



Protocol

# A Standardized Guide to Developing an Online Grocery Store for Testing Nutrition-Related Policies and Interventions in an Online Setting

Pasquale E. Rummo<sup>1,\*</sup>, Isabella Higgins<sup>2,†</sup>, Christina Chauvenet<sup>3</sup>, Annamaria Vesely<sup>2</sup>,  
Lindsay M. Jaacks<sup>4</sup> and Lindsey Taillie<sup>2,5</sup>

<sup>1</sup> Department of Population Health, New York University Grossman School of Medicine, New York, NY 10016, USA

<sup>2</sup> Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; ihiggins@email.unc.edu (I.H.); avesely@email.unc.edu (A.V.); taillie@unc.edu (L.T.)

<sup>3</sup> Prevention Research Center, University of South Carolina, Columbia, SC 29208, USA; cchauvenet@sc.edu

<sup>4</sup> Global Academy of Agriculture and Food Security, University of Edinburgh, Midlothian EH25 9RG, UK; Lindsay.jaacks@ed.ac.uk

<sup>5</sup> Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

\* Correspondence: pasquale.rummo@nyulangone.org

† These authors contributed equally to this work.

**Abstract:** Simulated online grocery store platforms are innovative tools for studying nutrition-related policies and point-of-selection/point-of-purchase interventions in online retail settings, yet there is no clear guidance on how to develop these platforms for experimental research. Thus, we created a standardized guide for the development of an online grocery store, including a detailed description of (1) methods for acquiring and cleaning online grocery store data, and (2) how to design a two-dimensional online grocery store experimental platform. We provide guidance on how to address product categorization, product order/sorting and product details, including how to identify outliers and conflicting nutritional information and methods for standardizing prices. We also provide details regarding our process of “tagging” food items that can be leveraged by future studies examining policies and point-of-selection/point-of-purchase interventions targeting red and processed meat and fruits and vegetables. We experienced several challenges, including obtaining accurate and up-to-date product information and images, and accounting for the presence of store-brand products. Regardless, the methodology described herein will enable researchers to examine the effects of a wide array of nutrition-related policies and interventions on food purchasing behaviors in online retail settings, and can be used as a template for reporting procedures in future research.

**Keywords:** food policy; experimental design; red meat; fruits and vegetables; sustainable diet; food labels; food marketing and promotion



**Citation:** Rummo, P.E.; Higgins, I.; Chauvenet, C.; Vesely, A.; Jaacks, L.M.; Taillie, L. A Standardized Guide to Developing an Online Grocery Store for Testing Nutrition-Related Policies and Interventions in an Online Setting. *Int. J. Environ. Res. Public Health* **2021**, *18*, 4527. <https://doi.org/10.3390/ijerph18094527>

Received: 27 January 2021

Accepted: 22 April 2021

Published: 24 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Diets rich in energy, added sugars, processed meat and unprocessed red meat, and low in whole grains, nuts, fruits and vegetables, are positively associated with risk of poor cardio-metabolic health and some cancers [1–4]. A higher-quality diet is associated with a lower risk of obesity, diabetes, cardiovascular disease and some cancers [4], emphasizing the importance of a healthy eating pattern in chronic disease prevention. However, very few adults in the US consume the recommended amounts of fruit, vegetables and red meat, potentially due to cost, food and taste preferences and cultural norms [5–8]. Nutrition-related policies and interventions that affect food retail, particularly at the point-of-selection (where consumers make decisions about what to buy) and point-of-purchase (where the purchasing decision actually occurs), are important tools for reducing the risk of nutrition-related diseases. Specific strategies include, for example, financial incentives and

disincentives (e.g., discounts and taxes), warning labels, product placement and healthy default foods and beverages (e.g., pre-set menu options) [9–11]. In more recent years, interest in these practices as strategies to reduce the environmental harm caused by food production, such as greenhouse gas emissions [12], has also grown.

Online food retail has grown quickly in the past few years, with a marked increase due to the COVID-19 pandemic [13,14]. This sudden shift provides a unique opportunity to leverage evidence-based strategies to encourage healthy food purchases (e.g., fruits and vegetables) and discourage less healthy food purchases (e.g., red meat and sugar-sweetened beverages) in online food retail settings. Compared to brick-and-mortar stores, it is easier to modify the design of online environments to influence consumer decision-making, such as modifying prices, editing front-of-package labeling and adding default shopping cart items. With support from software developers, behavioral health and nutrition researchers can evaluate the impact of nutrition-related policies and interventions on consumer behavior in online store experiments.

Researchers in several countries (e.g., New Zealand, Singapore, UK, USA), for example, have collaborated with software developers to build a platform for online behavioral research studies, which has the appearance and functionality of a real grocery store website, including browsing, search, product pages, shopping cart and checkout [15–21]. These experimental online stores are also linked to food databases, with nutritional information collected from food product labels. In addition to two-dimensional online stores, researchers have also created three-dimensional “virtual” supermarkets and other online retail stores (e.g., clothing) [22–25], which are designed to mimic a real shopping experience in a brick-and-mortar store. These “virtual” online stores are built to test the impact of various retail strategies, such as food swaps, price changes and nutrition labels, on purchase intentions. However, to our knowledge, there are no detailed methods on how to create two-dimensional online grocery stores for experimental research, which is important for consistency in the literature and maximizing generalizability by creating realistic shopping experiences.

To fill this gap, we sought to create standardized methods for the development of a grocery store designed to mimic an online shopping experience for the study of nutrition-related policies and interventions in an online retail setting, including a detailed description of (1) methods for acquiring and cleaning online grocery store data, and (2) how to design a two-dimensional online store for experimental research, including product categorization, product order and sorting, product images, product details (e.g., price, nutritional information) and tagging specific products for nutrition-related interventions, using fruits and vegetables and red meat as two examples.

## 2. Materials and Methods

### 2.1. Online Grocery Store Experimental Platform

To create an online grocery store, we used an online shopping platform called Gorilla (<https://gorilla.sc>) developed by UK-based Cauldron Science, Ltd. (<http://cauldron.sc>), a company with expertise in designing platforms for online behavioral research studies. Gorilla includes a tool called the Shop Builder, which enables researchers to create an online store that can be modified according to the researchers’ goals. The Gorilla platform also provides the functionality to randomize participants to conditions and administer surveys and tasks. The platform works with third-party panel vendors to recruit participants (e.g., CloudResearch). After providing informed consent, eligible participants are directed to the online store to complete specified shopping tasks, with instructions tailored to experimental conditions (e.g., budget amount, target products). The experimental platform allows for a variety of experimental stimuli, including warning labels, taxes, discounts, default shopping cart items and suggested swaps for specific products. Patients or the public were not involved in the design, conduct, reporting or dissemination plans of our guide.

## 2.2. Online Grocery Store Data

### 2.2.1. Data Source

To obtain online grocery store data, we contracted a company that provides custom web-scraping solutions (Marquee Data), which enabled us to compile a large dataset (<20,000 observations) in a timely manner (two weeks) at a low cost (<\$3000). A similar approach has been utilized previously by other researchers (18). The data were scraped from one of the largest online grocery stores in the USA in August 2020. We selected this store because it has a large presence throughout the USA and it has affordable prices. The product data included: department, aisle (i.e., department sub-categories), shelf (i.e., aisle sub-categories), product name, price, serving size, servings per container, calories, fat, saturated fat, trans fat, carbohydrates, sugar, sodium, fiber, protein, and iron, ingredients list, UPC, shelf rank and the product images (Supplementary Table S1). The store brand name was removed from product names to minimize the influence of brand loyalty, but the products themselves were retained to provide products at a range of prices, as store brands are often cheaper. The department (broadest category), aisle and shelf (most granular category) were used by the store to designate the location of a product, and the same product could be located in multiple departments, aisles and shelves. The serving size, servings per container, calories, fat, saturated fat, trans fat, carbohydrates, sugar, sodium, fiber, protein and iron were scraped from the nutritional facts panel (NFP) data displayed on the product's page. We also scraped "shelf rank", which indicates a product's order on a shelf when sorted by best seller.

We chose to scrape data at a single time-point from a single zip code because (1) it was cost-efficient, (2) it ensured all study participants were exposed to the same dataset, (3) it was too time-consuming to clean multiple datasets and (4) it ensured that the relative prices of the products were consistent for all participants. We chose to scrape data from zip code 63119 (St. Louis, MO, USA) because it represents to 95.3% the US cost-of-living (i.e., 4.7% below the average cost-of-living in the US), [26] and retailers have selected St. Louis as a pilot location for food products (e.g., the Impossible Whopper at Burger King) [27]. The initial dataset had 23,314 products, including duplicate products appearing in more than one shelf category.

### 2.2.2. Data Cleaning

To ensure the scraped data were of high quality and to prepare the data for our planned experiments, we conducted several phases of data cleaning, including (1) defining the initial list of products and filling in missing NFP information, (2) performing a data audit, (3) tagging products and (4) online grocery store review.

## 2.3. Phase 1: Define Initial List of Products and Fill in Missing NFP Information

In the first phase of data cleaning, we decided which products to eliminate from the dataset and addressed missing nutritional information. We removed all alcoholic beverages as alcohol laws vary from state to state and some states do not allow the sale of any alcohol in grocery stores. We removed all non-edible products, such as cake toppers and cupcake liners, because our planned experiments are focused on modifying food and beverage purchases. In addition, we removed the shelves "New in Pantry" and "New in Fresh Bakery" because all products were duplicates. After these exclusions, approximately 12% of the 20,320 remaining products did not contain calorie information.

As a first step, we eliminated freshly-made and bakery products missing calorie information because nutritional information was not available on the grocery store's website or other sources. We also deleted products missing calorie information if they had a shelf rank of greater than 60 because our store only displayed 60 products per page and we did not expect participants to browse more than one page per shelf. We did not, however, eliminate any products missing calorie information corresponding to the food groups in our planned research studies (i.e., the "Fruits & Vegetables," "Meat & Seafood" or "Deli" departments).

For the remaining products, we extracted all missing nutritional information from the product's image on the grocery store's website. If the information was not available, we extracted data from the product manufacturer's website. For fresh produce, we collected missing nutritional information from the FoodData Central website, an integrated, research-focused data system that provides data on nutrients and other food components [28]. We searched for items using as much detail as possible and selected the produce item that was the closest match, including product name, variety (e.g., "Granny Smith"), method of preparation (e.g., "raw," "sliced," "peeled," "in syrup," etc.) and other details (e.g., added vitamins). If variety, method of preparation or other details were not identified in our search, we selected the generic option. FoodData Central has several data types, including "SR Legacy," which provides nutrient and food component values that are derived from analyses, calculations and the published literature. The "Food and Nutrient Database for Dietary Studies (FNDDS)" contains data on nutrient and food component values derived from the National Health and Nutrition Examination Survey. "Branded Foods" provides values for nutrients in branded and private-label foods from a public-private partnership. Because fresh fruits and vegetables were not branded, we selected "Survey (FNDDS)" or "SR Legacy" items instead of "Branded", when possible. If the produce serving size was missing, we referenced the American Heart Association's serving size recommendations [29].

#### 2.4. Phase 2: Data Audit

In the second phase, we completed a data audit to identify potential systematic issues with the dataset. Four members of the study staff each reviewed 100 randomly-selected products from the dataset ( $n = 400$  total products) and assessed the extent to which the information from the scraped data matched the information on the grocery store's website. We found that NFP informational text from the grocery store's website often differed from the NFP data in the image from the product manufacturer. We decided to keep the NFP information from the grocery store's website because transcribing data from >10,000 images would be time-consuming and error-prone. To exclude advertisements and NFP images, we decided to only use the first image scraped for each product (which was always a front-of-package image). To identify extreme values of price, calories, sugar, sodium and saturated fat, we reviewed the top 1% of values for each variable within each department, and the bottom 1% for price values. We did not review the bottom 1% of values for the other variables as it is common for products to have low or zero values for calories, sugar, sodium and saturated fat. To identify outliers, we first compared the price and nutritional information of products in our dataset to the grocery store's website. If the prices matched but still appeared extreme, we used the price from the website of a different large online grocery store. If the prices were different, we updated the value using the other store's price. If the values for other variables matched but still appeared extreme, we used the values from the NFP image or the product manufacturer's website, if more plausible. In total, the price, calories, sugar, sodium and/or saturated fat of 1031 products (5%) were modified and 140 products (<1%) were removed; our final dataset included 20,180 products (Table 1).

**Table 1.** N (%) of products by department.

Department	Product [N (%)]
Beverages	2793 (13.8%)
Bread & Bakery	954 (4.7%)
Deli	740 (3.7%)
Eggs & Dairy	2004 (9.9%)
Frozen	2192 (10.9%)
Fruits & Vegetables	914 (4.5%)
Meat & Seafood	719 (3.6%)
Organic Shop	743 (3.7%)
Pantry	6264 (31.0%)
Snacks & Candy	2857 (14.2%)
<b>Total</b>	<b>20,180</b>

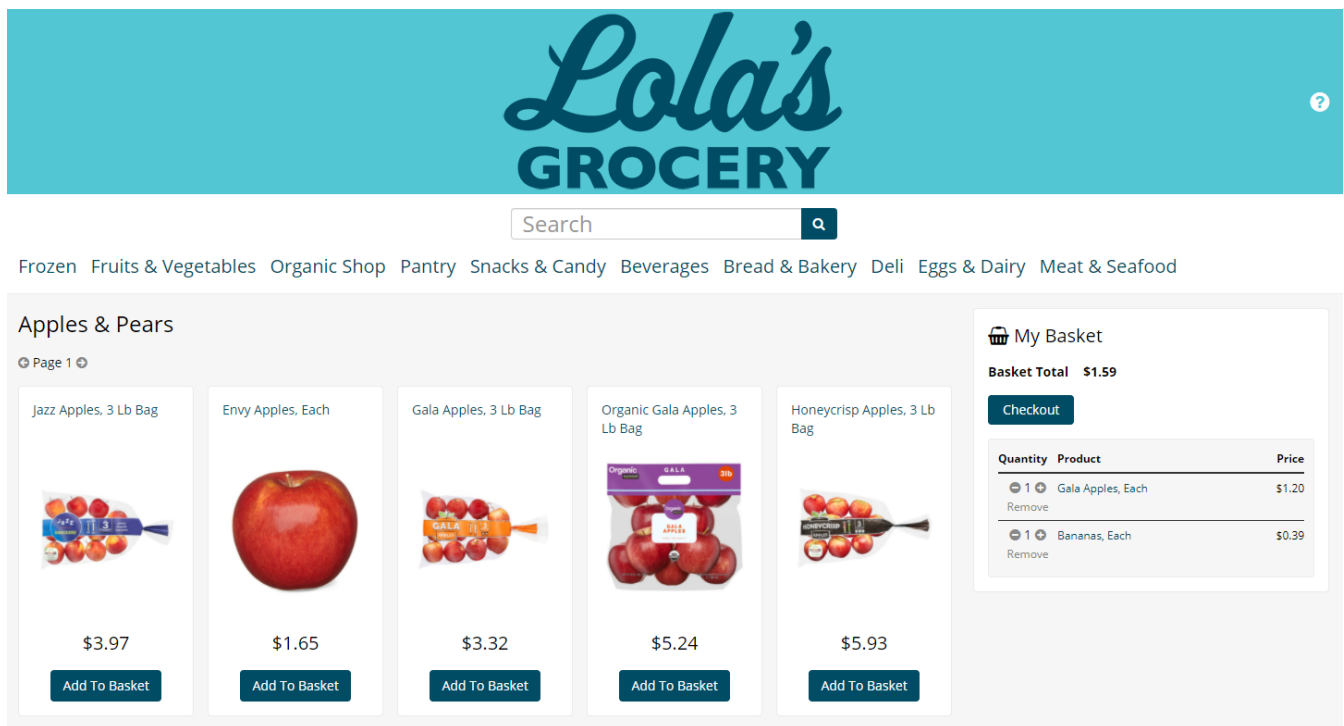
### 2.5. Phase 3: Online Grocery Store Review

We uploaded the final dataset to the online grocery store experimental platform and reviewed the first page of each shelf (Figure 1). First, we reviewed the first five products on the product page of each shelf and confirmed that the product name, image and ingredients list corresponded to the same/correct product. We also searched for incorrect applications of our point-of-selection strategies, image errors and missing and/or inaccurate product information. During our review, we found few errors regarding the application of the point-of-selection strategies, product name, image and ingredients list (<0.5%). However, we found that the servings per container and/or serving size were missing for several products. To mitigate this issue, we reviewed the departments, aisles and shelves with products relating to our planned studies. If the serving size and/or servings per container were missing, we extracted all missing nutritional information from the NFP on the grocery store's website, or, if not available, from the NFP image. If the information was missing from both sources, we searched the FoodData Central website for missing fruit and vegetable product data and the product manufacturers' websites for other product data [28]. In total, we updated the NFP for 814 (4%) products.

In addition to data cleaning, we "tagged" products that we wished to modify in our planned experiments. Tagging is accomplished on the "back end" of the server (versus the "front end" or user interface). The goals of two of our planned studies using the online grocery store were to (1) assess the impact of taxes and health and environmental warning labels on consumers' decisions to purchase red meat, and (2) examine the extent to which an economic incentive and a behavioral "nudge" increased fruit and vegetable purchases among low-income adults in the US. To prepare the dataset for these experiments, we tagged relevant products with a "0" or "1" for pre-specified intervention manipulation variables. See Supplementary Tables S2 and S3 for details of the tagging procedures.

After tagging all products, three staff members reviewed 200 randomly-selected products (n = 600 total products) to ensure the tagging was accurate, and did not identify errors in product tagging. In total, 1718 (9%) products were tagged as eligible for the red meat tax and health and environment warning labels, including 407 (55%) products in the "Deli" department and 395 (55%) products in the "Meat & Seafood" department. Among the red meat products, 59% contained beef, 67% contained pork, 26% contained beef and pork, 0.2% contained other red meat and 91% were processed. In total, 831 (4%) products were tagged as eligible for the fruit and vegetable discount, including 538 (59%) products in the "Fruits and Vegetables" department and 71 (10%) products in the "Organic Shop" department. Among all products tagged as eligible for the fruit and vegetable intervention, 14%, 16% and 71% were frozen, canned and fresh produce, respectively.





**Figure 1.** Online grocery store appearance.

## 2.6. Online Grocery Store Design

To maximize the external validity of our experimental findings, our goal was to make online shopping tasks as realistic as possible. To accomplish this goal, we created our own store name, “Lola’s,” and worked with a graphic designer to create a unique logo (Figure 1). In the future, we plan to develop additional graphic images to augment the store’s realistic appearance (e.g., ad banners). To mimic the appearance and functionality of a typical US grocery store’s website, we organized products by department, aisle and shelf (Supplementary Table S1). The department names are listed across the top of the screen, below the store logo. The aisle names appear when a department is clicked on, and, subsequently, shelf names appear when an aisle is clicked on. Then the participant can select a shelf to view individual products, with a maximum of 60 products per page, which are sorted according to the “shelf rank” (i.e., the popularity of the item), as scraped from the online retailer. Participants can also view products by typing words into the search bar, which will display individual products containing the inputted word(s). When a participant clicks on a product, the product appears on a new webpage, where a participant can view the product’s name and price, a front-of-package image, NFP text and an ingredient list (Figure 2).

The online grocery store has a shopping basket so participants can make multiple selections. The shopping basket shows the quantity of the product selected and plus and negative signs, which allow the participant to add or remove products from their basket. The basket also shows the cost of each product and the total cost of all products. The platform allows researchers to require that participants have a minimum and/or maximum number of items in their basket before checking out. The platform also has an optional shopping list feature, which enables researchers to outline which items participants should select from the online grocery store (i.e., task-based shopping). The shopping list includes check-boxes next to each item so participants can keep track of products in their basket and which products they still need to select, if desired. The shopping basket and shopping list are anchored to the right side of the screen (Figure 1; Supplementary Figure S1).



## Nutrition Facts

Amount per serving

**Calories 70 cal**

About 18 servings per container

<b>Serving Size</b>	50 g
<b>Total Fat</b>	5 g
Saturated Fat	1.5 g
Trans Fat	0 g
<b>Sodium</b>	70 mg
<b>Total Carbohydrates</b>	0 g
Fiber	
Sugar	
<b>Protein</b>	6 g

## Details

**Ingredients**  
EGG

**Figure 2.** Individual product page on online grocery store.

To discover potential issues with our initial design of the store, we recruited six participants and instructed them to conduct a household shop with a budget of \$50 while ‘thinking out loud.’ Based on participants’ feedback, we worked with the software developer to increase the upload time of the website, make the placement and size of the search bar more prominent, and update the “Add to basket” button to better convey when a product was added to a participant’s shopping cart.

### 3. Discussion

Through this guide, we are building the capacity to examine the impact of nutrition-related policies and interventions on food purchasing behaviors in online retail settings. Though several researchers have used online grocery stores to test healthy eating interventions [15–21], this guide is the first to provide a comprehensive description of how to create an online grocery store platform, which will minimize differences in the experimental platform of studies testing similar research questions (e.g., number and placement of items). Specifically, we provide guidance on how and where to obtain nutritional information for products missing such data (e.g., fresh fruits and vegetables), how to identify outliers and conflicting information and methods for standardizing the price of food items. Though the data and software used to develop our online store are proprietary, researchers will be able to mimic our approach by obtaining a contract and licensing fee from Marquee and Cauldron, respectively. In addition, we provided specific details regarding our process of tagging food items that can be leveraged by future studies examining nutrition-related policies and interventions targeting red and processed meat and fruit and vegetable purchases. Importantly, the methodology described herein can be used as a template for reporting

procedures in future research, including procedures for data collection and cleaning, data sources and platform design.

We experienced several challenges in the process of obtaining product information and product images for our online grocery store. Because we used a commercial vendor to acquire data from a real US grocery store, it was not feasible to develop a “living” dataset with up-to-date products. Thus, our data reflects a single point in time, which is potentially problematic because manufacturers are constantly creating new products (e.g., meat substitutes) and reformulating the ingredients in pre-existing products. However, our results remain internally valid, given the randomization of study participants, and our goal is to update the products prior to initiating new experimental research. Another challenge was accounting for the presence of store-brand products. For example, we decided to remove store-brand names from the product name but did not create our own store-brand name; this was to avoid potential effects of brand loyalty on purchases, which may be higher for online purchasing [30]. This may reduce the generalizability of our results for a segment of the population, given the role of brand loyalty in guiding some purchase decisions. Yet, the exclusion of store-brand products means our store has wider appeal for all participants, who shop at a variety of stores, and our database includes similar products at several price points, so participants have a large breadth of options. We also identified discrepancies between the NFP text provided by the online grocery store and the image provided by the product manufacturer. We decided to use the NFP text provided by the online grocery store because it was feasible to scrape such data (vs. the NFP image provided by the manufacturer) and thus was less prone to human error, but it is possible that these data did not reflect the most up-to-date nutritional information.

A key strength of our approach is its adaptability. Though we designed the online grocery store to be compatible with our specific research goals, we also made decisions to maximize its flexibility, such as filling in missing information for all products and nutrients. This flexibility will allow us and other researchers to test a wide variety of strategies, including point-of-selection (e.g., product labels, price changes) and point-of-purchase (e.g., shopping cart defaults), on a wide variety of outcomes (e.g., dollars spent, calories purchased, proportion of items accepted); this also allows us to tailor our products, including the sample size (>20,000 products), to the study goals, upload times and for other considerations. Our decision to acquire data from a specific grocery store in a specific zip code minimized potential issues related to the availability of products sold by third-party vendors, discrepancies in list price versus offer price, and differences in prices by region. As such, though prices will not reflect the cost-of-living in the area where study participants shop for groceries, any differences in prices between our experimental store and a real store will be equal across experimental conditions. We also combined secondary data collection procedures with primary data collection efforts, which enabled us to include products that lack NFP data in our online grocery store (e.g., fresh fruits and vegetables) and solicit participants’ feedback during the design phase. However, it was not feasible to identify and correct extreme values for all nutrients, and it was not possible to add personalized features linked to user details. The latter is an important limitation because our experimental platform will not be capable of mimicking unique features of online grocery stores, such as the promotion of products based on previous purchases. For this reason, it will be important to conduct validation studies to compare consumers’ food and beverage purchases in simulated versus real online grocery stores, including participants with varying levels of digital literacy and various food preferences. Lastly, a key weakness of our guidance is the proprietary nature of our data and our license to use the experimental online grocery store platform, which may prevent other researchers from exactly mimicking our store’s products and the design of our store.

#### 4. Conclusions

By following and adhering to *a priori* methods for creating online grocery stores—such as our approach to scraping and cleaning data, obtaining missing nutritional information



and standardizing price data—public health researchers can maximize the comparability and reproducibility of their online experimental research. This will facilitate future meta-analyses of effect estimates stemming from this work. The need for a template will only grow as more and more consumers in the US are turning to online grocery shopping, especially as national emergencies (e.g., virus outbreaks, climate change) engender new patterns of grocery shopping. Our guidance and methods are also purposefully adaptable, allowing for flexibility in design and capacity in future studies, which is important as the form and function of online grocery stores change over time and between regions/countries: the online grocery shopping experience will continue to evolve and experimental research platforms will need to be continuously updated to remain realistic. Other steps for strengthening the generalizability of our proposed and future research include developing additional graphic images to make the shopping experience more realistic (e.g., ad banners), and improving the search function of the online grocery store (e.g., growing our list of unique search terms). In sum, the dissemination of our methods and the findings from our proposed research goals will be used to inform nutrition-related policies and interventions designed to promote healthier food shopping behaviors in online grocery stores.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ijerph18094527/s1>, Figure S1: Shopping list and shopping basket, Table S1: Data scraping request, Table S2: Procedures for tagging meat products, Table S3: Procedures for tagging eligible fruit and vegetables

**Author Contributions:** Conceptualization, P.E.R., L.M.J. and L.T.; methodology, P.E.R., I.H., L.M.J. and L.T.; formal analysis, I.H.; resources, P.E.R., L.M.J. and L.T.; data curation, I.H. and C.C.; writing—original draft preparation, P.E.R. and I.H.; writing—review and editing, P.E.R., I.H., C.C., A.V., L.M.J. and L.T.; visualization, P.E.R. and I.H.; supervision, P.E.R., L.M.J. and L.T.; project administration, I.H. and C.C.; funding acquisition, P.E.R., L.M.J. and L.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Institute of Aging of the National Institutes of Health (grant number 1K01AG064146-01) and the Wellcome Trust (grant number 16042/Z/19/Z). L.T. and I.H. were also supported by the Carolina Population Center’s NIH Center grant P2C HD050924. The funding sources had no role in the study design; in the collection, analysis, or interpretation of data; in the writing of the report; nor in the decision to submit the article for publication.

**Institutional Review Board Statement:** Study protocols were approved by the New York University Grossman School of Medicine Institutional Review Board and the University of North Carolina at Chapel Hill Institutional Review Board.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data may be obtained from a third party and are not publicly available.

**Acknowledgments:** We would like to thank Jamie Halula, Sophia Hurr, Rhea Naik, Carmen Preston and Carla Seet for their help in cleaning the dataset and reviewing the online grocery store. Additionally, we would like to thank Emily Busey for designing our online grocery store logo, Bridgett Hollingsworth for providing guidance on how to tag and process red meat, Yiqing Zhang for providing insight into the data scraping process and Wilma Waterlander for her general feedback on store development and the process evaluation. We would also like to thank the researchers involved in developing the Woods supermarket platform for their support, including Susan Jebb and Brian Cook at the Nuffield Department of Primary Care Health Sciences at the University of Oxford.

**Conflicts of Interest:** No conflicts of interest or financial disclosures were reported by the authors of this paper.

## References

1. Aune, D.; Giovannucci, E.; Boffetta, P.; Fadnes, L.T.; Keum, N.; Norat, T.; Greenwood, D.C.; Riboli, E.; Vatten, L.J.; Tonstad, S. Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality—a systematic review and dose-response meta-analysis of prospective studies. *Int. J. Epidemiol.* **2017**, *46*, 1029–1056. [CrossRef] [PubMed]
2. Boeing, H.; Bechthold, A.; Bub, A.; Ellinger, S.; Haller, D.; Kroke, A.; Leschik-Bonnet, E.; Müller, M.J.; Oberritter, H.; Schulze, M.; et al. Critical review: Vegetables and fruit in the prevention of chronic diseases. *Eur. J. Nutr.* **2012**, *51*, 637–663. [CrossRef] [PubMed]
3. Rouhani, M.H.; Haghghatdoost, F.; Surkan, P.J.; Azadbakht, L. Associations between dietary energy density and obesity: A systematic review and meta-analysis of observational studies. *Nutrition* **2016**, *32*, 1037–1047. [CrossRef] [PubMed]
4. Zhong, V.W.; Van Horn, L.; Greenland, P.; Carnethon, M.R.; Ning, H.; Wilkins, J.T.; Lloyd-Jones, D.M.; Allen, N.B. Associations of Processed Meat, Unprocessed Red Meat, Poultry, or Fish Intake With Incident Cardiovascular Disease and All-Cause Mortality. *JAMA Intern. Med.* **2020**, *180*, 503–512. [CrossRef] [PubMed]
5. Cheah, I.; Sadat Shimul, A.; Liang, J.; Phau, I. Drivers and barriers toward reducing meat consumption. *Appetite* **2020**, *149*, 104636. [CrossRef]
6. Livingstone, K.M.; Burton, M.; Brown, A.K.; McNaughton, S.A. Exploring barriers to meeting recommendations for fruit and vegetable intake among adults in regional areas: A mixed-methods analysis of variations across socio-demographics. *Appetite* **2020**, *153*, 104750. [CrossRef] [PubMed]
7. Moore, L.V.; Thompson, F.E. Adults Meeting Fruit and Vegetable Intake Recommendations—United States, 2013. *Mmwr. Morb. Mortal. Wkly. Rep.* **2015**, *64*, 709–713.
8. Zeng, L.; Ruan, M.; Liu, J.; Wilde, P.; Naumova, E.N.; Mozaffarian, D.; Zhang, F.F. Trends in Processed Meat, Unprocessed Red Meat, Poultry, and Fish Consumption in the United States, 1999–2016. *J. Acad. Nutr. Diet.* **2019**, *119*, 1085–1098.e1012. [CrossRef]
9. Afshin, A.; Peñalvo, J.L.; Del Gobbo, L.; Silva, J.; Michaelson, M.; O’Flaherty, M.; Capewell, S.; Spiegelman, D.; Danaei, G.; Mozaffarian, D. The prospective impact of food pricing on improving dietary consumption: A systematic review and meta-analysis. *PLoS ONE* **2017**, *12*, e0172277. [CrossRef] [PubMed]
10. Anzman-Frasca, S.; Mueller, M.P.; Sliwa, S.; Dolan, P.R.; Harellick, L.; Roberts, S.B.; Washburn, K.; Economos, C.D. Changes in children’s meal orders following healthy menu modifications at a regional US restaurant chain. *Obesity* **2015**, *23*, 1055–1062. [CrossRef]
11. Niebylski, M.L.; Redburn, K.A.; Duhaney, T.; Campbell, N.R. Healthy food subsidies and unhealthy food taxation: A systematic review of the evidence. *Nutrition* **2015**, *31*, 787–795. [CrossRef]
12. Willett, W.; Rockström, J.; Loken, B.; Springmann, M.; Lang, T.; Vermeulen, S.; Garnett, T.; Tilman, D.; DeClerck, F.; Wood, A.; et al. Food in the Anthropocene: The EAT-Lancet Commission on healthy diets from sustainable food systems. *Lancet* **2019**, *393*, 447–492. [CrossRef]
13. Kamel, M.-A.; de Montgolfier, J.; Caine, S.; Ringer, J.; Puzio, S. How to Ramp Up Online Grocery—Without Breaking the Bank. Available online: <https://www.bain.com/insights/how-to-ramp-up-online-grocery-without-breaking-the-bank/> (accessed on 19 October 2020).
14. Business Insider. Online Grocery Report. 2020. Available online: <https://www.businessinsider.com/online-grocery-report-2020> (accessed on 18 November 2020).
15. Epstein, L.H.; Finkelstein, E.; Raynor, H.; Nederkoorn, C.; Fletcher, K.D.; Jankowiak, N.; Paluch, R.A. Experimental analysis of the effect of taxes and subsidies on calories purchased in an on-line supermarket. *Appetite* **2015**, *95*. [CrossRef]
16. Epstein, L.H.; Finkelstein, E.A.; Katz, D.L.; Jankowiak, N.; Pudlewski, C.; Paluch, R.A. Effects of nutrient profiling and price changes based on NuVal@scores on food purchasing in an online experimental supermarket. *Public Health Nutr.* **2016**, *19*, 2157–2164. [CrossRef]
17. Finkelstein, E.A.; Ang, F.J.L.; Doble, B.; Wong, W.H.M.; van Dam, R.M. A randomized controlled trial evaluating the relative effectiveness of the multiple traffic light and Nutri-score front of package nutrition labels. *Nutrients* **2019**, *11*, 2236. [CrossRef]
18. Forwood, S.E.; Ahern, A.L.; Marteau, T.M.; Jebb, S.A. Offering within-category food swaps to reduce energy density of food purchases: A study using an experimental online supermarket. *Int. J. Behav. Nutr. Phys. Act.* **2015**, *12*, 1–10. [CrossRef]
19. Koutoukidis, D.A.; Jebb, S.A.; Ordóñez-Mena, J.M.; Noreik, M.; Tsiountsioura, M.; Kennedy, S.; Payne-Riches, S.; Aveyard, P.; Piernas, C. Prominent positioning and food swaps are effective interventions to reduce the saturated fat content of the shopping basket in an experimental online supermarket: A randomized controlled trial. *Int. J. Behav. Nutr. Phys. Act.* **2019**, *16*, 50. [CrossRef] [PubMed]
20. Riches, S.P.; Aveyard, P.; Piernas, C.; Rayner, M.; Jebb, S.A. Optimising swaps to reduce the salt content of food purchases in a virtual online supermarket: A randomised controlled trial. *Appetite* **2019**, *133*, 378–386. [CrossRef] [PubMed]
21. Waterlander, W.E.; Jiang, Y.; Nghiem, N.; Eyles, H.; Wilson, N.; Cleghorn, C.; Genç, M.; Swinburn, B.; Mhurchu, C.N.; Blakely, T. The effect of food price changes on consumer purchases: A randomised experiment. *Lancet Public Health* **2019**, *4*, e394–e405. [CrossRef]
22. Waterlander, W.E.; Jiang, Y.; Steenhuis, I.H.M.; Mhurchu, C.N. Using a 3D virtual supermarket to measure food purchase behavior: A validation study. *J. Med. Internet Res.* **2015**, *17*, e107. [CrossRef] [PubMed]

23. Hoenink, J.C.; Mackenbach, J.D.; Waterlander, W.; Lakerveld, J.; van der Laan, N.; Beulens, J.W.J. The effects of nudging and pricing on healthy food purchasing behavior in a virtual supermarket setting: A randomized experiment. *Int. J. Behav. Nutr. Phys. Act.* **2020**, *17*, 98. [CrossRef] [PubMed]
24. Stuber, J.M.; Hoenink, J.C.; Beulens, J.W.J.; Mackenbach, J.D.; Lakerveld, J. Shifting toward a healthier dietary pattern through nudging and pricing strategies: A secondary analysis of a randomized virtual supermarket experiment. *Am. J. Clin. Nutr.* **2021**. [CrossRef] [PubMed]
25. Papagiannidis, S.; Pantano, E.; See-To, E.W.; Bourlakis, M. Modelling the determinants of a simulated experience in a virtual retail store and users' product purchasing intentions. *J. Mark. Manag.* **2013**, *29*, 1462–1492. [CrossRef]
26. Cost of Living in St. Louis, Missouri. Available online: [https://www.bestplaces.net/cost\\_of\\_living/city/missouri/st\\_louis](https://www.bestplaces.net/cost_of_living/city/missouri/st_louis) (accessed on 18 November 2020).
27. Maze, J. Burger King Expands Its Impossible Whopper Test. 2019. Available online: <https://www.restaurantbusinessonline.com/food/burger-king-expands-its-impossible-whopper-test> (accessed on 18 November 2020).
28. US Department of Agriculture. FoodData Central. Available online: <https://fdc.nal.usda.gov/> (accessed on 18 November 2020).
29. American Heart Association. Fruits and Vegetables Serving Sizes Infographic. Available online: <https://www.heart.org/en/healthy-living/healthy-eating/add-color/fruits-and-vegetables-serving-sizes> (accessed on 18 November 2020).
30. Dawes, J.; Nenycz-Thiel, M. Comparing retailer purchase patterns and brand metrics for in-store and online grocery purchasing. *J. Mark. Manag.* **2014**, *30*, 364–382. [CrossRef]