ingredient at the needed level within the targeted timeframe while yielding a positive sensory experience.

In order to design effective interventions that can be disseminated and implemented, the amount and type of NNS considered acceptable in metabolic and controlled feeding studies should be determined. Knowledge of the sensory characteristics of NNS within different types of carriers will allow researchers to choose the appropriate carrier for these substances in RCT and improve participant compliance. Sensory evaluation allows researchers and food manufacturers to measure and categorize responses to food and beverage products in a blinded setting that limits biases. Affective tests allow researchers to assess consumers’ preference or acceptance of a product based on a variety of characteristics, such as sweetness or bitterness\(^\text{10}\). Check-All-That-Apply (CATA) emotional term questionnaires provide researchers with a method to describe and categorize products with a sensitive measure\(^\text{11}\). Paired preference tests can be used to determine consumers’ preference for one item versus another\(^\text{10}\).

Previous investigations have examined the taste interactions between various stimuli in combination with NNS. Steiner (1979) determined that water alone (no added ingredients) elicits neutral sensory responses\(^\text{12}\); this suggests water is a suitable carrier beverage for dietary interventions. Wang et al. (2016) determined the influence of iron and other minerals on sensory perceptions of NNS in water, finding that water with higher mineral content can intensify perceptions of sweetness\(^\text{13}\). Leitch et al. (2015) found that NNS were considered to be acceptable as an alternative to sucrose when used in still beverages (tea)\(^\text{14}\). However, these studies utilized NNS dosages that would not be sufficient to study the acute metabolic effects of NNS.
The ideal carrier for an RCT would deliver the targeted compound (i.e., NNS) without creating another metabolic effect (e.g., an insulin or satiety response) that interferes with or confounds study outcomes. Beverages are reported to be less satiating than foods\textsuperscript{15}, potentially inferring that differences in sensory characteristics are likely to be reported by consumers. Water may be more easily compared with other beverages while eliminating other taste interactions as well as interactions from other nutrient components, such as carbohydrates; however, it may not be a preferable or tolerable carrier for participants to consume. Applesauce has an intrinsic sweetness to it that may make it a preferable carrier for NNS; however, because of the carbohydrate content, this may add additional challenges in interpreting outcomes.

The objective of this investigation was to determine an appropriate carrier (water or applesauce) for a high dose of specific non-nutritive sweeteners (aspartame, sucralose, or stevia) in order to inform future RCT research study designs. It was hypothesized that, when a controlled dose was administered, participants would prefer applesauce as a carrier compared to water. It was expected that both carriers of all NNS types (at high dose) would receive low affective attribute and preference scores as well as fewer positive emotional terms based on the relatively high NNS content of the samples. Finally, it was expected that participants who reported habitual NNS consumption would report higher affective attribute scores.

Methods

Subjects

Eligible participants were adults 18 years or older. Participants were excluded if they had phenylketonuria (PKU) due to the use of aspartame as part of the intervention, were pregnant or breastfeeding, or had an intolerance of NNS. Participants were recruited from the local
community (Blacksburg, Virginia). Seventy participants were enrolled in the study. This study was reviewed by the Western Institutional Review Board (WIRB) and was classified as exempt under 45 CFR §46.101(b)(6) and 21 CFR 56.104(d) (Virginia Tech IRB #18-063, WIRB #1-1074443-1). Participants reviewed and signed an informed consent document prior to receiving any samples.

Sample Preparation and Delivery

Pure sucralose, without bulking agents, was received from Sucral (Tate & Lyle; London, United Kingdom). Pure aspartame, without bulking agents, was received from Wego Chemical & Mineral Co. (Great Neck, NY). Truvia brand stevia packets, with erythritol, (Cargill; Wayzata, MN) were purchased at the local supermarket. Unsweetened applesauce, distilled water, drinking water, and unsalted soda crackers were purchased from Kroger (Kroger brand; Cincinnati, OH).

The NNS dosage was administered by adding one of the three types of NNS to unsweetened applesauce and distilled water at an established concentration level for each sweetener. Across each type of sweetener, the dosage was representative of the sweetness present in 3 sweetener packets per 4 ounces of the carrier sample (30 mg per ounce for aspartame, 8.25 mg per ounce for sucralose, and 6.75 mg (or 0.75 Truvia packet) per ounce for stevia). These weights indicate the pure form of the sweeteners without bulking agents. While aspartame and sucralose were used in the pure form, Truvia packets contain erythritol as a sweetening and bulking agent. The Truvia packets were chosen because dietary assessment software (Nutrition Data System for Research [NDS-R] 2015) does not measure stevia but does measure erythritol and contains the Truvia brand packets. However, these packets were used with the assumption that each packet contained 9 mg of stevia.
When preparing the applesauce samples, the sweeteners were dissolved into 200 ml (6.8 fluid ounces) of distilled water and then combined with 57 ounces of applesauce to ensure complete dispersion of the sweetener. This procedure did not have a noticeable impact on the texture or consistency of the applesauce. Water samples were prepared by combining the NNS directly into distilled water and mixing until dissolved and there was no visible precipitation. The NNS quantities were chosen to reflect amounts that participants could plausibly consume in a typical day (i.e., three sweetener packets) and concentration levels that could possibly be used in samples in research studies with the goal of consuming the dose in one serving. These amounts were also within the Food and Drug Administration’s acceptable daily intake\(^\text{16}\).

**Study Design**

Enrolled participants were seated in individual booths in the Virginia Tech Sensory Evaluation Laboratory. Participants were given three trays of NNS samples, with one pair of samples on each tray. Each tray included one applesauce sample and one water sample, containing the same sweetener type, unsalted soda crackers, a cup of drinking water, and an empty cup for expectoration. Participants were advised to cleanse their palates by rinsing their mouths and taking a small bite of the soda cracker between samples. One-ounce samples of each carrier were served at room temperature in two-ounce opaque plastic cups; each sample was labeled with a randomized 3-digit code. Participants were informed that they could swallow or expectorate the samples after each tasting. The order of sweeteners that each participant received was planned in a balanced order, where the tray order of aspartame, sucralose, and stevia was randomized such that each sweetener was presented in each position (first, second, third) an equal number of times across all participants. Carrier presentation was also planned in a balanced
order such that an equal number of participants tasted water samples first and an equal number of
participants taste applesauce samples first.

For each of the three pairs of samples, participants completed a CATA emotional term
questionnaire for each sample\textsuperscript{11,13,14}, a five-point affective attribute questionnaire to determine
overall acceptability for each sample (1 = dislike extremely, 5 = like extremely), as well as
attribute (sweetness, bitterness, metallic) acceptability and intensity liking. Finally, a paired
preference questionnaire for a comparison of the water and applesauce samples was completed
(Figure 1). A five-point scale was used to reduce participant burden and fatigue over the course
of six samples; investigators recognized that participants were not trained and might experience
decision fatigue over a broader number of categories for the number of research questions and
samples. Between the first and second pairs of samples, participants were provided a rest while
they completed a brief demographic questionnaire (sex, age, race). Between the second and third
pairs of samples, participants completed a NNS food frequency questionnaire (NNS-FFQ) to
assess habitual NNS consumption patterns\textsuperscript{17}. Data from the NNS-FFQ was used to quantify
habitual sucralose, aspartame, saccharin, acesulfame potassium, and erythritol intake over the
previous month. Erythritol intake was used as a proxy for stevia intake due to the lack of data
available on stevia in NDS-R.
Figure 1. Sensory evaluation questionnaires utilized in this investigation

![Emotional Term Ballot](image)

Product: ____________________________

Select all the word(s) from the list below that describe how you feel right now about the product you have just sampled. Check ALL that apply.

- Active
- Enthusiastic
- Mild
- Tame
- Adventurous
- Fear
- Nostalgic
- Tender
- Affectionate
- Free
- Peaceful
- Understanding
- Aggressive
- Friendly
- Please
- Warm
- Angry
- Good
- Pleasant
- Whole
- Bored
- Good-natured
- Polite
- Mild
- Calm
- Guilty
- Quiet
- Worried
- Content
- Happy
- Sad
- Daring
- Interested
- Safe
- Disgusted
- Joyful
- Satisfied
- Disgust
- Loving
- Secure
- Energetic
- Merry
- Steady

---

![Affective Testing Score Card](image)

Product: ____________________________

- Please rinse your mouth before starting.
- Considering all characteristics, please indicate your overall opinion by checking one box.

![Dislike extremely](image)

- ![Dislike](image)
- ![Neither like nor dislike](image)
- ![Like](image)
- ![Like extremely](image)

Comments: Please indicate WHAT in particular you liked or disliked about this product.

- Liked ____________________________
- Disliked ____________________________

Retaste the product as needed and check the box for your response for both questions (LIKING and INTENSITY).

![Sweetness](image)

- ![Dislike extremely](image)
- ![Dislike](image)
- ![Neither like nor dislike](image)
- ![Like](image)
- ![Like extremely](image)

![Bitterness](image)

- ![Dislike extremely](image)
- ![Dislike](image)
- ![Neither like nor dislike](image)
- ![Like](image)
- ![Like extremely](image)

![Metallic](image)

- ![Dislike extremely](image)
- ![Dislike](image)
- ![Neither like nor dislike](image)
- ![Like](image)
- ![Like extremely](image)

---

![Preference Score Card](image)

Instructions:

1. Taste the product on the left first, and the product on the right second. Now that you’ve tasted both products, which one do you prefer? Please choose one:

- ![□](image)
- ![□](image)

2. Please comment on the reasons for your choice:

---

69
Statistical Analyses

Descriptive statistics were used to analyze demographics, CATA emotional terms, and preference data (mean ± standard deviation and frequencies). Preference frequencies were totaled and Roessler’s significance tables were used to determine if significant differences were present\(^1\). Mean acceptance scores for each NNS type (aspartame, sucralose, and stevia) and each carrier (water and applesauce) were calculated and paired samples t-tests were evaluated to determine the differences between each acceptance and attribute intensity score for each carrier. One-way analysis of variance (ANOVA) was conducted to measure mean differences between NNS consumers and non-consumers. Statistical analyses were conducted using SPSS statistical analysis software (v. 25). An *a priori* significance level was set at p≤0.05.

Results

Participants

Among the 70 participants who enrolled in the study, 67 participants completed the investigation. The sample was predominately Caucasian (73%) females (70%) with age ranging from 18 to 62 years old with a mean age of 26.2±8.9 years and a median age of 22.0 years.

Carrier Preference Rating

Applesauce was significantly preferred as a carrier over water for all sweetener types (p≤0.001; Table 1), with at least 75% of participants preferring the applesauce carrier for each sweetener type. No differences in preferences were detected based on sex or ethnicity.
Table 1. Carrier preference for each non-nutritive sweetener type (n=67)

<table>
<thead>
<tr>
<th>Non-nutritive Sweetener Type</th>
<th>Water Preference n (%)</th>
<th>Applesauce Preference n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sucralose</td>
<td>11 (16.4%)</td>
<td>56 (83.6%)***</td>
</tr>
<tr>
<td>Aspartame</td>
<td>14 (20.9%)</td>
<td>53 (79.1%)***</td>
</tr>
<tr>
<td>Stevia</td>
<td>17 (25.4%)</td>
<td>50 (74.6%)***</td>
</tr>
</tbody>
</table>

***p≤0.001; Note: Significance established based on Roessler et al., significance tables.

Carrier and Intensity Relationship to Acceptability

Mean acceptability scores were calculated for each of the six sweetener-carrier combinations to determine the liking (overall and for each attribute) and intensity. This was interpreted as the interaction of the NNS within the carrier matrix, with inherent product characteristics. In addition to measuring liking, it was important to measure the acceptability of the intensity of the attribute since the NNS dose was at a high level and this knowledge could help identify challenges that might influence compliance. Mean acceptability scores for the two samples of each sweetener-carrier combination (e.g., water with sucralose and applesauce with sucralose) were compared (Table 2). For all NNS types, mean overall liking scores were higher for applesauce samples compared to water samples. When comparing the sucralose-loaded applesauce and water, significant mean differences favoring applesauce samples were detected for attribute liking, with the exception of metallic liking, and intensity acceptability. For aspartame samples, significant mean differences favoring applesauce samples were detected for overall affective score, sweetness liking, and sweetness intensity. For stevia samples, significant mean differences favoring applesauce samples were detected for all affective attributes with the exception of metallic liking and metallic intensity. Metallic liking scores for applesauce were
slightly, but non-significantly, lower than for water and metallic intensity scores for applesauce were significantly lower than for water, which was not observed among other sweetener samples.

**Table 2. Mean hedonic scores for sucralose (a), aspartame (b), and stevia (c) with water or applesauce (n=67)**

<table>
<thead>
<tr>
<th>Sweetener Type</th>
<th>Water with Sweetener Mean (SD)</th>
<th>Applesauce with Sweetener Mean (SD)</th>
<th>Mean Difference Mean (SE)a</th>
<th>Correlation (r)b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Liking Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Sucralose</td>
<td>Overall Liking 2.8 (1.0)</td>
<td>3.6 (0.9)</td>
<td>0.8 (0.2)***</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Sweetness 2.9 (1.2)</td>
<td>3.7 (1.0)</td>
<td>0.8 (0.2)***</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Metallic 2.9 (0.7)</td>
<td>3.1 (0.7)</td>
<td>0.2 (0.1)</td>
<td>0.29*</td>
</tr>
<tr>
<td></td>
<td>Bitterness 3.1 (0.7)</td>
<td>3.3 (0.8)</td>
<td>0.2 (0.1)*</td>
<td>0.39**</td>
</tr>
<tr>
<td></td>
<td>Intensity Acceptability Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sweetness - Intensity 2.7 (1.1)</td>
<td>3.4 (1.1)</td>
<td>0.7 (0.2)***</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Metallic - Intensity 2.9 (0.7)</td>
<td>3.1 (0.7)</td>
<td>0.2 (0.1)*</td>
<td>0.49***</td>
</tr>
<tr>
<td></td>
<td>Bitterness - Intensity 3.1 (0.7)</td>
<td>3.3 (0.7)</td>
<td>0.2 (0.1)*</td>
<td>0.33**</td>
</tr>
<tr>
<td>b) Aspartame</td>
<td>Overall Liking 2.6 (1.1)</td>
<td>3.4 (1.2)</td>
<td>0.8 (0.2)***</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Sweetness 2.7 (1.2)</td>
<td>3.2 (1.2)</td>
<td>0.5 (0.2)**</td>
<td>0.41**</td>
</tr>
<tr>
<td></td>
<td>Metallic 2.9 (0.5)</td>
<td>2.9 (0.7)</td>
<td>0.0 (0.1)</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Bitterness 3.1 (0.5)</td>
<td>3.1 (0.7)</td>
<td>0.0 (0.1)</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>Intensity Acceptability Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sweetness - Intensity 2.6 (1.2)</td>
<td>3.1 (1.1)</td>
<td>0.8 (0.2)***</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Metallic - Intensity 2.9 (0.5)</td>
<td>3.0 (0.8)</td>
<td>0.5 (0.2)**</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Bitterness - Intensity 3.1 (0.5)</td>
<td>3.2 (0.7)</td>
<td>0.0 (0.1)</td>
<td>0.20</td>
</tr>
<tr>
<td>c) Stevia</td>
<td>Overall Liking 2.5 (1.3)</td>
<td>2.9 (1.2)</td>
<td>0.4 (0.2)***</td>
<td>0.31*</td>
</tr>
<tr>
<td></td>
<td>Sweetness 2.7 (1.3)</td>
<td>3.1 (1.3)</td>
<td>0.4 (0.2)***</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Metallic 2.9 (0.9)</td>
<td>2.8 (1.0)</td>
<td>0.1 (0.1)</td>
<td>0.61***</td>
</tr>
<tr>
<td></td>
<td>Bitterness 2.7 (1.0)</td>
<td>3.0 (1.0)</td>
<td>0.3 (0.1)*</td>
<td>0.47***</td>
</tr>
<tr>
<td></td>
<td>Intensity Acceptability Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sweetness - Intensity 2.4 (1.2)</td>
<td>3.0 (1.2)</td>
<td>0.6 (0.2)***</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Metallic - Intensity 2.9 (0.9)</td>
<td>2.7 (1.0)</td>
<td>0.2 (0.1)*</td>
<td>0.62***</td>
</tr>
<tr>
<td></td>
<td>Bitterness - Intensity 2.8 (1.0)</td>
<td>3.0 (1.0)</td>
<td>0.2 (0.1)*</td>
<td>0.47***</td>
</tr>
</tbody>
</table>

Note: Affective scores were evaluated by 5-point scale; 1 = extremely dislike, 2 = dislike, 3 = neither like nor dislike, 4 = like, 5 = extremely like; attribute (sweetness, bitterness, metallic) acceptability and intensity were measured for each sweetener-carrier combination;

a * Indicates that means within a row are significantly different;
b * Indicates a significant correlation between carriers for a specific NNS type;
*p≤0.05, **p≤0.01, ***p≤0.001
Emotional Term Analysis

Frequencies for each emotional term (43 total terms) were totaled for each carrier type (water or applesauce) and sweetener (sucralose, aspartame, or stevia). Thirteen terms were selected with 20% frequency or more for at least one sweetener-carrier sample and were classified as “frequently selected”. With the exception of the term “disgusted”, all other frequently selected terms were positive, such as “happy”, or neutral, such as “mild”. These terms were graphed in radar plots to compare for overall similarities and differences between carriers by sweetener (Figure 2).
Figure 2. Distribution of frequently selected emotional terms among panelists (n=67)

Note: The displayed terms were selected by >20% of participants for at least one sample. Terms denoted with one asterisk were shared terms (≤8% difference among the two carrier groups within a given NNS type). Terms denoted with two were unique terms (≥12% difference among the two carrier groups within a given NNS type).

**NNS Consumer vs. Non-Consumer Analysis**

Based on the NNS-FFQ, participants reported a mean NNS intake of 81.6±41.7 mg per day, excluding erythritol. Participants were classified as NNS consumers or non-consumers according to the data provided in the NNS-FFQ. Using a published novel method\(^9\), participants who reported daily consumption of the NNS equivalent of 1 fluid ounce of diet soda from all foods and beverages were categorized as NNS consumers. This intake level corresponds to 3 mg...
acesulfame potassium, 17 mg aspartame, 12 mg saccharin, or 6 mg sucralose per day. Erythritol was not included in this analysis due to the fact that it is not used in stevia-sweetened soda.

Three participants were excluded from the NNS-FFQ analysis due to incomplete responses on the questionnaire. Among the participants who completed the NNS-FFQ (n=64), 42 (65.6%) were identified as NNS consumers. One-way analysis of variance (ANOVA) was conducted to measure mean differences between NNS consumers and non-consumers (Table 3). No significant differences between groups were detected for any of the sweetener-carrier combinations based on participants’ habitual NNS consumption.

Table 3. Mean overall affective scores among NNS consumers and non-consumers (n=64)

<table>
<thead>
<tr>
<th>Sweetener-Carrier</th>
<th>NNS Consumers (n=42) Mean (SD)</th>
<th>NNS Non-Consumers (n=22) Mean (SD)</th>
<th>One-Way ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sucralose-Water</td>
<td>2.9 (1.0)</td>
<td>2.7 (1.1)</td>
<td>F=0.9; p=0.36</td>
</tr>
<tr>
<td>Sucralose-Applesauce</td>
<td>3.8 (0.8)</td>
<td>3.4 (1.1)</td>
<td>F=3.1; p=0.09</td>
</tr>
<tr>
<td>Aspartame-Water</td>
<td>2.8 (1.0)</td>
<td>2.4 (1.1)</td>
<td>F=2.0; p=0.17</td>
</tr>
<tr>
<td>Aspartame-Applesauce</td>
<td>3.6 (1.1)</td>
<td>3.0 (1.3)</td>
<td>F=3.4; p=0.07</td>
</tr>
<tr>
<td>Stevia-Water</td>
<td>2.5 (1.2)</td>
<td>2.5 (1.5)</td>
<td>F=0.0; p=0.99</td>
</tr>
<tr>
<td>Stevia-Applesauce</td>
<td>3.1 (1.2)</td>
<td>2.5 (1.3)</td>
<td>F=3.0; p=0.09</td>
</tr>
</tbody>
</table>

Note: Affective scores were evaluated by 5-point scale; 1 = extremely dislike, 2 = dislike, 3 = neither like nor dislike, 4 = like, 5 = extremely like; attribute (sweetness, bitterness, metallic) acceptability and intensity were measured for each sweetener-carrier combination.

Discussion

This investigation provides researchers with an improved understanding of how to identify an appropriate carrier for delivering NNS to participants in metabolic and controlled feeding studies aiming to assess the health effects of NNS. Participants’ overwhelming
preference for applesauce samples rather than water samples aligned with the hypothesis. While water has neutral sensory properties that may be helpful for understanding the flavor profile of the sweeteners themselves\textsuperscript{12}, it is not likely to be well-tolerated as a carrier for a high dosage of NNS for research studies. Qualities of the applesauce, such as the texture and acidity, may have made it more tolerable for participants to consume\textsuperscript{20}, although this may not be as effective with stevia samples and the nutrient content of applesauce may conflict with other aspects of metabolic study designs.

Participants’ overall preference for either water or applesauce as the NNS carrier was primarily influenced by their perceptions of sweetness, rather than bitterness and metallic flavors. Among all sweetener types, significant mean differences favoring applesauce samples in hedonic scores for sweetness were detected, indicating that this quality was most likely to align with their overall preference for carrier type. Additionally, bitterness and metallic attributes between carrier types were significantly correlated for sucralose and stevia and t-tests revealed smaller mean differences, suggesting that these qualities were perceived similarly between carrier types and were less impactful on overall liking. Previous investigations have found that bitterness was the most impactful in predicting sweetener preference\textsuperscript{13,21}; however, this investigation found that sweetness was more correlated with overall liking. Moreover, the emotional term questionnaire revealed that applesauce samples were more likely to illicit positive terms, such “happy”, “good”, and “pleased”, while the term “disgusted” was a frequently selected term for all three water samples and there were fewer frequently selected positive terms for water.

Researchers should aim to use a sweetener-carrier combination with at least a neutral score, in this case a 3.0 (neutral/neither like nor dislike), or greater. Scores lower than 3.0 may
indicate that participants are less likely to be compliant with study procedures. In this investigation, all overall liking scores for water samples as well as stevia-applesauce samples were below a 3.0, suggesting that these samples would not be well-suited for metabolic research studies.

To our knowledge, this is the first sensory evaluation on NNS to measure both NNS preferences and habitual intake (via NNS-FFQ). It is important to note, that while no significant acceptability differences for any of the sweetener-carrier combinations were found between NNS consumers and non-consumers, all values were greater or equal for all sweetener-carrier combinations for NNS consumers. This is valuable insight suggesting that habitual NNS consumers would be more likely to rate NNS products more favorably than those who do not habitually consume NNS. These slight, but statistically insignificant, differences in acceptability would be important to recognize when designing a NNS intervention.

As previously mentioned, beverages and foods have different satiety effects when consumed, with foods producing a greater satiety response than beverages. This was observed in this investigation, where the emotional terms associated with satiety were more frequently selected for applesauce samples compared to water samples. The term “satisfied” was selected by 25% of participants for applesauce-aspartame samples, but only 5% of water-aspartame samples. The term “content” was selected by 42% of applesauce-sucralose samples and only 18% of water-sucralose samples. NNS have been studied for their impact on satiety levels, with some demonstrating that NNS consumption is associated with decreased satiety while others found that NNS did not impact satiety levels. However, it is yet to be determined how qualities of NNS influence how the carriers themselves are perceived.
Another aspect of NNS that should be explored are reports of side-effects, such as migraines\textsuperscript{25-27} and gastrointestinal symptoms. While these symptoms could not be measured in this sensory panel study, it would be important to know if consumers experience these side-effects when receiving a high-dose of sweetener within a delivery vehicle. Understanding these side effects will allow researchers to better select sweetener-carrier combinations for researcher studies involving NNS intake as an intervention.

\textit{Limitations}

This study asked participants to taste each sample and allowed participants to expectorate after each sample. Investigators estimate that participants consumed half of each one-ounce sample. However, in future research settings, participants would be asked to consume an entire sample, such as four ounces or more at a time. Therefore, this study should be considered a preliminary investigation in the acceptability of various carriers for NNS intervention research. Future investigations should consider using larger samples of sweetened products and ask that participants consume the entire sample.

Another limitation is the lack of understanding of why participants felt the way they did about each sweetener-carrier combination. Qualitative analysis may allow researchers to better understand the reasons for participants’ strong preference for applesauce. Future analysis should include qualitative analysis of participants’ comments.

\textit{Conclusions}

This investigation allows researchers to understand which carrier is more tolerable for participants to consume high doses of NNS and to design future research studies to determine
health outcomes associated with NNS consumption. While water alone has neutral sensory qualities, participants in this investigation preferred NNS combined with applesauce rather than water. Applesauce is likely a more appropriate and tolerable carrier for NNS in research studies. This information will allow researchers to strengthen their future study designs and improve participant compliance in a given intervention. An acceptable method for delivering NNS to research participants will be valuable for designing intervention studies aimed at examining how NNS consumption impacts health outcomes.
References


Chapter 4 – Non-nutritive Sweetener Urinary Biomarker Methodology

Abstract

Introduction: Non-nutritive sweeteners (NNS) are low-calorie sweeteners frequently recommended as a substitute for added sugars to reduce caloric intake and manage blood glucose levels; however, the lack of objective NNS dietary assessment methods limits researchers’ understanding of the impact of these recommendations on health. Developing an objective dietary biomarker to measure NNS intake will allow researchers and clinicians to inferentially test the impact of NNS on health and better design research studies on NNS-related health outcomes.

Objective: The objective of this preliminary analysis was to develop the methodology for a NNS urinary biomarker with the ability to detect the presence or absence of specific types of NNS in 24-hour urine samples from free-living adults.

Methods: Adult participants completed 24-hour urine samples while consuming an ad libitum diet. 24-hour urine samples were assessed for NNS presence using ultra performance liquid chromatography (UPLC-MS/MS). This rapid method analyzes each sample in 5.5 minutes, measuring the presence of NNS and related metabolites (saccharin, acesulfame potassium, sucralose, steviol glucuronide, and erythritol). The method includes using an isotopically labeled internal standard for quality control.

Results: Among 9 adults, (mean age=35.4±15.0 years, range: 21-61), the UPLC-MS/MS method observed broad differences, ranging from very strong presence of NNS to not detectable.

Conclusion: This preliminary investigation determined that the NNS urinary biomarker is able to detect the presence of NNS among free-living adult populations. Future investigations should establish a lower limit of detection to measure quantitative amounts of NNS in urine samples.
compared to recalls. The ability to objectively measure NNS intake will help researchers and clinicians evaluate the accuracy of self-reported assessment methods.
Introduction

Non-nutritive sweetener (NNS) use is surrounded by controversy and average consumers have difficulty making decisions on whether or not to consume them. NNS can be found in a wide variety of beverages and food products, including light yogurts, sugar-free coffee creamers, and sugar-free desserts. The number of products that utilize NNS in their products has increased rapidly, with over 6,000 new products released between 1999 and 2004. Recent analysis determined that 4% of food and beverage products in the United States utilize NNS.

The impact of NNS on health outcomes is controversial for many reasons. Early studies on saccharin intake in lab animals indicated there may be a risk for certain types of cancer; though, these results were never seen in humans. NNS have also been studied for their impact on both weight loss and weight gain; however, most studies pointing to weight gain are observational studies, while intervention studies tend to show weight loss. NNS have also been studied for their impact on compensatory appetite, cardiovascular disease, diabetes, and migraines. More recently, NNS are being investigated for their impact on gut microbiota. The use of animal models and the reliance on self-reported, observational data to examine health outcomes of NNS pose a problem for understanding the true health impact of these additives.

Despite their widespread adoption by the food industry, the average consumer may be unaware of how many products contain NNS, and thus, unable to accurately report their consumption using traditional subjective, self-report methods of dietary assessment, such as 24-hour dietary recalls or food frequency questionnaires. NNS are included in the ingredients list on the nutrition facts label, which may be difficult for the average consumer to interpret, especially for food items with a large number of ingredients. Additionally, there are a variety of
terms used by the food industry to label products with NNS, such as “diet”, “light”, “sugar-free”, “reduced-sugar”, and “no sugar added,” contributing to the confusion for consumers. These factors all add to the difficulty consumers may have in determining and reporting how much NNS they are consuming.

The currently available research studies on NNS often measure diet soda intake as a proxy for NNS consumption, which excludes a variety of commonly consumed NNS food and beverage products. Further illustrating this point is the large discrepancy in reported NNS consumptions rates in the literature, with reported consumption rates ranging from 15% to 50%. Thus, further investigation is needed to provide objective dietary assessment methods to overcome these reporting discrepancies. Determining true consumption rates of NNS is an important step for designing and conducting robust studies aimed at improving understanding NNS use and associated health outcomes.

Currently, the “gold standard” for measuring subjective NNS intake is multiple dietary recalls or food records, which are then analyzed with dietary analysis software (e.g., Nutrition Data System for Research software [NDS-R]). Dietary assessment for all foods has known challenges, which are compounded in the case of NNS. The FDA does not require specific amounts of NNS used in food products to be reported, but rather requires ingredients to be listed in order of predominance within the product, making it difficult to quantify NNS intake, even with validated dietary analysis software. With new products being developed using different combinations and amounts of NNS, it is difficult for software to keep up with the new products being released. Furthermore, if consumers are unable to identify the products they are consuming as artificially sweetened, dietary recalls may fail to capture NNS in consumers’ diets.
Though there are many aspects of NNS consumption and related health outcomes to be investigated, an accurate and effective method for measuring NNS intake must be developed in order to improve research study design. Without evidence showing that NNS consumption can be accurately measured, it is difficult to assess their true impact on the health of consumers. Developing methods to measure NNS intake will improve researchers’ understanding of how much of the population consumes NNS and allow investigators to inferentially examine potential health outcomes of NNS consumption.

Biomarkers can be used to accurately assess dietary intake of specific substances without the limitations of self-reported data\(^36\). There are many benefits to incorporating biomarkers into dietary assessment and broader nutrition studies. As mentioned, biomarkers are useful for validating other dietary assessment methods, such as food frequency questionnaires and 24-hr dietary recalls. Their objective nature minimizes the errors associated with self-reported measures\(^37\). Depending on the tissue specimen required (urine, blood, fingernails, etc.), biomarkers may be considered less of a burden to participants than self-reported methods\(^38\). While the number of biomarkers available is increasing, current use is limited until new methods are developed\(^39\).

Biomarkers are substances or indicators that can be used to objectively measure dietary intake\(^40\). These markers of intake or nutritional status may also be used to determine the biological and health outcomes of dietary intake\(^36\). At this time, there is a need for the further development of more nutrition biomarkers. The National Academy of Medicine (formally the Institute of Medicine) has encouraged the development and expansion of nutritional biomarkers for the purpose of improving dietary assessment methods, including validating other dietary assessment methods and measuring dietary intake\(^39\). Because objective measures are not
influenced by opinion or feelings, biomarkers allow for greater validity and neutrality in dietary assessment\textsuperscript{39}.

Given that many NNS are expected to be excreted unchanged, either in whole or in part, in the urine\textsuperscript{41-43}, a 24-hour urinary biomarker could serve as a useful method for determining quantities of NNS intake, including specific NNS types. A NNS biomarker would be useful in research and clinical settings, allowing researchers to design intervention studies aimed at determining both the long-term and short-term effects of NNS use. Developing a NNS biomarker has potential to improve the quality of research studies conducted on NNS and related health outcomes. This preliminary analysis seeks to develop the methodology of a NNS biomarker using ultra performance liquid chromatography (UPLC) mass spectrometry.

\textbf{Methods}

\textit{Study Design}

Adult participants (n=125) were recruited to participate in this observational cross-sectional study. Eligible participants were English-speaking adults ages 18 years or older. Participants were recruited through traditional methods, including flyers and listservs. This study was approved by the Virginia Tech Institutional Review Board (IRB #15-682, approved September 25, 2015). Participants provided written informed consent before enrollment. Participants completed three visits over the course of two weeks (\textbf{Figure 1}). Measures included the collection of three 24-hour dietary recalls and the completion of two 24-hour urine samples. During the first visit to the laboratory, participants provided demographic information (age, sex, etc.), and height, without shoes, was measured in centimeters using a research-grade digital
stadiometer, and weight, in light clothing without shoes, was measured to the nearest 0.1 kg using a calibrated digital Tanita scale (Model: TBF-310GS; Tokyo, Japan).

**Figure 1.** Study design and timeline of a non-nutritive sweetener biomarker validity study

Participants were provided two 3-liter 24-hour urine collection containers for each urine collection and detailed instructions for collecting their sample during visit 1. During the day prior to the second visit, participants collected a 24-hour urine sample in 3-liter containers provided to them. For each urine sample, the urine containers contained 6.75 ml of a thymol 10% MeOH solution. This thymol solution acts as a preservative, allowing for the sample to be kept at room temperature while controlling for bacteria\(^4\).

When participants returned for their second visit, total urine volume was measured and recorded. A trained research assistant, supervised by a PhD level Registered Dietitian Nutritionist, collected a 24-hour dietary recall from the previous day. Between the second and third visit, participants completed one unannounced dietary recall via phone call. At the final visit, total urine volume was measured and recorded, and a third 24-hour dietary recall was collected from the previous day.
Among the 125 participants who completed the study, 9 participants were chosen for this preliminary analysis due to reporting either high total NNS intake or no NNS intake via 24-hour dietary recalls. Two of the 9 participants from the preliminary analysis were chosen to demonstrate how the UPLC chromatograms can be interpreted. UPLC analysis was conducted to detect inclusions in participants’ 24-hr urine samples. Specifically, UPLC analysis was used to detect the presence of saccharin, acesulfame potassium, sucralose, and stevia and related metabolites. Urine samples were stored in 1.8 ml cryovials in a -80°C freezer until UPLC analysis was conducted.

**Development of UPLC Methodology**

Urine samples were prepared for ULPC analysis by combining 1.5 mL urine with 0.15 mL internal standard [1 mg/mL $^{13}$C$_6$-glucose in water] in a labeled 15 mL centrifuge tube, leaving the cap off. Tubes were covered with a 1-inch square of Parafilm. Using a needle, 2-3 holes were made in the Parafilm. The urine tubes were frozen in a −80°C freezer or on dry ice and then were set to freeze dry overnight. The following day, 200 μL extraction solvent (50% MeOH, 50% 1.5 M formic acid in water) were added to the tubes and tubes were capped. Tubes were vortexed for 30 seconds and sonicated for 5 minutes. Liquid was transferred to a labeled 1.5 or 2 mL microcentrifuge tube and tubes were centrifuged for 3 minutes (17,000 x g). The supernatant was filtered (0.2 μm PTFE syringe filter) into a labeled LC-MS vial containing a 150 μL glass insert and vials were capped and analyzed by UPLC-MS/MS.

UPLC separation was performed on a Waters Acquity H-class (Milford, MA) equipped with an Acquity UPLC BEH Amide column (2.1 mm × 50 mm, 1.7 μm particle size) + a BEH amide guard column (2.1 mm × 5 mm, 1.7 μm particle size). Binary gradient elution was
performed at 0.7 mL/min using acetonitrile:water (65:35) with 0.2% v/v triethylamine (TEA) (phase A) and 0.1% formic acid in water (phase B). The following linear gradient was employed: 1% A (0 min), 99% A (3 min), 1% A (3.05 min), and 1% A (5.5 min). Column and sample temperatures were 35 and 10°C, respectively. Detection by MS/MS was performed on a Waters Acquity Triple Quadrupole Detector (TQD). Negative-mode electrospray ionization (−)-ESI was performed with capillary voltage of −4 kV, and source and desolvation temperatures of 150 and 450°C, respectively. Desolvation and cone gasses were N2 at flow rates of 900 and 50 L/h, respectively. For MS/MS, the collision gas was Ar (0.1 mL/min). Urinary levels of the following compounds will be quantified using Multiple Reaction Monitoring (MRM): Sucralose, Saccharin, Acesulfame Potassium, Erythritol, and Stevia components (measured as their deconjugated and then Phase-II metabolite steviol glucuronide).

**Table 1.** Optimized mass spectrometry parameters for the quantitation of non-nutritive sweeteners (NNS)

<table>
<thead>
<tr>
<th>Compound</th>
<th>Parent [M−H]− (m/z)</th>
<th>Daughter (m/z)</th>
<th>Cone (V)</th>
<th>Collision (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saccharin</td>
<td>181.78</td>
<td>41.92</td>
<td>100</td>
<td>16</td>
</tr>
<tr>
<td>Sucralose</td>
<td>394.73</td>
<td>358.92</td>
<td>98</td>
<td>8</td>
</tr>
<tr>
<td>Erythritol</td>
<td>120.83</td>
<td>88.90</td>
<td>18</td>
<td>6</td>
</tr>
<tr>
<td>Acesulfame K</td>
<td>161.76</td>
<td>81.93</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>Steviol glucuronide</td>
<td>493.20</td>
<td>317.30</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>13C6-glucose (IS)</td>
<td>184.57</td>
<td>91.8</td>
<td>16</td>
<td>8</td>
</tr>
</tbody>
</table>

The cone voltages, collision energy, and MRM transitions were developed and optimized for each analyte using the Intellistart function of Waters MassLynx v4.1 software. Peak widths were ~4 s, and AutoDwell was employed with required points-per-peak set at 12. The interscan delay time was 0.02 s.
Results

Preliminary analysis was conducted among 9 adults (mean age=35.4±15.0 years, range: 21-61). The UPLC-MS/MS method was capable of detecting broad differences, ranging from very strong presence of native NNS and steviol glucuronide to not detectable, which was also observed for NNS intake reported in the 24-hr dietary recalls. Two participants’ samples were selected for demonstration based on their self-reported NNS intake via 24-hour dietary recall. The NNS consumer reported a total of 271.5mg of NNS, which included sucralose, ace k, and aspartame in the dietary recall, whereas the NNS non-consumer reported 0 mg of NNS. UPLC raw chromatogram outputs were able to distinguish between a NNS consumer and non-consumer (Figure 2).

Based on this analysis, NNS can be detected in 24-hour urine samples based on ad libitum dietary intake (i.e., without a dose administration). NNS was detected even at low intensities as seen in raw chromatogram outputs. In the chromatograms included, the detection intensity of each compound is plotted over the course of the 5-minute retention time. Acesulfame potassium was detected at an intensity level about 10,000 times higher in the NNS consumer’s sample compared to the NNS non-consumer’s sample. Steviol glucuronide was not detected in the NNS non-consumer and may have been detected in the NNS consumer. Erythritol, sucralose, and saccharin chromatograms indicate that the detection level needs to be explored, as clear peaks were not identified.
Figure 2. Raw chromatogram UPLC outputs from a non-nutritive sweetener (NNS) consumer and a NNS non-consumer

Discussion

The ability to objectively measure individual NNS levels in the urine will be useful in future clinical studies that seek to examine the impact of NNS dietary intake on health outcomes. One of the primary benefits of this methodology will be the ability to quantify intake of individual types of NNS, rather than studying NNS as a whole. Currently, research investigations typically assess NNS as one group rather than individual compounds. NNS are each distinct chemical compounds with unique metabolic pathways when consumed, with some being broken
down by the body while other are excreted in the urine or feces\textsuperscript{41,42,45}. As such, they likely have differing impacts on the health of the consumer. Because NNS are often used in a variety of combinations, being able to capture multiple types of NNS intake will allow researchers to understand different health impacts associated with each type.

This NNS UPLC methodology will be beneficial in overcoming errors and bias due to self-reported assessment methods. As previously mentioned, there is a wide discrepancy among observational studies regarding what percentage of the population consume NNS. Researchers are currently limited to self-report methods for capturing NNS intake, where it is known that research participants tend to underreport intake foods perceived to be socially undesirable\textsuperscript{46}. This methodology will allow researchers to gather more accurate dietary intake information as well as provide an object measure against which to validate subjective (i.e. self-report) assessment methods.

Previous investigations have developed mass spectrometry methodology for individual NNS types (Table 2). These investigations will be useful for quantifying NNS content in urine samples as well as for anticipating mean recovery percentages. This methodology expands upon previous investigations by developing a single methodology that can be used to detect five different NNS. To our knowledge, this is the first UPLC methodology to comprehensively assess five NNS metabolites.
Table 2. Spiking range and mean recovery percentage of previous NNS biomarker investigations

<table>
<thead>
<tr>
<th>NNS Type</th>
<th>Study Type</th>
<th>Spiking Range</th>
<th>Mean Recovery Percentage (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saccharin</td>
<td>Observational/Ad Libitum</td>
<td>0.4-40.8 mg/L</td>
<td>84.6% (65.7-103.5%)</td>
</tr>
<tr>
<td>Saccharin</td>
<td>Intervention/Dose-Response</td>
<td>0.8-80.2 mg/L</td>
<td>83.7% (63.2-116.4%)</td>
</tr>
<tr>
<td>Acesulfame Potassium</td>
<td>Observational/Ad Libitum</td>
<td>0.6-64.1 mg/L</td>
<td>74.0% (51.5-93.6%)</td>
</tr>
<tr>
<td>Acesulfame Potassium</td>
<td>Intervention/Dose-Response</td>
<td>1.2-118.6 mg/L</td>
<td>67.6% (50.1-91.6%)</td>
</tr>
<tr>
<td>Stevia (Steviol Glucuronide)</td>
<td>Intervention/Dose-Response</td>
<td>-</td>
<td>75.50%</td>
</tr>
<tr>
<td>Stevioside and rebaudioside A</td>
<td>Intervention/Dose-Response</td>
<td>-</td>
<td>Reb A 59%</td>
</tr>
<tr>
<td>Stevioside and rebaudioside A</td>
<td>Intervention/Dose-Response</td>
<td>-</td>
<td>Stevioside 62%</td>
</tr>
<tr>
<td>Sucralose (Radioactive 14C-Sucralose)</td>
<td>Intervention/Dose-Response</td>
<td>-</td>
<td>Urine: 14.5% (8.9-21.8%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Feces: 78.3% (69.4-89.6%)</td>
</tr>
</tbody>
</table>

**Limitations**

This investigation is not without limitations. This study relies on self-reported dietary data from 24-hour dietary recalls. As previously discussed, self-reported data has limited validity due to reporting error. To help address this limitation, graduate level research assistants were trained to administer 24-hour dietary recalls and state-of-the-art dietary analysis software was used. Additionally, to establish rapport and improve data integrity, the same research assistant completed all dietary recalls and completed the data entry.

Additionally, the study used in the initial investigation did not utilize para-aminobenzoic acid (PABA) as a way to measure the completeness of participants’ 24-hour urine samples. When administrated to participants in known quantities, PABA is known to be excreted within a certain percentage range in complete 24-hour urine samples. To mitigate this concern, researchers gave participants detailed instructions for completing urine samples and provided reminder emails the day prior to sample collection. Moreover, participants who reported more than a 2-hour margin of error (i.e., less than 22 hours and greater than 26 hours of collection
time), will be excluded from the future validity analysis of this methodology. Future investigations should consider using PABA to verify completeness of 24-hour urine samples.

Another limitation is the inability of UPLC analysis to detect aspartame in the urine. Aspartame is a commonly consumed NNS, but because it is metabolized into aspartic acid, phenylalanine, and methanol\textsuperscript{64}, it is not expected to be detectable in the urine.

Next Steps and Future Directions

This preliminary investigation determined that this NNS urinary biomarker is able to detect the presence of NNS and metabolites in free-living adults. However, further analysis must be completed to thoroughly assess the validity of this biomarker. Future work will include data acquisition, processing, and quantification using Waters MassLynx v4.1 software. Urinary sweetener concentrations will be calculated based on peak area ratios compared to the internal standard and calculated based on standard curves prepared from authentic standards (compounds obtained from Sigma, St. Louis, MO). Calculated urinary NNS concentrations will be converted to total mass excreted based on total urine volume.

Statistical analyses will be conducted using IBM SPSS statistical analysis software (v. 25 for Windows, 2017, SPSS Inc., Chicago, IL). Descriptive statistics (mean ± standard deviation and frequencies) will be utilized to analyze participant demographics. The validity of the NNS biomarker will be measured by comparing NNS quantities detected by the biomarker to those reported in the participants’ dietary recall from dietary intake reported on the previous day. The specificity, or false positive rate, will be determined by analyzing NNS quantities detected by the biomarker for those participants who reported no NNS intake in their dietary recalls. Pearson’s
correlations (r), paired sample t-tests, and Bland-Altman analyses will be conducted to measure comparative validity and specificity of the NNS biomarker.

Future investigations should establish a lower limit of detection to measure quantitative amounts of NNS in urine samples and compare these amounts to dietary recalls using a sample size adequately powered for this analysis. The ability to objectively measure NNS intake will help researchers and clinicians inferentially examine the relationship between NNS intake and health outcomes.
References


Chapter 5 – A systematic review of study design elements and outcomes of non-nutritive sweetener and weight status research: Where do we go from here?

Abstract

Background: Non-nutritive sweeteners (NNS), or artificial sweeteners, have been suggested as alternatives to added dietary sugars to reduce overall energy intake and aid in weight management. Results from studies on NNS and weight status have been inconclusive thus far, with results suggesting both weight loss and weight gain. Study design types and methods vary widely among studies making difficult to draw conclusions regarding the impact of NNS on weight status.

Objective: The objective of this systematic review was to determine if study design type (observational versus randomized controlled trial) and study methods, including type of NNS examined, mode of NNS delivery, NNS intake assessment method, and weight assessment method, were associated with different outcomes in studies on NNS and weight-related outcomes among adult populations.

Design: A systematic literature review found 41 relevant articles for inclusion.

Results: The articles identified included 25 observational studies, 14 of which were cross-sectional studies and 11 of which were longitudinal. Among the observational studies, 76% reported an association of higher weight outcomes with NNS intake. Of the 16 randomized controlled trials, 81% indicated NNS interventions resulted in improved weight outcomes across NNS types.

Conclusion: This review demonstrates that the wide variety of study methods used make comparing weight-related outcomes across studies challenging. In order to truly evaluate how
NNS impact weight status, researchers need to implement RCT examining individual NNS types, using appropriate NNS sources, and with specific NNS intake and weight assessment methods.
**Introduction**

Non-nutritive sweeteners (NNS), or artificial sweeteners, are substances containing a concentrated sweet flavor with negligible caloric content\(^1\). Since their introduction to food products, they have been marketed for their ability to reduce overall energy intake and assist in weight management when replacing caloric foods and sugar-sweetened beverages. Because these high-intensity sweeteners range between 160 and 1,000 times sweeter than sugar\(^1\), only a very small amount of NNS substance is needed to taste the sweet flavor. However, conflicting evidence exists on whether or not NNS is associated with weight outcomes. Despite their lack of caloric content, there is substantial controversy over whether or not NNS contribute to weight gain\(^2\).

Body weight and other weight-related measurements, such as waist circumference, total body fat percentile, and abdominal adiposity, are anthropometric indicators of cardiometabolic risk. These measurements have been associated with non-communicable diseases including type 2 diabetes, hypertension, and heart disease\(^3,4\). With obesity rates continuing to rise globally\(^5\), added sugars, frequently in the form of sugar-sweetened beverages, are often considered a target for reducing energy intake\(^6,7\), as sugar-sweetened beverages contribute approximately 7\% of total caloric intake among adults\(^8\). Diet sodas, which are sweetened with a variety of NNS, are often used as a substitute for sugar-sweetened beverages. Therefore, many researchers and healthcare practitioners have suggested that replacing added sugars in the diet with NNS, frequently in the form of diet sodas, could be a feasible method to reduce energy intake, thereby decreasing the prevalence of overweight and obesity\(^9,10\).

Conducting conclusive research on NNS and weight-related outcomes is challenging for many reasons. Ideally, double-blind, randomized controlled trials (RCT) should be used\(^11\), as
they allow researchers to control for confounding variables. However, RCT on human subjects are expensive and burdensome to carry out\textsuperscript{12}. Thus, researchers often rely on observational studies on this topic because of the relative ease with which large sample sizes can be collected as well as the availability of large prospective cohort study data. However, observational data cannot determine causation, due to the potential of confounding variables and reverse causality\textsuperscript{13}.

Studies on NNS consumption are currently limited due to self-reported dietary intake information, which has known levels of error and reporting bias\textsuperscript{11}. Furthermore, there are limitations due to the current practice of examining NNS intake as a whole rather than distinguishing between types of NNS\textsuperscript{14}. Due to these difficulties, research using a variety of study design characteristics can be found in the literature.

There are six NNS that are currently approved by the U.S. Food and Drug Administration (FDA): saccharin, aspartame, acesulfame potassium, neotame, advantame, and sucralose\textsuperscript{15}. Stevia and luo han guo, also known as monk fruit, are classified by the FDA as “Generally Recognized As Safe,” or GRAS, but have not been evaluated for FDA approval\textsuperscript{16}. As each NNS has a distinct chemical profile and metabolic pathway, impacts of each on weight-related outcomes are likely to differ\textsuperscript{1}. This important consideration is often overlooked, and research studies that associate total NNS intake with weight outcomes may be contributing to misunderstandings on the topic among researchers, clinicians, and consumers.

Thus far, research findings on the impact of NNS on weight outcomes have been conflicting, with numerous studies identifying associations with weight loss as well as weight gain. Previous reviews have explored study design with NNS and weight outcomes, in addition to other metabolic risk factors\textsuperscript{17,18}. Miller et al.,\textsuperscript{17} determined that RCT were more likely to find improved weight outcomes, while observational studies were less conclusive. On the other hand,
Azad et al., found that observational studies were more likely to find higher weight outcomes while RCT were less conclusive. This review expands upon these previous investigations by examining additional study design elements and methods and how study characteristics may impact outcomes.

The objective of this systematic review is to determine if study design type (observational study versus randomized controlled trial) and study methods, including type of NNS studied, mode of delivery of NNS (food versus beverage), NNS intake assessment method, and weight assessment method, are associated with different outcomes in studies on NNS and weight status among adult populations. In addition, this review provides recommendations for strengthening NNS study designs and methods for future investigations.

Methods

A literature search was conducted for relevant articles published using PubMed (MEDLINE), PubMed Central, Web of Science, CINAHL, and COCHRANE databases to search for relevant articles. A variety of search terms were required due to the many ways that NNS are referenced and the various measures of weight status and weight-related outcomes that are reported in the literature. The following search terms were used to identify relevant articles:

“Non-Nutritive Sweetener” OR “Non-Nutritive Sweeteners” OR “Nonnutritive Sweetener” OR “Nonnutritive Sweeteners” OR “NNS” OR “Artificial Sweetener” OR “Artificially Swe...
Beverages” OR “Diet Food” OR “Diet Foods” OR “Saccharin” OR “Aspartame” OR “Acesulfame Potassium” OR “Ace K” OR “Acesulfame K” OR “Sucralose” OR “Stevia” OR “Stevioside” OR “Steviol” AND “Weight” OR “Weight Status” OR “Weight Loss” OR “Weight Gain” OR “Weight Change” OR “Weight Maintenance” OR “Body Composition” OR “Body Mass Index” OR “BMI” OR “Waist Circumference” OR “Body Fat”.

This systematic review was limited to clinical trials, randomized control trials, observational studies, prospective cohort studies, and reviews published prior to May 15, 2018. Additional articles were identified based on relevant review papers identified17,18. To be included in the review, articles had to 1) include human subjects, 2) include adults aged 18 years or older, 3) use randomized controlled trial or observational (cross-sectional or longitudinal) study designs, 4) measure NNS intake, and 5) measure weight status, weight loss, weight gain, body mass index (BMI), total body fat percentile, waist circumference, and/or abdominal adiposity. Exclusion criteria included 1) studies on non-human subjects, 2) studies on special populations (e.g., pregnant women, diabetic participants), 3) studies on sugar alcohols, and 4) publications in non-English languages. Sugar alcohols were not included in this review, as they do contain nutritive content19 and are not typically categorized as NNS. These criteria were selected to target studies on NNS and weight-related outcomes among general adult populations without specific disease states.

An initial search in the five databases returned 3,147 articles. Articles were screened according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines20. A flow diagram of the screening and eligibility process is shown in Figure 1. After duplicates were removed, each article’s title and abstract (n=1,913) were reviewed for inclusion and exclusion criteria. Full text articles (n=114) were downloaded and
reviewed for inclusion and exclusion criteria.

**Figure 1.** Study selection process flow diagram

Studies that were included after the full text review (n=41) were assessed for quality and risk of bias. Observational studies were screened using the National Institutes of Health Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies\(^\text{21}\). Studies were evaluated based on fourteen criteria. The Cochrane Risk of Bias Tool\(^\text{22}\) was used to assess bias among the randomized controlled trials. Studies were evaluated based on seven criteria to
measure potential selection, reporting, performance detection, and attrition biases. Results from risk of bias analyses are summarized in the Appendix B and Appendix C.

Results

Forty-one articles were identified as relevant for inclusion in this review. Study design, study methods, and a brief summary of weight-related findings are summarized in Table 1 and Table 2. The term “weight-related outcomes” is used throughout this analysis to reflect the variety of measurements used to capture measures of body weight. The number of studies that utilized each type of study design, NNS type, NNS mode of delivery, and NNS intake assessment method are shown in Table 3. Also included in Table 3 is an indication of weight-related outcomes reported (indicating higher weight-related results, lower-related, or neither). Weight-related outcome frequencies and percentages based on NNS methods by study design type (observational or cross-sectional) are included in Table 4.
Table 1. Observational studies evaluating NNS intake and weight status

<table>
<thead>
<tr>
<th>Study Population</th>
<th>Study Sample Size (n)</th>
<th>Study Design</th>
<th>NNS Type Examined</th>
<th>NNS Mode of Delivery</th>
<th>NNS Intake Assessment Method</th>
<th>Body Weight Assessment Method</th>
<th>Study Duration</th>
<th>Weight-Related Outcome Results</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults; 18 years</td>
<td>n = 120 (51 Heavy NNS Beverage Consumers; 69 Non-Habitual NNS Beverage Consumers)</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>FFQ</td>
<td>Self-Reported</td>
<td>N/A</td>
<td>↑ BMI</td>
<td>Heavy NNS beverage consumers reported significantly higher BMI (25.7±4.5 kg/m²) compared to non-NNS users (22.4±2.4 kg/m²).</td>
</tr>
<tr>
<td>Adults; 45-60 years; SU.VI.MAX cohort</td>
<td>n = 4278 (2299 males; 1979 females)</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Food and Beverages</td>
<td>24-hour dietary recalls</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↑ BW ↑ BMI ↑ Waist/Hip Ratio</td>
<td>Weight, BMI, and waist/hip ratio were significantly higher among high NNS consumers compared to non-consumers.</td>
</tr>
<tr>
<td>Adults; ≥20 years; NHANES 1999-2010</td>
<td>n = 23965 (12052 males; 11913 females)</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>24-hour dietary recalls</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↑ Body Weight</td>
<td>Compared to those who drink SSB, overweight (19%) and obese (22%) adults drink more diet beverages than healthy-weight adults (11%).</td>
</tr>
<tr>
<td>Adults; ≥18 years; NHANES 2003-2004</td>
<td>n = 3828 (1757 males; 2071 females)</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Coffee/Tea</td>
<td>FFQ</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↑ BMI ↑ WC</td>
<td>NNS use in coffee/tea was associated with higher BMI (28.2 vs 27.1 in men, 28.4 vs 27.1 in women; p≤0.05); Women who use NNS in their coffee/tea had a 1.7 kg/m² higher BMI, and men who use NNS in their coffee/tea had a 4.7 cm higher waist circumference compared to those who did not use NNS in coffee/tea.</td>
</tr>
<tr>
<td>Study Population</td>
<td>Study Sample Size (n)</td>
<td>Study Design</td>
<td>NNS Type Examined</td>
<td>NNS Mode of Delivery</td>
<td>NNS Intake Assessment Method</td>
<td>Body Weight Assessment Method</td>
<td>Study Duration</td>
<td>Weight-Related Outcome Results</td>
<td>Outcomes</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Adults; ≥18 years; NWCR&lt;sup&gt;27&lt;/sup&gt;</td>
<td>n = 434 (123 males; 311 females)</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>FFQ</td>
<td>Self-Reported</td>
<td>N/A</td>
<td>↑ BMI</td>
<td>Percentage of participants reporting habitual consumption of NNS beverages was significantly different between BMI category (2.57% of normal weight, 21.1% of overweight, and 35.1% of obese, overall P=0.0447).</td>
</tr>
<tr>
<td>Adults; ≥18 years; consuming ≥200kcal of SSB/day&lt;sup&gt;14&lt;/sup&gt;</td>
<td>n = 301 (NNS Consumer = 100; NNS Non-Consumer = 201)</td>
<td>Observational; Cross-sectional</td>
<td>Combination</td>
<td>Food and Beverages</td>
<td>24-hour dietary recalls</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↑ BW↑ BMI</td>
<td>Mean body weight was higher among NNS consumers (94.4±28.6) compared to non-consumers (88.5±23.4) (P=0.06). Mean BMI was higher among NNS consumers (34.7±10.6) compared to non-consumers (32.1±8.2) (P=0.02).</td>
</tr>
<tr>
<td>Adults; ≥18 years&lt;sup&gt;28&lt;/sup&gt;</td>
<td>n = 125 (54 males, 71 females)</td>
<td>Observational; Cross-sectional</td>
<td>Combination</td>
<td>Food and Beverages</td>
<td>24-hour dietary recalls</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↑ BMI↑ WC</td>
<td>BMI was significantly higher among NNS consumers (F=10.3 (0.002), X² = 9.0 (0.03) compared to non-consumers. Total NNS intake and some individual NNS types [...saccharin, sucralose and ace k] were significant predictors of waist circumference.</td>
</tr>
<tr>
<td>Adults; 18-60 years; ENRICA Study&lt;sup&gt;29&lt;/sup&gt;</td>
<td>n = 9178</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>Surveys and Questionnaires</td>
<td>Clinical Measurement</td>
<td>2-4 years</td>
<td>↑ BMI</td>
<td>At baseline, more frequent NNS beverage consumption was associated with higher BMI (≥1/week = 27.4±5.4; 1-6/week = 26.9±6.1; &lt;1/week = 26.1±6.4)</td>
</tr>
</tbody>
</table>

113
<table>
<thead>
<tr>
<th>Study Population</th>
<th>Study Sample Size (n)</th>
<th>Study Design</th>
<th>NNS Type Examined</th>
<th>NNS Mode of Delivery</th>
<th>NNS Intake Assessment Method</th>
<th>Body Weight Assessment Method</th>
<th>Study Duration</th>
<th>Weight-Related Outcome Results</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult females; 18-24 years&lt;sup&gt;30&lt;/sup&gt;</td>
<td>n = 185</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Food and Beverages</td>
<td>Surveys and Questionnaires</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↓ BMI</td>
<td>43% of normal weight, 49% of overweight, and 29% of obese college-aged women report consuming sugar-free versions of foods/drinks; 31% of normal weight and overweight and 5% of obese college-aged women report consuming NNS.</td>
</tr>
<tr>
<td>Adults; 18-70 years&lt;sup&gt;31&lt;/sup&gt;</td>
<td>n = 791</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>FFQ</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↑ WC ↑ BMI ↑ TBF %</td>
<td>Increased frequency of NNS beverage intake was associated with greater waist circumference, BMI, and TBF%, but was not associated with variation in VAT or SAT mass or VAT%.</td>
</tr>
<tr>
<td>Adults&lt;sup&gt;32&lt;/sup&gt;</td>
<td>n = 303 (172 weight loss maintainers; 131 normal weight)</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Food and Beverages</td>
<td>24-hour dietary recalls</td>
<td>Self-Reported</td>
<td>N/A</td>
<td>↓ BW</td>
<td>Adults successfully maintaining weight loss report consuming three times more daily servings of NNS beverages (0.91 vs. 0.37; P=0.003) than normal weight individuals.</td>
</tr>
<tr>
<td>Adults; ≥18 years; NHANES 2009-2012&lt;sup&gt;33&lt;/sup&gt;</td>
<td>n = 11098 (5465 males; 5633 females)</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Food and Beverages</td>
<td>24-hour dietary recalls</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↑ BMI</td>
<td>Prevalence of consuming NNS three or more times daily increased with BMI in adults (19.2% obese adults vs. 13.4% normal weight adults; P&lt;0.001)</td>
</tr>
<tr>
<td>Adults; 18-65 years; morbid obesity&lt;sup&gt;34&lt;/sup&gt;</td>
<td>n = 100 (17 males, 83 females)</td>
<td>Observational; Cross-sectional</td>
<td>Not Specified</td>
<td>Beverages and Packets</td>
<td>FFQ</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>No significant changes</td>
<td>No associations found between NNS and BMI in this study. Authors believe this may be due to the weight status of participants, however, it could indicate a lack of weight-reducing effect of NNS.</td>
</tr>
<tr>
<td>Study Population</td>
<td>Study Sample Size (n)</td>
<td>Study Design</td>
<td>NNS Type Examined</td>
<td>NNS Mode of Delivery</td>
<td>NNS Intake Assessment Method</td>
<td>Body Weight Assessment Method</td>
<td>Study Duration</td>
<td>Weight-Related Outcome Results</td>
<td>Outcomes</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>----------------------</td>
<td>----------------------------</td>
<td>-----------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Adults; ≥20 years; NHANES 1988-1994</td>
<td>n = 15721 (7403 males; 8328 females)</td>
<td>Observational; Cross-sectional</td>
<td>Aspartame</td>
<td>Not Specified</td>
<td>Surveys and Questionnaires</td>
<td>Clinical Measurement</td>
<td>N/A</td>
<td>↑ SAAT</td>
<td>Aspartame intake was positively associated with abdominal adiposity (OR:1.8, 95% CI: 1.1-1.26).</td>
</tr>
<tr>
<td>Adults; ≥20 years; BLSA</td>
<td>n = 1454 (741 males; 713 females)</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Food and Beverages</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>Up to 28 years (median follow-up of 10 years)</td>
<td>↑ BMI ↑ WC ↑ SAAT</td>
<td>NNS users had 0.80 kg/m² higher BMI, 2.6 cm larger WC, and 36.7% higher prevalence and 53% higher incidence of abdominal obesity compared to non-users.</td>
</tr>
<tr>
<td>Adult females; 30-55 years; NHS</td>
<td>n = 31940</td>
<td>Observational; Longitudinal</td>
<td>Saccharin</td>
<td>Not Specified</td>
<td>FFQ</td>
<td>Self-Reported</td>
<td>8 years</td>
<td>↑ BW</td>
<td>Higher saccharin consumers had higher mean body weights (61.9 kg vs 67.0 kg) and higher weight gain over the first four years (0.37 kg/year vs 0.57 kg/year) and the second four years (0.35 kg/year vs 0.47 kg/year).</td>
</tr>
<tr>
<td>Adults; 25-64 years; SAHS</td>
<td>n = 5158 (baseline) n = 3682 (follow-up)  (1767 non-NNS users; 1604 NNS users)</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>Combination</td>
<td>Clinical Measurement</td>
<td>7-8 years</td>
<td>↑ BMI</td>
<td>BMI changes were 47% higher among NNS users than non-NNS users (+1.48 vs +1.01 kg/m², P&lt;0.0001).</td>
</tr>
<tr>
<td>Adults; 65+ years; SALSA</td>
<td>n = 749 baseline; 474 1-year follow-up; 413 2-year follow-up; 375 3-year follow-up</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Diet Soda</td>
<td>Surveys and Questionnaires</td>
<td>Clinical Measurement</td>
<td>Up to 4 years</td>
<td>↑ WC</td>
<td>Change in waist circumference was highest among high diet soda consumers (3.04 cm) and lowest among non-consumers (0.77 cm)</td>
</tr>
<tr>
<td>Study Population</td>
<td>Study Sample Size (n)</td>
<td>Study Design</td>
<td>NNS Type Examined</td>
<td>NNS Mode of Delivery</td>
<td>NNS Intake Assessment Method</td>
<td>Body Weight Assessment Method</td>
<td>Study Duration</td>
<td>Weight-Related Outcome Results</td>
<td>Outcomes</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>-------------------------------</td>
<td>-------------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Adults; 45-84 years; MESA&lt;sup&gt;40&lt;/sup&gt;</td>
<td>n = 2961</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>FFQ</td>
<td>Clinical Measurement</td>
<td>3-7 year follow-up</td>
<td>↑ WC ↑ BMI</td>
<td>More frequent NNS beverage consumption was associated with higher BMI (rarely/never = 27.3±0.1; &lt;1/week = 28.3±0.2; 1/week-1/day = 28.5±0.2; ≥1/day = 29.3±0.2; p&lt;0.001) and higher waist circumference (rarely/never = 95.6±0.3; &lt;1/week = 97.2±0.6; 1/week-1/day = 98.3±0.5; ≥1/day = 100.6±0.5; p&lt;0.001)</td>
</tr>
<tr>
<td>Adults; 27-64 years; NHS; NHS II; HPFS&lt;sup&gt;41&lt;/sup&gt;</td>
<td>n = 50013 females from NHS; n = 52987 females from NHS II; n = 21988 males from HPFS</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>FFQ</td>
<td>Self-Reported</td>
<td>4-year follow-up intervals</td>
<td>↓ BW</td>
<td>Each 1-cup increase in diet beverage intake was inversely associated with weight gain within each 4-year interval (-0.10 kg; 95% CI: 0.14, 0.06)</td>
</tr>
<tr>
<td>Adults; 18-64 years; PHHP&lt;sup&gt;42&lt;/sup&gt;</td>
<td>n = 465 (176 males; 289 females)</td>
<td>Observational; Longitudinal</td>
<td>Saccharin</td>
<td>Not Specified</td>
<td>FFQ</td>
<td>Clinical Measurement</td>
<td>4-year follow-up</td>
<td>↑ BW</td>
<td>Higher levels of saccharin consumption were associated with higher mean baseline weight (0 g/day = 69.5±1.0 kg; 0.1-28.2 g/day = 74.9±2.1 kg; &gt;28.2 g/day = 76.4±1.3 kg) and higher mean weight change (0 g/day = 0.4±0.4 kg; 0.1-28.2 g/day = 0.1±0.7 kg; &gt;28.2 g/day = 1.4±0.4 kg)</td>
</tr>
<tr>
<td>Adult females; 24-44 years; NHS II&lt;sup&gt;43&lt;/sup&gt;</td>
<td>n = 51603</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>FFQ</td>
<td>Clinical Measurement</td>
<td>4-year follow-up</td>
<td>↓ BW</td>
<td>Participants who increased their diet soft drink consumption from 1 drink or less per week to 1 drink or more per day gained less weight (+1.59 kg) compared with those who decreased their diet soft drink consumption from 1 drink or more per day to 1 drink or less per week (4.25 kg)</td>
</tr>
<tr>
<td>Study Population</td>
<td>Study Sample Size (n)</td>
<td>Study Design</td>
<td>NNS Type Examined</td>
<td>NNS Mode of Delivery</td>
<td>NNS Intake Assessment Method</td>
<td>Body Weight Assessment Method</td>
<td>Study Duration</td>
<td>Weight-Related Outcome Results</td>
<td>Outcomes</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>----------------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Adults; 27-64 years; NHS I; NHS II; HPFS[^43]</td>
<td>n = 117992 (21472 males; 96520 females)</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Diet Soda</td>
<td>FFQ</td>
<td>Self-Reported</td>
<td>Up to 24 years (4-year follow-up intervals)</td>
<td>↑ BW</td>
<td>Diet soda was associated with 0.31 (0.27, 0.35) (p&lt;0.001) weight change in prevalence analysis and 0.20 (0.09, 0.32) (p&lt;0.001) weight change in lagged change analysis.</td>
</tr>
<tr>
<td>Adult females; 50-69 years[^44]</td>
<td>n = 78694</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Not Specified</td>
<td>Surveys and Questionnaires</td>
<td>Self-Reported</td>
<td>6 years</td>
<td>↑ BW</td>
<td>No significant differences between percentages of NNS users and non-users who lost weight at any initial relative weight level. NNS users were significantly more likely to gain weight, regardless of initial weight (P&lt;0.001).</td>
</tr>
<tr>
<td>Adult females; 35-45 years[^46]</td>
<td>n = 170</td>
<td>Observational; Longitudinal</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>FFQ</td>
<td>Clinical Measurement</td>
<td>4 years</td>
<td>↓ BW</td>
<td>Females who consumed SSB predominantly gained significantly more body weight (+2.68±5.14 kg) than women who consumed NNS beverages (-0.05±4.40 kg) or no soft drinks (-0.50±5.05 kg), but did not differ significantly from those who consumed a mix of SSB and NNS beverages (+1.22±5.06).</td>
</tr>
</tbody>
</table>

Table 2. Randomized controlled trials evaluating NNS intake and weight status

<table>
<thead>
<tr>
<th>Study Population</th>
<th>Study Sample Size (n)</th>
<th>Study Design</th>
<th>NNS Type Examined</th>
<th>NNS Mode of Delivery</th>
<th>NNS Intake Assessment Method</th>
<th>Body Weight Assessment Method</th>
<th>Study Duration</th>
<th>Weight-Related Outcome Results</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults; 19-60 years(^{47})</td>
<td>n = 16 (8 males, 8 females)</td>
<td>Randomized Controlled Trial</td>
<td>Stevia</td>
<td>Powder mixed with hot beverage</td>
<td>None included</td>
<td>Clinical Measurement</td>
<td>3 weeks (1-week stevia intervention, 1-week placebo, 6 days washout)</td>
<td>↓ BW ↓ BMI</td>
<td>Small but insignificant reductions in weight (P=0.246) and BMI (P=0.249) after stevia intervention were detected.</td>
</tr>
<tr>
<td>Obese adult females; 20-60 years(^{48})</td>
<td>n = 163 (82 aspartame group; 81 no-aspartame)</td>
<td>Randomized Controlled Trial</td>
<td>Aspartame</td>
<td>Food and Beverages</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>19 weeks (active weight loss), 71 weeks (weight maintenance), 175 weeks (weight maintenance)</td>
<td>↓ BW</td>
<td>Participants consuming aspartame lost more weight (P=0.028) and regained less weight during maintenance and follow-up (P=0.046) than those not consuming aspartame.</td>
</tr>
<tr>
<td>Normal-weight adult females; 18-50 years(^{49})</td>
<td>n = 49 (13 control; 17 reduced-fat; 19 reduced-sugar)</td>
<td>Randomized Controlled Trial</td>
<td>Not Specified</td>
<td>Food and Beverages</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>10 weeks</td>
<td>No significant changes</td>
<td>Body weights did not change significantly based on intervention type within the 4-week intervention.</td>
</tr>
<tr>
<td>Normal weight adults; 18-60 years(^{50})</td>
<td>n = 100</td>
<td>Randomized Controlled Trial</td>
<td>Aspartame</td>
<td>Diet Soda and Capsules</td>
<td>None included</td>
<td>Clinical Measurement</td>
<td>12-week intervention</td>
<td>No significant changes</td>
<td>No main effect of aspartame treatment on body weight.</td>
</tr>
<tr>
<td>Overweight and obese adults; 20-60 years(^{51})</td>
<td>n = 59 (13 males; 46 females)</td>
<td>Randomized Controlled Trial</td>
<td>Aspartame</td>
<td>Food and Beverages</td>
<td>Combination</td>
<td>Clinical Measurement</td>
<td>12-week intervention with 1-year follow-up</td>
<td>↓ BW (females)</td>
<td>Females lost an average of 12.8 lbs in the control group and 16.5 lbs in the aspartame group; No significant differences in the male group.</td>
</tr>
<tr>
<td>Study Population</td>
<td>Study Sample Size (n)</td>
<td>Study Design</td>
<td>NNS Type Examined</td>
<td>NNS Mode of Delivery</td>
<td>NNS Intake Assessment Method</td>
<td>Body Weight Assessment Method</td>
<td>Study Duration</td>
<td>Weight-Related Outcome Results</td>
<td>Outcomes</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>------------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Obese adult females; 20-60 years $^{52}$</td>
<td>n = 163 (82 aspartame; 81 no-aspartame)</td>
<td>Randomized Controlled Trial</td>
<td>Aspartame</td>
<td>Food and Beverages</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>16 weeks (active weight loss), 71 weeks (weight maintenance), 175 weeks (follow-up)</td>
<td>↓ BW</td>
<td>During weight maintenance (week 71), participants in the aspartame group had an average 3.1% weight regain, compared to 4.9% in the control group (P=0.05). By week 156 (follow-up), participants in the aspartame group had regained an additional 2.4% (net weight loss from baseline of 5.1%) compared with a gain of 5.4% (net weight loss from baseline of 0.3%) in the no-aspartame group (P=0.01).</td>
</tr>
<tr>
<td>Overweight and obese adult females; 18-50 years; habitual NNS consumers $^{53}$</td>
<td>n = 62 (32 NNS beverage; 30 water)</td>
<td>Randomized Controlled Trial</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>24-week weight loss program</td>
<td>↓ BW</td>
<td>Participants receiving the water intervention had a greater decrease in weight (-8.8±1.9kg) compared to the diet intervention (-7.6±2.1kg) (P=0.015).</td>
</tr>
<tr>
<td>Overweight adults; 20-50 years $^{55}$</td>
<td>n = 47 (17 males; 30 females)</td>
<td>Randomized Controlled Trial</td>
<td>Aspartame</td>
<td>Diet Soda</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>26 weeks</td>
<td>↓ SAAT</td>
<td>Regular cola and milk resulted in an increase in the amount of SAAT, while diet soda and water resulted in a reduction in SAAT.</td>
</tr>
<tr>
<td>Overweight and obese adults; 21-65 years $^{57}$</td>
<td>n = 303 (154 NNS group; 149 water group)</td>
<td>Randomized Controlled Trial</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>12-week weight-loss program</td>
<td>↓ BW</td>
<td>The NNS beverage group had significantly higher weight loss (6.95±3.94kg) than the water group (4.09±3.74kg).</td>
</tr>
<tr>
<td>Overweight and obese adults; 21-65 years $^{56}$</td>
<td>n = 303 (154 NNS; 149 water)</td>
<td>Randomized Controlled Trial</td>
<td>Not Specified</td>
<td>Beverages</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>1-year weight maintenance program (after initial 12-week weight loss program)</td>
<td>↓ BW</td>
<td>Subjects receiving water intervention maintained more weight loss (2.45±5.59 kg) while those receiving NNS beverages (6.21±7.65kg) (P&lt;0.001).</td>
</tr>
<tr>
<td>Study Population</td>
<td>Study Sample Size (n)</td>
<td>Study Design</td>
<td>NNS Type Examined</td>
<td>NNS Mode of Delivery</td>
<td>NNS Intake Assessment Method</td>
<td>Body Weight Assessment Method</td>
<td>Study Duration</td>
<td>Weight-Related Outcome Results</td>
<td>Outcomes</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-----------------------</td>
<td>----------------------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>-----------------------------</td>
<td>-------------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Overweight adults; 20-50 years</td>
<td>n = 41</td>
<td>Randomized Controlled Trial</td>
<td>Not Specified</td>
<td>Food and Beverages</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>10 weeks</td>
<td>↓ BW</td>
<td>BW and TBF increased in the sucrose group (1.6kg and 1.3kg) and decreased in the NNS group (1.0kg and 0.3kg).</td>
</tr>
<tr>
<td>(21 sucrose; 20 NNS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↓ TBF</td>
<td></td>
</tr>
<tr>
<td>Normal weight adult females; 20-55 years</td>
<td>n = 133</td>
<td>Randomized Controlled Trial</td>
<td>Aspartame</td>
<td>Diet Soda</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>5 weeks (1-week baseline, 4-week intervention)</td>
<td>↓ BW</td>
<td>Women who consumed the sucrose drink gained some weight during the study and women consumed aspartame lost weight [F(10.20,1.86)=4.509; P&lt;0.05].</td>
</tr>
<tr>
<td>(20-55 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight adult females; 20-55 years</td>
<td>n = 53</td>
<td>Randomized Controlled Trial</td>
<td>Aspartame</td>
<td>Diet Soda</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>5 weeks (1-week baseline, 4-week intervention)</td>
<td>No significant changes</td>
<td>Added sucrose or aspartame beverages did not lead to significant weight.</td>
</tr>
<tr>
<td>(20-55 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight and obese adults; 18-65 years; consuming 280 kcal/day of SSBs</td>
<td>n = 318 (105 = control; 108 = water; 105 = NNS)</td>
<td>Randomized Controlled Trial</td>
<td>Beverages</td>
<td>24-hour dietary recalls</td>
<td>Clinical Measurement</td>
<td>26 weeks</td>
<td>↓ BW</td>
<td>Average percentage weight loss at 3-months was highest in the diet beverage group (-1.87%±0.32) compared to water (-1.31%±0.27), and control groups (-1.34%±0.27). Average percentage weight loss at six months was highest in the diet beverage group (-2.54%±0.45) compared to the water (-2.03%±0.40), and the control group (-1.76%±0.35).</td>
<td></td>
</tr>
<tr>
<td>(20-55 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal-weight adults; aged 22.9±3.7 years</td>
<td>n = 30 (21 males; 9 females; all participants received three interventions)</td>
<td>Randomized Controlled Trial</td>
<td>Aspartame</td>
<td>Diet Soda</td>
<td>Food records</td>
<td>Clinical Measurement</td>
<td>3 3-week intervention periods</td>
<td>↓ BW (males)</td>
<td>Female participants consuming HFCS-sweetened soda gained weight significantly (0.97±0.25kg, p&lt;0.01) and males gained slightly (0.52±0.23kg). Female participants consuming APM-sweetened soda, females gained weight slightly (0.25±0.29kg, p&lt;0.05) but males lost weight significantly (0.47±0.22kg, p&lt;0.05).</td>
</tr>
</tbody>
</table>

Note: NNS = Non-Nutritive Sweetener, BMI = Body Mass Index (kg/m2), SSB = Sugar-sweetened beverage, NHANES = National Health and Nutrition Examination Survey, NWCR = National Weight Control Registry, ENRICA = Estudio de Nutrición y Riesgo cardiovascular en España.
### Table 3. Study design element frequencies with weight outcomes

<table>
<thead>
<tr>
<th>Study Design Type</th>
<th># Studies</th>
<th>↑ Weight n (%)</th>
<th>↓ Weight n (%)</th>
<th>No change n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observational; Cross-Sectional</td>
<td>14 (34%)</td>
<td>11 (79%)</td>
<td>2 (14%)</td>
<td>1 (7%)</td>
</tr>
<tr>
<td>Observational; Longitudinal</td>
<td>11 (27%)</td>
<td>8 (73%)</td>
<td>3 (27%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Randomized Controlled Trial</td>
<td>16 (39%)</td>
<td>0 (0%)</td>
<td>13 (81%)</td>
<td>3 (19%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of NNS Examined</th>
<th># Studies</th>
<th>↑ Weight</th>
<th>↓ Weight</th>
<th>No change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartame</td>
<td>9 (22%)</td>
<td>1 (11%)</td>
<td>6 (67%)</td>
<td>2 (22%)</td>
</tr>
<tr>
<td>Saccharin</td>
<td>2 (4%)</td>
<td>2 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Stevia</td>
<td>1 (2%)</td>
<td>0 (0%)</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Multiple Individual Types</td>
<td>2 (4%)</td>
<td>2 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Not Specified</td>
<td>27 (66%)</td>
<td>14 (52%)</td>
<td>11 (41%)</td>
<td>2 (7%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NNS Mode of Delivery</th>
<th># Studies</th>
<th>↑ Weight</th>
<th>↓ Weight</th>
<th>No change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beverages</td>
<td>24 (59%)</td>
<td>10 (42%)</td>
<td>12 (50%)</td>
<td>2 (8%)</td>
</tr>
<tr>
<td>Other Diet Beverages</td>
<td>15 (37%)</td>
<td>6 (40%)</td>
<td>9 (60%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Diet Soda</td>
<td>8 (20%)</td>
<td>3 (38%)</td>
<td>3 (38%)</td>
<td>2 (25%)</td>
</tr>
<tr>
<td>Coffee/Tea</td>
<td>1 (2%)</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Food and Beverages</td>
<td>13 (32%)</td>
<td>5 (39%)</td>
<td>6 (46%)</td>
<td>2 (15%)</td>
</tr>
<tr>
<td>Not Specified</td>
<td>4 (9%)</td>
<td>4 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NNS Intake Assessment Method</th>
<th># Studies</th>
<th>↑ Weight</th>
<th>↓ Weight</th>
<th>No change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveys and Questionnaires</td>
<td>5 (12%)</td>
<td>4 (80%)</td>
<td>1 (20%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Food Frequency Questionnaires (FFQ)</td>
<td>12 (29%)</td>
<td>7 (58%)</td>
<td>3 (25%)</td>
<td>1 (8%)</td>
</tr>
<tr>
<td>24-hour dietary recalls</td>
<td>7 (17%)</td>
<td>5 (71%)</td>
<td>2 (29%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Food records</td>
<td>13 (32%)</td>
<td>1 (8%)</td>
<td>10 (77%)</td>
<td>2 (15%)</td>
</tr>
<tr>
<td>Combination</td>
<td>2 (4%)</td>
<td>1 (50%)</td>
<td>1 (50%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>None Included</td>
<td>2 (4%)</td>
<td>0 (0%)</td>
<td>1 (50%)</td>
<td>1 (50%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Body Weight Assessment Method</th>
<th># Studies</th>
<th>↑ Weight</th>
<th>↓ Weight</th>
<th>No change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Measurement</td>
<td>34 (83%)</td>
<td>14 (41%)</td>
<td>16 (47%)</td>
<td>4 (12%)</td>
</tr>
<tr>
<td>Self-Reported</td>
<td>7 (17%)</td>
<td>5 (71%)</td>
<td>2 (41%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>
Table 4. Weight outcome frequencies based on study design types and methods

<table>
<thead>
<tr>
<th></th>
<th>Observational\textsuperscript{a,b}</th>
<th>Randomized Controlled Trial\textsuperscript{c}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weight (\uparrow) n (%)</td>
<td>Weight (\downarrow) n (%)</td>
</tr>
<tr>
<td><strong>NNS Type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specified</td>
<td>5 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Not Specified</td>
<td>14 (74%)</td>
<td>5 (26%)</td>
</tr>
<tr>
<td><strong>NNS Mode of Delivery</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beverages</td>
<td>10 (77%)</td>
<td>3 (23%)</td>
</tr>
<tr>
<td>Food and Beverages</td>
<td>5 (71%)</td>
<td>2 (29%)</td>
</tr>
<tr>
<td>Not Specified</td>
<td>4 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td><strong>NNS Assessment Method</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surveys/FFQ</td>
<td>12 (75%)</td>
<td>4 (25%)</td>
</tr>
<tr>
<td>Recalls/Food records</td>
<td>6 (86%)</td>
<td>1 (14%)</td>
</tr>
<tr>
<td>Combination</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>None included</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

\textbf{Note:} \(\uparrow\) indicates higher weight-related outcomes; \(\downarrow\) indicates lower weight-related outcomes
\textsuperscript{a}Observational studies include both cross-sectional and longitudinal studies
\textsuperscript{b}One observational study found no significant weight changes. No longitudinal studies found no significant weight changes.
\textsuperscript{c}Three randomized controlled trials found no significant weight changes.
**Study Design**

**Observational**

The articles identified included 25 observational studies, 14 of which were cross-sectional studies\textsuperscript{14,23-35} and 11 of which were longitudinal\textsuperscript{36-46}. Among the cross-sectional studies, 11 of 14 (79\%) reported higher weight-related outcomes for NNS consumers. Five studies\textsuperscript{23,27,29,31,33} associated higher NNS intake with higher BMI measurements. For example, one study found that heavy NNS beverage consumers (>825 ml/day) reported significantly higher BMI compared to non-NNS users\textsuperscript{23}. Only two cross-sectional studies reported lower weight outcomes\textsuperscript{30,32}. The first, an investigation among college-aged females, found that normal weight and overweight women were more likely than obese women to report consuming NNS and sugar-free versions of products\textsuperscript{30}. Another study found that participants who were successful in maintaining weight loss reported consuming three times as many NNS beverages than normal weight individuals\textsuperscript{32}. Four of these cross-sectional studies relied upon National Health and Nutrition Examination Survey (NHANES) data from various time periods\textsuperscript{25,26,33,35}.

The 11 longitudinal studies\textsuperscript{36-46} ranged in length from 3 years to 28 years. The majority of these (8 of 11 studies) found associations between NNS use and weight gain over time while the remaining 3 longitudinal studies reported lower weight-related outcomes. Among those finding higher weight-related outcomes, two studies found positive associations between NNS consumption and BMI and waist circumference\textsuperscript{36,40}. Conversely, in a four-year study among females, less weight gain was reported among participants who reporting increasing their diet beverage intake\textsuperscript{45}. Another study comparing NNS to SSB found that females who reported consuming NNS beverages
found that females who reported consuming NNS beverages gained less weight over time than those drinking SSB\textsuperscript{46}.

**Randomized Controlled Trials**

Of the 16 randomized controlled trials\textsuperscript{47-62}, the majority of NNS interventions (13 of 16) resulted in improved weight-related outcomes. None of the studies found higher weight-related outcomes and three showed NNS intake did not make a significant difference in weight-related outcomes.

Among those with improved weight outcomes, one study enrolled overweight and obese adults in a RCT where participants were given either water or NNS beverages to replace habitual sugar-sweetened beverage consumption as part of a weight loss program. Authors found average percentage weight loss at three months was greatest in the diet beverage group compared to the water and the control group.\textsuperscript{61} Another study on overweight adults who were instructed to consume either sucrose-sweetened or NNS foods and beverages over the course of a ten-week intervention period, found that body weight and body fat percentile increased in the sucrose group and decreased in the NNS group\textsuperscript{58}. A RCT using a 12-week weight loss program and a one-year weight maintenance program found that participants in a NNS beverage group (group consuming >24 fluid ounces of NNS beverages per day) lost more weight and maintained greater weight loss than those advised to consume water (consuming >24 fluid ounces of water per day) as part of their weight loss intervention\textsuperscript{56,57}.

In a study examining the impact of beverages sweetened with high-fructose corn syrup compared to beverages sweetened with aspartame in normal weight adults,
investigators found conflicting results, with males drinking aspartame-sweetened beverages losing a significant amount of weight and females gaining weight, although it was not a significant change\textsuperscript{62}. In another study comparing interventions using NNS beverages and water in a 24-week weight loss program, findings indicated that water was more effective than NNS beverages in improving weight loss outcomes; however, the NNS intervention was still associated with improved weight-related outcomes\textsuperscript{53}.

Among the studies that did not find significant weight-related changes, one study found that NNS interventions did not have significant impacts on weight outcomes of participants compared to low-fat interventions over the course of a four-week intervention\textsuperscript{49}. Two studies by Reid et al. in 2007\textsuperscript{60} and 2010\textsuperscript{59}, examined the impact of aspartame-sweetened diet soda over the course of a four-week intervention. The 2010 study, including overweight women, did not find that aspartame was associated with weight change, while the 2007 study, including normal weight women, found that aspartame was marginally associated with greater weight loss.

\textit{Type of NNS Examined}

The majority of studies in this review did not specify the NNS type (66\%), but rather looked at NNS as a whole (Table 3). Aspartame was studied individually more than any other NNS type, with nine of the studies examining the relationship between aspartame consumption and weight-related outcomes\textsuperscript{35,48,50-52,55,59,60,62}. Among those not specifying NNS type, 74\% of observational studies (14 of 19 studies) reported higher weight-related outcomes compared to 100\% of RCT (6 studies) reporting improved weight outcomes (Table 4). RCT were more likely to specify the type of NNS being
examined, with 7 out of 13 (54%) of RCT specified NNS type. Moreover, 100% of RCT (7 studies) specifying NNS type found lower weight-related outcomes (Table 4).

Among studies that looked at other individual NNS types, two observational studies in this review found associations between saccharin and weight gain. Specifically, one study using 8 years of data from the Nurses’ Health Study on adult females found that higher saccharin consumers had higher body weights compared to non-consumers\textsuperscript{37}. Similarly, among participants in the four-year Pawtucket Hearth Health Program, investigators found that saccharin intake was associated with higher weight status\textsuperscript{42}. One study investigated the impact of stevia on weight status, finding that there was a small, but not significant, reduction in body weight and BMI after a one-week stevia intervention\textsuperscript{47}.

Two observational studies analyzed consumption of multiple individual types of NNS. One of these studies analyzed participants diets based on the type of NNS, including sucralose, aspartame, acesulfame potassium, and saccharin, and correlated total NNS intake with weight status, finding that NNS consumers were more likely to have higher body weights and BMI\textsuperscript{14}. Another study also collected data on a combination of NNS, finding that total NNS intake as well as saccharin, sucralose, and acesulfame potassium were associated with a higher waist circumference measurements\textsuperscript{28} and additional analysis was conducted based on NNS type.

As previously mentioned, RCT were more likely to specify the NNS types within the study methods. In a RCT among obese females participating in a weight loss program, researchers found that, those who received aspartame as part of their intervention lost significantly more weight and better maintained their weight loss than
those who did not supplement their weight loss program with aspartame\cite{48}. A 1994 study by Kanders et al.,\cite{51} also found that participants regained less weight during weight maintenance when utilizing aspartame as part of their diet. With the exception of two studies that did not find significant weight changes, all other RCT studies on aspartame found reduced weight-related outcomes.

**NNS Mode of Delivery**

Regardless of study design type, 59\% of studies (24 of 41 studies) examined NNS beverages consumption and relation to weight-related outcomes. Of these, 8 studies looked specifically at diet soda consumption and one looked at coffee and tea with NNS additives. Among those using beverages as a proxy of NNS intake, 77\% of observational studies (10 of 13 studies) found higher weight-related outcomes and 100\% of RCT (9 studies) found lower weight-related outcomes (Table 4). When looking at studies that included foods and beverages in their analysis, 71\% of observational studies (5 of 7 studies) found higher weight-related outcomes, compared to 100\% of RCT (4 studies) finding lower weight-related outcomes (Table 4).

Among the observational studies included, one of the cross-sectional studies found that NNS use in coffee and/or tea was associated with higher BMI and waist circumference\cite{26}. Another observational study found that those in the overweight and obese category were significantly more likely to report consumption of diet sodas with 2.6\% of normal weight, 21.1\% of overweight, and 35.1\% of obese individuals reporting habitual diet soda consumption; however, this association was not present between BMI
and all NNS beverages\textsuperscript{27}. Another observational study found that overweight and obese adults consume more diet beverages compared to normal weight adults\textsuperscript{25}.

When looking at RCT included in this review, one study using a diet soda intervention found that diet soda consumption resulted in reduced subcutaneous abdominal adipose tissue among overweight adults\textsuperscript{55}. Participants in another RCT were given aspartame-sweetened pudding or milk shakes, as well as instructions to consume NNS sweeteners, beverages, and gelatins, as part of a weight loss program. Findings revealed that females, but not males, who consumed aspartame lost more weight compared to those not receiving an aspartame intervention\textsuperscript{51}. Additionally, four studies\textsuperscript{35,37,42,44} did not specify if NNS in the form of food or beverage was measured.

\textit{NNS Intake Assessment Method}

This investigation found that the most commonly used dietary assessment methods for measuring NNS intake were food records (used by 32\% of studies) and food frequency questionnaires (FFQ) (used by 29\% of studies). The observational studies used a variety of assessment methods, most often utilizing FFQ and surveys. Among observational studies that used surveys or FFQ to measure NNS intake, 75\% (12 of 16 studies) found higher weight-related outcomes (Table 4). The majority of the RCT (85\%, 11 of 13 studies) utilized food records or 24-hours dietary recalls to measure NNS intake, with 100\% (11 studies) of these RCT finding lower weight-related outcomes (Table 4).

Among the observational studies, 12 used FFQ to gauge habitual NNS consumption and other dietary patterns. FFQ were more frequently used in studies involving more participants and for a longer period of time. Two studies analyzed subsets
of data from the Nurses’ Health Study I, the Nurses’ Health Study II, and the Health Professionals Follow-Up Study. One study exploring associations between NNS beverages and weight status found that each one-cup increase in diet beverage intake was inversely associated with weight gain over time. Another study on diet soda intake determined that diet soda was associated with increased weight in both prevalence and lagged change analysis. The studies by Hedrick et al. and Hess et al., which both reported on a variety of NNS types, used 24-hour dietary recalls, which typically allow researchers to gather more detailed dietary information, including specific amounts of NNS consumed. A study using both 24-hour dietary recalls and survey questions to determine NNS intake found that BMI increases were greater among NNS users than non-users.

As mentioned, RCT were more likely to use dietary recalls and food records to assess NNS intake. In one intervention study using food records, investigators found that regular soda and milk consumption resulted in an increase in subcutaneous abdominal adipose tissue, compared to decreased subcutaneous abdominal adipose tissue among participants consuming diet soda and water. In an investigation using both FFQ and 3-day food records to measure intake during a 12-week intervention, researchers discovered that females who consumed aspartame as part of their weight loss intervention lost more weight than those who did not consume aspartame, although there were no significant differences among males in the same study.
Body Weight Assessment Method

Seven of the studies included in this review, all of which were observational designs, used self-reported body weight measurements in their outcomes. Among these studies, 5 reported higher weight-related outcomes and 2 reported lower weight-related outcomes (Table 3). Using self-reported height and weight data, one study found that habitual consumers of NNS beverages reported significantly higher BMI levels than habitual non-consumers of NNS beverages\(^{23}\). Another study using self-reported weight data from the National Weight Control Registry found that those in the overweight and obese category were significantly more likely to report consumption of diet sodas\(^{27}\). Data from the Nurses’ Health Study, the Nurses’ Health Study II and the Health Professionals Follow-Up Study provided self-reported weight data for a study finding that diet soda was associated with an increased body weight\(^{43}\). Authors explained that because study participants were all health professionals, their self-reported weight measurements were likely accurate. The remaining studies used clinical measurements to assess body weight and other weight-related outcomes.

Conclusions

This review demonstrates how study design and method usage varies in NNS studies in the literature and how research conclusions can differ as a result, revealing that RCT are more likely to use study design methods with more specificity. Of the 16 randomized controlled trials, 81% indicated NNS interventions resulted in improved weight outcomes and 0% demonstrated higher weight outcomes as compared to 76% of observational studies finding higher weight outcomes. These findings indicate the
association between RCT study design and improved weight-related outcomes are likely to be more reliable in determining causation than the association between observational study design and higher weight-related outcomes. As researchers move forward in their investigations of NNS and weight-related outcomes, challenges to study design (including reverse causality, starting weight status, and addressing missing data) and assessment methodology (including NNS types examined, and NNS types and body weight assessment methods) must be addressed.

One of the major risks of relying upon observational studies to understand the relationship between NNS consumption and weight status is the potential bias of reverse causality. In observational studies on NNS consumption, overweight and obese participants are frequently selected as study participants. While NNS consumption may be associated with a higher weight status, it is difficult to determine if NNS is chosen due to their weight status or if their higher weight status is caused by NNS consumption. Using RCT study designs in studies of NNS on weight-related outcomes allows researchers to control for potential reverse causality.

Beyond concerns about reverse causality, comparing results across these studies, particularly in meta-analyses, is difficult due to the population of participants being not normally distributed in terms of body weight and BMI. Among the studies included in this review, 11 RCT specifically recruited participants who were overweight or obese at the start of the study to participant in weight loss interventions and 5 recruited normal weight adults. In addition, RCT vary in whether NNS are being used to replace added sugars or simply added to participants diets and whether they are being compared to water or SSB. As a letter by Sievenpiper et al., suggested, the impact of energy
displacement may confound these studies without determining the mechanisms by which weight status is changed.

Currently, the preferred method for quantifying NNS intake is multiple dietary recalls or food records\textsuperscript{64}, which are then analyzed with dietary analysis software (Nutrition Data System for Research software [NDS-R])\textsuperscript{65}. Dietary assessment is especially challenging for NNS due to the fact that the FDA does not require food manufacturers to report specific amounts of NNS used in products\textsuperscript{15}. With new products including different combinations and amounts of NNS frequently being introduced into the food supply, it is difficult for software to stay current with new products being released and for consumers to identify NNS products they consume. As a result, researchers have difficulty quantifying NNS intake and are often limited to studying NNS as a single category, rather than as individual compounds. However, NNS are each distinct chemical compounds with differing metabolic pathways when consumed\textsuperscript{66}. Because they are each metabolized differently\textsuperscript{67-69}, with some being broken down by the body while other are excreted in the urine or feces, they likely have unique impacts on energy balance and, ultimately, weight outcomes. It may be less challenging for participants to report and researchers to measure NNS as a comprehensive group; however, it is necessary in order to determine if each distinct type of NNS has differential impacts on weight-related outcomes. Novel assessment methods, such as a newly validated NNS-FFQ, may help mitigate the challenges of measuring habitual intake of individual NNS types\textsuperscript{70}.

The currently available research on individual NNS types lacks information on NNS that are commonly used. Based on an observational study in southwestern Virginia
by Hedrick et al.,\textsuperscript{14} sucralose is the most commonly consumed NNS; however, according to this review, aspartame has been studied individually more than other types of sweeteners. The popularity of stevia, a naturally-derived NNS, has been rising steadily since stevia’s introduction into the US market in 2008, with stevia’s market share rising from 5\% to 15\% between 2009 and 2013\textsuperscript{71}. Only one study in this review utilized stevia as part of its dietary intervention\textsuperscript{47}.

The lack of consistency in the mode of delivery of NNS among studies make it difficult to compare results across studies. Because foods and beverages differ in their effects on satiety\textsuperscript{72,73}, it may be more useful to design studies that first use pure forms of NNS to determine if the substances themselves have an impact on weight-related outcomes. Subsequent studies should study separate NNS foods and NNS beverages in interventions to distinguish between effects of the NNS and satiety. While consumers are more likely to consume NNS in the form of beverages\textsuperscript{74}, because NNS are being used in an increasing number of food products\textsuperscript{2}, it is important to know the impact that NNS have on satiety and ultimately the impact this has on consumers’ weight status. The common use of diet soda as a representation of total NNS intake for NNS consumption quantification poses additional challenges. Diet sodas utilize a variety of types and amounts of NNS in each diet soda (e.g., aspartame with acesulfame potassium, or sucralose with acesulfame potassium) making it difficult to compare across diet sodas. Moreover, using diet soda alone to measure NNS intake excludes individuals who consume NNS in other food and beverage sources. In one cross-sectional study, authors found that if diet soda alone were used to measure NNS intake, only 22\% of participants
were classified as NNS consumers, compared to 50% if all NNS food and beverage sources were included in the analysis\textsuperscript{28}.

A limitation due to self-reported data is the likelihood of reporting bias when measuring weight-related outcomes. In previous analysis, Briefel et al.,\textsuperscript{75} found under-reporting of energy intake to be more common among overweight and obese populations. Lin et al.,\textsuperscript{76} determined that obese women were more likely to under-report their weight and underweight women were more likely to over-report their weight. Utilizing self-reported dietary intake and weight assessment methods are useful in large-scale, epidemiological studies; however, there are limitations associated with these methods of data collection and efforts should be made to use more specific and objective measurements when it is feasible.

In order to better understand the impact that NNS have on consumers’ weight status, researchers must ultimately determine the mechanisms by which NNS affect weight status. To achieve this, controlled-feeding studies should be conducted, rather than measuring ad libitum diet. Controlled-feeding studies allow researchers to reduce potential confounding variables, such as compensatory appetite, from altering outcomes. The negative energy balance induced by the lack of caloric content in NNS is the most obvious way that NNS could impact weight status; however, future investigations must explore other possible causal mechanisms impacting the relationship between NNS and weight-related outcomes, such as insulin response, gut microbiota, taste perceptions, or changes to other dietary intake patterns.

This systematic review confirmed previous findings that observational studies were more likely to report higher weight-related outcomes and RCT were more likely to
report lower weight-related outcomes. Expanding upon previous reviews conducted, study methods, including the type of NNS measured, NNS mode of delivery, and NNS intake and body weight assessment methods, were analyzed, revealing that RCT were more likely to use study design methods with more specificity. These findings indicate the association between RCT study design and lower weight-related outcomes are likely to be more reliable in determining causation, than the association between observational study design and higher weight-related outcomes.

**Future Directions**

A number of study design and method recommendations can be made based on the findings of this systemic review. Most importantly, RCT should be implemented, rather than observational study designs. More specifically, controlled feeding studies are needed to better understand the impact of NNS on participants’ weight status while controlling for other dietary factors. As RCT studies are not always feasible, the following recommendations are applicable for RCT and observational studies alike:

1) Specify individual NNS types, preferably focusing on those most frequently consumed NNS types (e.g., sucralose and stevia), to be used in interventions.

2) Select the dietary source of NNS based on the specific aim of the study (e.g., satiety from food versus beverage) and address potential influence of delivery mode. Studies should refrain from using diet soda as a proxy for assessing NNS intake and instead use all foods and beverages.
3) Use specific and objective measurements to assess NNS dietary intake. Novel dietary assessment methods need to be developed to better assess NNS consumption (e.g., food frequency questionnaires and dietary biomarkers).

4) Explore other causal mechanisms impacting the relationship between NNS and weight-related outcomes, such as the impact of NNS on insulin response, gut microbiota, taste perceptions, or changes to other dietary intake patterns.
References


10. Hu FB. Resolved: there is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases. *Obes Rev.* 2013;14(8):606-619.


Appendix A. Institutional Review Board Protocol Approvals

MEMORANDUM

DATE: September 25, 2015
TO: Valisa Ellen Hedrick, Emily Myers
FROM: Virginia Tech Institutional Review Board (FWA00000572, expires July 29, 2020)

PROTOCOL TITLE: Validation of New Dietary Assessment Methods
IRB NUMBER: 15-682

Effective September 25, 2015, the Virginia Tech Institution Review Board (IRB) Chair, David M Moore, approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As: Expedited, under 45 CFR 46.110 category(ies) 2,4,7
Protocol Approval Date: September 25, 2015
Protocol Expiration Date: September 24, 2016
Continuing Review Due Date*: September 10, 2016

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.
MEMORANDUM

DATE: September 1, 2016

TO: Valisa Ellen Hedrick, Emily Myers, Kevin Davy, Jose Manuel Rivero

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)

PROTOCOL TITLE: Validation of New Dietary Assessment Methods

IRB NUMBER: 15-682

Effective August 31, 2016, the Virginia Tech Institution Review Board (IRB) Chair, David M Moore, approved the Continuing Review request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at: http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As: Expedited, under 45 CFR 46.110 category(ies) 2,4,7
Protocol Approval Date: September 25, 2016
Protocol Expiration Date: September 24, 2017
Continuing Review Due Date*: September 10, 2017

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal/work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.
MEMORANDUM

DATE: September 1, 2016

TO: Valisa Ellen Hedrick, Emily Myers, Kevin Davy, Jose Manuel Rivero

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)

PROTOCOL TITLE: Validation of New Dietary Assessment Methods

IRB NUMBER: 15-682

Effective August 31, 2016, the Virginia Tech Institutional Review Board (IRB) Chair, David M Moore, approved the Amendment request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As: Expedited, under 45 CFR 46.110 category(ies) 2, 4, 7

Protocol Approval Date: September 25, 2015

Protocol Expiration Date: September 24, 2016

Continuing Review Due Date*: September 10, 2016

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.
MEMORANDUM

DATE: August 31, 2017

TO: Valisa Ellen Hedrick, Emily Myers, Kevin Davy, Jose Manuel Rivero

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)

PROTOCOL TITLE: Validation of New Dietary Assessment Methods

IRB NUMBER: 15-682

Effective August 30, 2017, the Virginia Tech Institution Review Board (IRB) Chair, David M Moore, approved the Continuing Review request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at: http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As: Expedited, under 45 CFR 46.110 category(ies) 2,4,7
Protocol Approval Date: September 25, 2017
Protocol Expiration Date: September 24, 2018
Continuing Review Due Date*: September 10, 2018

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal/work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution
<table>
<thead>
<tr>
<th>Date*</th>
<th>OSP Number</th>
<th>Sponsor</th>
<th>Grant Comparison Conducted?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this IRB protocol is to cover any other grant proposals, please contact the IRB office (irbadmin@vt.edu) immediately.
MEMORANDUM

DATE: July 19, 2018

TO: Susan E Duncan, Emily Myers, Ali Wang

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)

PROTOCOL TITLE: Sweetener sensory effects in two delivery systems

IRB NUMBER: 18-063

The Virginia Tech Institution Review Board (IRB), acknowledges the Amendment request for the above-mentioned research protocol.

This acknowledgement recognizes the item(s) identified in the Special Instructions section.

NOTE: Please ensure that required Amendments are submitted to WIRB for review and approval. WIRB guidance is provided on page 49 of the Guide for Researchers. The section is titled Changes to Research / Additional Document Submissions. The document is located at: http://wirb.com/Documents/Guide%20for%20Researchers.pdf#page=2
IRB SPECIAL INSTRUCTIONS:

This Amendment Acknowledgement includes WIRB determination letter, WIRB smart form, and WIRB-approved consent form(s).

<table>
<thead>
<tr>
<th>Date*</th>
<th>OSP Number</th>
<th>Sponsor</th>
<th>Grant Comparison Conducted?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this IRB protocol is to cover any other grant proposals, please contact the IRB office (irbadmin@vt.edu) immediately.
## Appendix B. Risk of bias assessment for observational studies

<p>| Study                  | Research question/objective clearly stated? | Population clearly specified and defined? | Participation rate of eligible persons at least 50%? | Inclusion/exclusion criteria pre-specified and applied uniformly? | Sample size justification, or variance and effect estimates provided? | Exposure(s) of interest measured prior to the outcome(s) being measured? | Sufficient timeframe? | Exposure measures clearly defined, valid, reliable, etc.? | Exposure(s) assessed more than once over time? | Outcome measures clearly defined, valid, reliable, etc.? | Outcome assessors blinded to the exposure status of participants? | Loss to follow-up after baseline 20% or less? | Potential confounding variables measured and adjusted statistically? | Funding Source                                                                 | Reviewer Comments                                                                 |
|------------------------|---------------------------------------------|------------------------------------------|--------------------------------------------------|-----------------------------------------------------------------|---------------------------------------------------------------------|------------------------------------------------------------------------|------------------------|----------------------------------------------------------------|---------------------------------------------------------------|----------------------------------------------------------------|--------------------------------------------------------------------------|-----------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|
| Appleton et al, 2001  | Ok                                          | Ok                                       | N/A                                               | Unclear                                                        | N/A                                                                | Ok                                                                     | N/A                    | Ok                                                             | Unclear                                                        | N/A                                                            | Ok                                                                 | Unclear                                                      | Ok                                                                 | School of Psychology, University of Leeds                                   | Moderate NNS Consumers excluded; Self-Reported BW and height used for BMI  |
| Bellisle et al, 2001  | Ok                                          | Ok                                       | N/A                                               | Unclear                                                        | N/A                                                                | Ok                                                                     | N/A                    | Ok                                                             | Unclear                                                        | N/A                                                            | Ok                                                                 | Unclear                                                      | Ok                                                                 | No specific funding source                                                                                             | Longitudinal data, but only cross-sectional analysis used            |
| Bleich et al, 2014    | Ok                                          | Ok                                       | N/A                                               | Unclear                                                        | N/A                                                                | Ok                                                                     | N/A                    | Ok                                                             | N/A                                                            | Ok                                                             | N/A                                                               | Unclear                                                      | Ok                                                                 | National Heart, Lung, and Blood Institute                              |
| Bouchard et al, 2010  | Ok                                          | Ok                                       | N/A                                               | Unclear                                                        | N/A                                                                | Ok                                                                     | N/A                    | Ok                                                             | N/A                                                            | Ok                                                             | N/A                                                               | Unclear                                                      | Ok                                                                 | No specific funding source                                                                                             |
| Catenacci et al, 2014 | Ok                                          | Ok                                       | N/A                                               | Unclear                                                        | N/A                                                                | Ok                                                                     | N/A                    | Ok                                                             | N/A                                                            | Ok                                                             | N/A                                                               | Unclear                                                      | Ok                                                                 | Partial funding for this study was from an unrestricted gift from the Coca Cola Company to the University of Colorado Foundation |
| Hedrick et al, 2017   | Ok                                          | Ok                                       | N/A                                               | Unclear                                                        | Ok                                                                  | N/A                                                                    | Ok                     | Unclear                                                        | N/A                                                            | Unclear                                                        | N/A                                                               | Unclear                                                      | National Institutes of Health and the American Heart Association                                |                                                                 |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Research question/objective clearly stated?</th>
<th>Population clearly specified and defined?</th>
<th>Participation rate of eligible persons at least 50%?</th>
<th>Inclusion/exclusion criteria pre-specified and applied uniformly?</th>
<th>Sample size justification, power description, or variance and effect estimates provided?</th>
<th>Exposure(s) measured prior to the outcome(s) being measured?</th>
<th>Different levels of the exposure measured for continuous outcomes?</th>
<th>Exposure measures clearly defined, valid, reliable, etc.?</th>
<th>Exposure(s) assessed more than once over time?</th>
<th>Outcome measures clearly defined, valid, reliable, etc.?</th>
<th>Outcome assessors blinded to the exposure status of participants?</th>
<th>Loss to follow-up after baseline 20% or less?</th>
<th>Potential confounding variables measured and adjusted statistically?</th>
<th>Funding Source</th>
<th>Reviewer Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hess et al, 2018</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>USDA National Institute of Food and Agriculture</td>
<td>Ok</td>
<td></td>
</tr>
<tr>
<td>Lana et al, 2015</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>No</td>
<td>Sanofi-Aventis and Spanish Ministry of Health FIS grants</td>
<td>Ok</td>
<td></td>
</tr>
<tr>
<td>Malinauskas et al, 2006</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>Ok</td>
<td>N/A</td>
<td>Unclear</td>
<td>Yes</td>
<td>N/A</td>
<td>No</td>
<td>Undergrad research assistantships, East Carolina University Honors Program</td>
<td>Unclear</td>
<td></td>
</tr>
<tr>
<td>Odegaard et al, 2012</td>
<td>Ok</td>
<td>N/A</td>
<td>Unclear</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>N/A</td>
<td>Unclear</td>
<td>N/A</td>
<td>No</td>
<td>N/A</td>
<td>Yes</td>
<td>No specific funding source</td>
<td>Ok</td>
<td></td>
</tr>
<tr>
<td>Phelan et al, 2009</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>N/A</td>
<td>Unclear</td>
<td>N/A</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
<td>National Institutes of Health and American Beverage Corporation</td>
<td>Ok</td>
<td></td>
</tr>
<tr>
<td>Sylvetsky et al, 2017</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>N/A</td>
<td>Unclear</td>
<td>N/A</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
<td>Department of Exercise and Nutrition Sciences at the George Washington University and Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases</td>
<td>Ok</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Research question/objective clearly stated?</td>
<td>Population clearly specified and defined?</td>
<td>Participation rate of eligible persons at least 50%?</td>
<td>Sample size justification, power description, or variance and effect estimates provided?</td>
<td>Sufficient timeframe?</td>
<td>Different levels of the exposure measured for continuous outcomes?</td>
<td>Exposure measures clearly defined, valid, reliable, etc.?</td>
<td>Exposure(s) assessed more than once over time?</td>
<td>Outcome measures clearly defined, valid, reliable, etc.?</td>
<td>Outcome assessors blinded to the exposure status of participants?</td>
<td>Loss to follow-up after baseline 20% or less?</td>
<td>Potential confounding variables measured and adjusted statistically?</td>
<td>Funding Source</td>
<td>Reviewer Comments</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------------------------</td>
<td>----------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td></td>
</tr>
<tr>
<td>Winther et al, 2017</td>
<td>Ok</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>N/A</td>
</tr>
<tr>
<td>Wulaningsih et al, 2017</td>
<td>Ok</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>Ok</td>
<td>N/A</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>N/A</td>
</tr>
<tr>
<td>Chia et al, 2016</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>N/A</td>
</tr>
<tr>
<td>Colditz et al, 1990</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>National Institutes of Health</td>
</tr>
<tr>
<td>Fowler et al, 2008</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>National Institutes of Health and the United States Department of Agriculture</td>
</tr>
<tr>
<td>Fowler et al, 2015</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>National Institute on Aging, National Institute of Diabetes and Digestive and Kidney Diseases, and National Center for</td>
</tr>
<tr>
<td>Study</td>
<td>Research question/objective clearly stated?</td>
<td>Population clearly specified and defined?</td>
<td>Participation rate of eligible persons at least 50%?</td>
<td>Inclusion/exclusion criteria pre-specified and applied uniformly?</td>
<td>Sample size justification, power description, or variance and effect estimates provided?</td>
<td>Exposure(s) of interest measured prior to the outcome(s) being measured?</td>
<td>Different levels of the exposure measured for continuous outcomes?</td>
<td>Exposure measures clearly defined, valid, reliable, etc.?</td>
<td>Exposure(s) assessed more than once over time?</td>
<td>Outcome measures clearly defined, valid, reliable, etc.?</td>
<td>Outcome assessors blinded to the exposure status of participants?</td>
<td>Loss to follow-up after baseline 20% or less?</td>
<td>Potential confounding variables measured and adjusted statistically?</td>
<td>Funding Source</td>
<td>Reviewer Comments</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------------------------</td>
<td>------------------------------------------</td>
<td>----------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Nettleton et al, 2009</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>Research Resources, and CTSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pan et al, 2013</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>National Institutes of Health</td>
<td>Carried forward missing data points for continuous variables and created a missing categorical variable</td>
<td></td>
</tr>
<tr>
<td>Parker et al, 1997</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>National Heart, Lung, and Blood Institute</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schulze et al, 2004</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
<td>National Institutes of Health, European Association for the Study of Diabetes/ American Diabetes Association Trans-Atlantic fellowship, fellowship of the German Academic Exchange Service (DAAD), and American Heart Association Established Investigator Award</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Research question/objective clearly stated?</td>
<td>Population clearly specified and defined?</td>
<td>Participation rate of eligible persons at least 50%?</td>
<td>Inclusion/exclusion criteria pre-specified and applied uniformly?</td>
<td>Sample size justification, power description, or variance and effect estimates provided?</td>
<td>Exposure(s) of interest measured prior to the outcome(s) being measured?</td>
<td>Sufficient timeframe?</td>
<td>Different levels of the exposure measured for continuous outcomes?</td>
<td>Exposure measures clearly defined, valid, reliable, etc.?</td>
<td>Exposure(s) assessed more than once over time?</td>
<td>Outcome measures clearly defined, valid, reliable, etc.?</td>
<td>Outcome assessors blinded to the exposure status of participants?</td>
<td>Loss to follow-up after baseline 20% or less?</td>
<td>Potential confounding variables measured and adjusted statistically?</td>
<td>Funding Source</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------------------</td>
<td>-------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>--------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>--------------------------------------------------------------</td>
<td>--------------------------------------------------------------</td>
<td>--------------------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Smith et al, 2015</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
</tr>
<tr>
<td>Stellman et al, 1986</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
</tr>
<tr>
<td>Tucker et al, 2015</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>N/A</td>
<td>Ok</td>
</tr>
</tbody>
</table>

Note: *Using NIH Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies*
### Appendix C. Risk of bias assessment for randomized controlled trials

<table>
<thead>
<tr>
<th>Study</th>
<th>Selection Bias Random Sequence Generation</th>
<th>Selection Bias Allocation Concealment</th>
<th>Reporting Bias Selective Reporting</th>
<th>Performance Bias Blinding (participants and personnel)</th>
<th>Detection Bias Blinding (outcomes assessment)</th>
<th>Attrition Bias Incomplete Outcome Data</th>
<th>Other Bias</th>
<th>Funding Source</th>
<th>Reviewer Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al-Dujaili et al, 2017</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>None</td>
<td>No specific funding source</td>
<td>Dietary assessment method not included to measure food/beverage intake.</td>
</tr>
<tr>
<td>Blackburn et al, 1997</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
<td>None</td>
<td>NutraSweet Company and the Center for Nutritional Research Charitable Trust</td>
<td>Participants were not blinded to treatment group due to the group counseling element of the interventions. No evidence of blinded assessors.</td>
</tr>
<tr>
<td>Gatenby et al, 1997</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>Competitive Strategic grant from the Biotechnology and Biological Sciences Research Council</td>
<td>Participants were not blinded to treatment group; no mention of randomization or blinding; Participants who misreported intake were excluded from analysis.</td>
</tr>
<tr>
<td>Higgins et al, 2017</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>Ajinomoto Co., Inc. and Clinical and Translational Sciences Institute</td>
<td>Dietary assessment method not included to measure food/beverage intake.</td>
</tr>
<tr>
<td>Study</td>
<td>Selection Bias Random Sequence Generation</td>
<td>Selection Bias Allocation Concealment</td>
<td>Reporting Bias Selective Reporting</td>
<td>Performance Bias Blinding (participants and personnel)</td>
<td>Detection Bias Blinding (outcomes assessment)</td>
<td>Attrition Bias Incomplete Outcome Data</td>
<td>Other Bias</td>
<td>Funding Source</td>
<td>Reviewer Comments</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------------------------------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Kanders et al, 1988</td>
<td>Ok</td>
<td>No</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>NutraSweet Company, the Center for Nutritional Research Charitable Trust, and National Institutes of Health</td>
<td>Participants were informed they may be receiving an aspartame intervention.</td>
</tr>
<tr>
<td>Kanders et al, 1994</td>
<td>Ok</td>
<td>No</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>NutraSweet Company and the Center for Nutritional Research Charitable Trust</td>
<td>Participants were informed they may be receiving an aspartame intervention.</td>
</tr>
<tr>
<td>Madjd et al, 2015</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>None</td>
<td>School of Life Sciences, University of Nottingham and the Digestive Disease Research Institute (DDRI), affiliated with Tehran University of Medical Sciences</td>
<td>Participants were blinded, but not personnel; Power analyses suggested n=67.</td>
</tr>
<tr>
<td>Madjd et al, 2017</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>None</td>
<td>School of Life Sciences, University of Nottingham and the Digestive Disease Research Institute (DDRI), affiliated with Tehran University of Medical Sciences</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Selection Bias</td>
<td>Selection Bias</td>
<td>Reporting Bias</td>
<td>Performance Bias</td>
<td>Detection Bias</td>
<td>Attrition Bias</td>
<td>Other Bias</td>
<td>Funding Source</td>
<td>Reviewer Comments</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>----------------</td>
<td>------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Maersk et al, 2012</td>
<td>Ok</td>
<td>No</td>
<td>Ok</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Ok</td>
<td>None</td>
<td>Last value carried forward for missing data or left blank.</td>
</tr>
<tr>
<td>Peters et al, 2014</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>The American Beverage Association</td>
<td>Intent-to-treat analyses used</td>
</tr>
<tr>
<td>Peters et al, 2016</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>The American Beverage Association</td>
<td>Intent-to-treat analyses used</td>
</tr>
<tr>
<td>Raben et al, 2002</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>Danish Research and Development Programme for Food Technology 1990-1994 and Danisco Sugar</td>
<td>Participants were informed they may be receiving a NNS intervention, but not the purpose of the study.</td>
</tr>
<tr>
<td>Reid et al, 2007</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>Biotechnology and Biological Sciences Research Council (BBSRC)</td>
<td>Half of participants were misinformed about the type of beverage they would receive in the study.</td>
</tr>
<tr>
<td>Reid et al, 2010</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>None</td>
<td>BBSRC and the Sugar Bureau</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Selection Bias</td>
<td>Selection Bias Allocation Concealment</td>
<td>Reporting Bias Selective Reporting</td>
<td>Performance Bias Blinding (participants and personnel)</td>
<td>Detection Bias Blinding (outcomes assessment)</td>
<td>Attrition Bias Incomplete Outcome Data</td>
<td>Other Bias</td>
<td>Funding Source</td>
<td>Reviewer Comments</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------</td>
<td>----------------------------------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>------------------------------------------</td>
<td>------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>Tate et al, 2012</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>Ok</td>
<td>No</td>
<td>Ok</td>
<td>None</td>
<td>Nestle Waters USA</td>
<td>Intent-to-treat analyses used</td>
</tr>
<tr>
<td>Tordoff et al, 1990</td>
<td>Ok</td>
<td>Unclear</td>
<td>Ok</td>
<td>Unclear</td>
<td>No</td>
<td>Ok</td>
<td>None</td>
<td>USDA’s Competitive Research Grants Program</td>
<td>All participants received all three treatment interventions</td>
</tr>
</tbody>
</table>

Note: *Using Cochrane Risk of Bias Assessment