32nd International Conference of Agricultural Economists
2-7 August 2024 I New Delhi I India

Novel approaches to analyze consumer behavior and policies to promote healthy and sustainable consumption

Katare Bhagyashree[1], Jacqueline Yenerall[2], Shuoli Zhao[3], Xuejian Wang[1]

1: Department of Agricultural Economics, Purdue University
2. Department of Agricultural and Resource Economics, University of Tennessee
3: Department of Agricultural Economics, University of Kentucky
Corresponding author email: bkatare@purdue.edu

Abstract

Technological advancements, such as online grocery shopping, have significantly transformed consumer retail environments and experiences. Effectively studying consumer behavior in these new environments requires the use of novel methodological approaches, which will also aid in the development of interventions to encourage healthy and sustainable consumption. This paper begins by providing an overview of the current literature on novel approaches to analyzing consumer behavior. To contribute to this literature, the paper also examines consumer decision-making pathways within online grocery shopping platforms. Specifically, the paper focuses on exploring the consumers' digital footprints, such as page visits, product additions and removals, and interactions with information labels to identify patterns and interests in consumer responses to healthy and sustainable consumption. The study investigates potential heterogeneities in consumers' socio-demographics and attitudes, aiming to provide insights for shaping online shopping environments to promote healthy and sustainable food choices. Findings highlight the potential benefits of integrating consumer search tracking data with environment design to facilitate informed and conscious food choices.

JEL Codes: D12, D9

Novel approaches to analyze consumer behavior and policies to promote healthy and sustainable consumption

Abstract

Technological advancements, such as online grocery shopping, have significantly transformed consumer retail environments and experiences. Effectively studying consumer behavior in these new environments requires the use of novel methodological approaches, which will also aid in the development of interventions to encourage healthy and sustainable consumption. This paper begins by providing an overview of the current literature on novel approaches to analyzing consumer behavior. To contribute to this literature, the paper also examines consumer decision-making pathways within online grocery shopping platforms. Specifically, the paper focuses on exploring the consumers' digital footprints, such as page visits, product additions and removals, and interactions with information labels to identify patterns and interests in consumer responses to healthy and sustainable consumption. The study investigates potential heterogeneities in consumers' socio-demographics and attitudes, aiming to provide insights for shaping online shopping environments to promote healthy and sustainable food choices. Findings highlight the potential benefits of integrating consumer search tracking data with environment design to facilitate informed and conscious food choices.

Keywords: online shopping, machine learning, big data, consumer behavior, facial recognition technology

JEL Codes: D12, D9

1 INTRODUCTION

The consumer retail environment has been significantly transformed by digitization, greatly changing the consumer shopping experience. A notable example is the rise of online grocery shopping, enabling consumers to engage in routine shopping from the comfort of their homes. This shift also provides retailers with new opportunities to target consumers with tailored advertisements and price promotions based on the analysis of high volumes of data generated from online and in-person shopping (Burris, 2023; Rajagopal, 2023). These technological advancements add complexity to consumer behavior as consumers increasingly engage in multichannel shopping, which can include in-person shopping, online shopping through a retailer's website or specialized platforms such as Instacart, as well as seeking product information through social media.

Policymakers continue to be concerned with promoting healthy and sustainable consumption to support public health and environmental sustainability. Given the complexities of modern consumer behavior, novel analytical approaches are needed to investigate the role of technological advancements in healthy and sustainable consumption. The emergence of big data and machine learning, geospatial analysis, and social media has revolutionized the study of consumer behavior, particularly in the context of food consumption. Big data analytics, for instance, can reveal seasonal consumption trends, the impact of promotions on buying habits, and preferences for specific food attributes such as organic or locally sourced products (Huang et al., 2023; Singh & Glińska-Neweś, 2022). Using information available on social media, researchers can gauge public opinion on various food products and trends through sentiment analysis and thematic analysis of posts, comments, and shares (Cash et al., 2022, Mogaji et al., 2021; Samoggia et al., 2020). Machine learning can predict consumer food preferences and behaviors with high accuracy, enabling personalized marketing strategies (Shen et al., 2021). Similarly, facial recognition technology is emerging as a powerful tool for studying consumer reactions and engagement (Mordor Intelligence, 2020).

These novel analytical approaches have the potential to provide deeper insights into how consumer behaviors respond to technological advancements as well as identify and test interventions designed to support healthy and sustainable consumption. Different strategies can be employed to encourage healthy and sustainable consumption, from nudges and incentives to comprehensive regulatory frameworks. By integrating advanced analytical techniques with targeted policy interventions, there is an opportunity to significantly impact public health and environmental sustainability. Deeper insights, more accurate predictions, and enhanced consumer engagement strategies can enable companies to sustainably meet consumer needs. By leveraging these new technologies, the food industry can drive innovation, improve customer satisfaction, and contribute to healthier and more sustainable consumer behaviors.

This paper aims to contribute to this goal by highlighting novel approaches to consumer behavior analysis. First, we provide an overview of the current literature on novel approaches to analyzing consumer behavior. Second, we empirically examine the decision-making pathways of consumers as they navigate a simulated online grocery shopping platform. Respondents are randomly assigned to one of three treatments or a control group to investigate their influence on several measures of online shopping behavior: page visits, product additions and removals, and interactions with information labels. Findings highlight the potential benefits of integrating consumer search tracking data with environment design to facilitate informed and conscious food choices.

2 OVERVIEW OF CURRENT LITERATURE

2.1 Machine Learning and Big Data Analytics

Machine learning (ML) has become a powerful tool for analyzing consumer behavior, using sophisticated techniques to uncover patterns and predict future actions from larger datasets. Big data analytics has amplified this potential by providing the infrastructure to handle and analyze large and complex datasets that exceed the capabilities of traditional methods (Lu et al., 2019). Integrating big data with advanced ML models offers creative tools for marketing analytics, guiding managerial decisions in customer relationship management, product placement, pricing, promotion, and personalization (Wedel & Kannan, 2016).

The synergy of machine learning and big data analytics is powerful, as ML algorithms rely on large and complex datasets, and big data analytics provide the means to efficiently collect, store, and process these datasets. This combination provides a deeper and more accurate insight into consumer behavior, facilitating more effective and personalized marketing strategies (Chandra et al., 2022; Zhou et al., 2021). Wang et al. (2019) highlighted the scalability of ML and big data in promoting sustainable consumer behavior across different regions and demographics, and such advancements are critical for adapting to changing consumer preferences and behavior.

One important application of ML in consumer behavior is predictive analytics, where ML algorithms can analyze individual preferences and behaviors to provide personalized recommendations for healthier and more sustainable choices. Mitchell et al. (2021) used Attributable Components Analysis, an ML technique, and personal health data to provide personalized nutrition goal recommendations. Amin et al., (2021) used ML methods, such as random forests and least absolute shrinkage and selection Operator (LASSO) to predict the access to healthful food retailers. Predictive models help identify patterns in consumer behavior, allowing businesses to anticipate future actions and tailor their strategies accordingly. Recent research has employed ML techniques (Random Forest, Decision Tree, K-Nearest Neighbor, etc.) to predict the likelihood of consumers adopting sustainable products, providing insights for targeted green marketing campaigns (Choudhury et al., 2024).

Using large-scale data from various sources, such as social media, transaction records, and wearable devices, big data analytics provides comprehensive analysis for a holistic view of consumer behavior, identifying key factors influencing health and sustainability choices

(Erevelles et al., 2016). By analyzing big data, researchers can segment consumers based on their behavior and preferences, allowing for more targeted and effective interventions. Wedel & Kannan, (2016) provide a review of methods and studies of how big data analytics could segment consumers to promote sustainable purchasing behavior effectively. Big data allows for real-time monitoring and analysis of consumer behavior, enabling timely interventions. For example, studies have suggested the use of real-time data from social media and mobile applications to encourage healthy eating habits among users (e Silva et al., 2018; Ma et al., 2022)

2.2 New technological innovations in studying consumer behavior

Emerging technologies offer novel methods to study how individuals make choices, interact with food products, and engage with food-related information. Virtual Reality (VR) and Augmented Reality (AR) have gained traction as tools to study consumer behavior in the food industry. Additionally, the adoption of VR by retailers, such as Amazon's VR kiosks or eBay's VR department store app, also demonstrates VR's potential to modify consumer experiences and behavior (Xi & Hamari, 2021).

VR immerses individuals in simulated environments, allowing researchers to observe their reactions to different food-related scenarios. Innocenti (2017) provided an early overview of VR in experimental economics, differentiating between low-immersive (LIVE) and high-immersive virtual environments (HIVE). LIVE involves interactions with real or virtual environments on screens, while HIVE uses technology like headsets to create three-dimensional environments (Innocenti, 2017; Mol, 2019). HIVE's realism can generate more natural responses in experimental settings. Furthermore, VR simulations can replicate shopping experiences in virtual supermarkets, enabling researchers to analyze consumer responses to product displays, packaging designs, and marketing strategies(Xi & Hamari, 2021). Similarly, AR overlays digital content in the real world, offering opportunities for interactive product demonstrations and personalized marketing experiences (Huang & Benyoucef, 2013). Studies utilizing VR and AR have provided insights into consumer preferences, sensory perceptions, and purchasing decisions in food environments (Dong et al., 2021; Heller et al., 2019).

The Internet of Things (IoT) and smart devices have revolutionized data collection and analysis in consumer behavior research. IoT-enabled devices, such as smart refrigerators, wearable sensors, and kitchen appliances, capture real-time data on individuals' food consumption patterns, dietary habits, and nutritional intake (Boland et al., 2019; Fujiwara et al., 2018). By integrating IoT data with machine learning algorithms, researchers can generate personalized recommendations, dietary interventions, and behavior change strategies tailored to individual preferences and health goals (Sundaravadivel et al., 2018). Furthermore, smart packaging equipped with RFID (Radio-frequency Identification) tags or sensors provides valuable insights into product freshness, shelf-life management, and supply chain logistics, influencing consumer perceptions of food quality and safety (Chen et al., 2020; Zuo et al., 2022).

Facial recognition technology has emerged as a promising tool for understanding consumer behavior, particularly in the context of food product consumption. By analyzing facial expressions, researchers can gain insights into consumers' emotional responses, preferences, and decision-making processes during food-related decision-making. One of the primary applications of facial recognition in consumer behavior research is the analysis of emotional responses to food stimuli. Facial recognition algorithms can detect micro-expressions, subtle changes in facial muscles that reflect underlying emotions such as happiness, surprise, disgust, or sadness (Ekman, 2002). For example, studies have used facial recognition technology to measure consumers' facial expressions while tasting various food samples or viewing food advertisements, providing valuable insights into the emotional drivers of food preferences and choices (Mellouk & Handouzi, 2020). Facial recognition software is being used to track users' facial expressions and gestures while interacting with food-related interfaces, such as restaurant menus, food delivery apps, or self-service kiosks (Marstone, 2017; Rankin, 2017). This allows researchers to understand how consumers navigate these interfaces and identify potential pain points for improvement.

2.3 Social Media and Consumer Behavior

Social media is defined as an online platform that allows for the "creation and exchange of user-generated content" (Kaplan & Haenlein, 2010), and includes platforms such as Facebook, X, WhatsApp, YouTube, and Instagram. Research using data from social media has steadily increased in the 2000s, likely due in part to the wealth of data that can be leveraged to gain insight into consumer preferences and trends (Rini et al., 2024; Titova et al., 2023; Tumasjan, 2024). This data is appealing to researchers for two main reasons. First, it generates big data sets with large volumes, variety, and complexity, as users share text, pictures, and videos frequently and with a wide audience (Li et al., 2021). Second, it represents a novel technology that has transformed consumer interactions with each other and with firms (Li et al., 2021).

While numerous studies have explored social media data to understand user engagement (Santini et al., 2020), methods for analyzing social media data (Bazzaz Abkenar et al., 2021; Xu et al., 2022; Zachlod et al., 2022), and specific marketing topics like influencer marketing(Vrontis et al., 2021) and content marketing (du Plessis, 2022), this discussion focuses on using social media to understand consumer behavior, particularly food-related behaviors. Titova et al. (2022) and Rini et al. (2024) review the literature on using social media to study dietary and food consumer behavior, respectively, noting a steady increase in articles over time, with Twitter being the most commonly used data source.

In food and consumer behavior research, social media data can be analyzed in a standalone analysis, integrated into broader studies, or combined with external data. Initial steps often involve identifying food types and related attitudes using thematic, content, sentiment analysis, natural language processing, AI, ML, or manual approaches (Titova et al., 2022). Sentiment analysis specifically discerns the text's positive, negative, or neutral sentiments. Standalone analysis of social media data has examined attitudes toward organic foods, perceptions of novel food technologies (e.g., cultured meat), and niche products employing topic modeling and sentiment analysis (Kouarfaté & Durif, 2023; Li & Hu, 2021; Singh & Glińska-Neweś, 2022).

When social media data is integrated into larger studies, the purpose is not always to study online consumer behavior (Rini et al., 2024). For instance, Pancer et al. (2022) explored the impact of mindset on healthy food engagement by incorporating recipe videos from a food-focused Facebook page, categorized by caloric content, into a between-subject experiment. Similarly, Liu & Lopez (2016) utilized a unique dataset from Nielsen that merged scanner data with social media conversations from platforms like Facebook, Twitter, and YouTube to examine consumer preferences for carbonated soft drinks. On the other hand, (Philp et al., 2022) investigated the concept of "Instagrammable" food by showing respondents Instagram posts from top restaurants and assessing their likelihood of involvement.

Researchers face several challenges when utilizing social media data, which should be addressed in future studies. Firstly, social media data may suffer from sample selection bias, as it predominantly represents individuals comfortable with technology, and social desirability bias, as users may tailor their content to seek peer approval (Titova et al., 2022). Secondly, with the rise of social media marketing and influencers, distinguishing between firm-generated and consumer-generated content becomes increasingly complex (Li et al., 2021).

2.4 Innovations in Choice Experiments to Study Consumer Behavior

Discrete choice experiments (DCE), which have a theoretical foundation in random utility theory and Lancaster's consumer theory, have been increasingly popular since the early 2000s (Lizin et al., 2022). One advantage of DCEs is their applicability in hypothetical scenarios, making them ideal for evaluating novel products or features not yet available in the market. Yet the reliance on hypothetical choices or stated preferences has raised concerns about their external validity and reliability (Lizin et al., 2022). Our discussion focuses on three innovations particularly relevant for studying new technologies: the inclusion of quantities, basket-based choice experiments, and the incorporation of new technology such as virtual reality (VR) and videos to simulate social media environments (Caputo & Scarpa, 2022; Fang et al., 2021).

DCEs traditionally study binary choices by having respondents select their most preferred option from a set of alternatives that vary based on attributes, including price. However, this approach does not reflect real-world scenarios where consumers can choose various quantities of multiple products. To address this, Dennis et al. (2021) introduced the open-ended choice experiment (OECE), which allows respondents to select a quantity for each alternative thereby capturing the ‘none of these’ alternatives commonly included in DCEs. They randomly assigned the respondents to either a DCE or OECE for three meat products: pork, beef, and chicken. The OECE's choice frequencies and preferences better matched observed consumer behavior from the Livestock Marketing Center. This showed that OECE is more consistent with economic theory and provides greater flexibility for studying consumer behaviors and estimating demand frameworks like the Almost Ideal Demand System (Dennis et al., 2021).

Lin et al. (2023) proposed an alternative experimental design to researcher-specified quantities in hypothetical experiments, incorporating consumers' reported purchase quantities into choice questions. This approach addresses the phenomenon of mental accounting in behavioral economics, where a consumer's marginal propensity to spend depends on the relationship between the market price and their mental budget for that expenditure category. If the price exceeds this budget, the propensity to purchase declines. In DCEs, lower quantities increase the likelihood that costs fall within this mental budget, which can lead to an overestimation of willingness to pay (WTP). Lin et al. (2023) randomly assigned participants to either a DCE with fixed quantities or one with quantities based on reported purchasing behavior to test this hypothesis. They found that when researcher-specified quantities were lower than consumers' actual purchases, the marginal utility of money decreased, resulting in an overestimated WTP. This study highlights the importance of aligning experimental designs with real consumer behavior to obtain more accurate results.

While both Lin et al. (2023) and Dennis et al. (2021) address the limitations of restricting quantities in DCEs they focused on a single product or a small set of similar products, which can be restrictive when studying broader consumer behavior or the value of different food products. Caputo & Lusk (2022) introduced the basket-based choice experiment (BBCE) to overcome these limitations. In their study, they included 21 distinct food items at three different price levels, allowing consumers to choose multiple items and revise their choices before finalizing. This approach enabled the estimation of a demand system to identify if products were complements or substitutes.

Other studies have explored the impact of technological advancements on consumer behavior (Hu et al., 2022; Liu & Lopez, 2016; Xi & Hamari, 2021) and addressed limitations in choice experiments using technology (Fang et al., 2020; Mol, 2019; Innocenti, 2017) or empirical innovation (Zhao and Yue, 2020). As discussed previously, in VR HIVE's realism can generate more natural responses in experimental settings. Studies have tested HIVE's ability to enhance choice experiment performance, such as (Mokas et al., 2021), who found that VR improved respondents' interpretation of complex information in a DCE for urban greenery compared to text or video.

3 EXPERIMENTAL SETUP

To contribute to the recent literature on novel methods in studying consumer behavior, this paper explores consumer decision-making in a digital environment, specifically as they navigate online grocery shopping platforms. The growing popularity of online grocery shopping is fundamentally transforming the grocery retail landscape and altering consumer preferences.

With over 45% of consumers now shopping for groceries online more frequently than before (Redman, 2021), this trend is rapidly reshaping the food system. Factors such as convenience, ease of use, customization, promotional offers, accessible information, and diverse product assortments significantly influence the adoption of online grocery shopping technology. Importantly, given the increasing policy and marketing interests in promoting healthy and sustainable consumption patterns, understanding how consumers interact with information in these digital spaces is crucial. This study seeks to address this gap by examining how different information cues (labeling, categorization, or combination) within the online grocery shopping environment influence consumer decision pathways.

We utilized a simulated online grocery store, the Open Science Online Grocery (OSOG) platform (Howe et al., 2022). The OSOG platform contains over 11,000 real-world food products, spanning a wide range of categories, brands, and price points, providing participants with a realistic choice of items. The platform closely replicates the experience of shopping at a typical online grocery store and provides features such as product browsing, searching, and the ability to add to a shopping cart and check out. This simulated environment allows for controlled experimentation while maintaining a high degree of ecological validity, as participants engage in decision-making processes similar to those in real-world grocery shopping. Together, this approach allows us to explore consumer responses to information cues related to health and sustainability in a setting that closely replicates real-world decision-making. The study was approved by the institutional review board at the author's university.

Sample: A nationally representative sample of 2,405 participants was recruited through an online panel. Participants were selected using quota sampling (Yang & Banamah, 2014), a non-probability sampling technique designed to ensure the sample reflects the proportions of specific characteristics (e.g., age, gender, income, race, geographic region) found in the target population. After completing the online grocery shopping in OSOG, participants also completed a survey to provide their demographic information (Table 1).

Research design: Participants were instructed to purchase items with a budget of $20 and a minimum spending requirement of $10. These limits were implemented to simulate constraints often encountered in real-world grocery shopping. The participants were randomly assigned to one of four experimental conditions within the OSOG platform: 1) *Labeling* group participants were shown color-coded carbon footprint labels with greenhouse gas (GHG) emission values, 2) *Categorization* group participants were shown a curated category of plant-based products, 3) the *Combined* group participants were shown both the carbon footprint label and the curated category, and 4) the control group participants were not shown any labels or special categories. The sample demographic characteristics are balanced over the four groups. Appendix Table A1 lists the plant- and animal-based products used in the experiment.

We focus on consumers' digital footprints, which have been investigated in the literature to study cognitive load (Wang et al., 2014) and consumer shopping behavior, such as size of order (Wang et al., 2015). However, this is the first study to investigate how information labels influence consumer digital footprints.[1] In this application, the user activity captured in the digital footprint provides insight into a consumer decision-making process as it relates to choices at the intersection of health and sustainability. Specifically, the digital footprints monitor consumers' behavior as they try to decide between animal-based foods and their plant-based alternatives while responding to GHG information treatments. Thus, the experiment provides an ideal opportunity to explore (almost) in real-time the consumer decision-making processes underlying the choice between healthier and more sustainable products (Bryant, 2022).

[1] For detailed information on the experimental design and product description, please refer to Katare and Zhao (*forthcoming*). The effect of these interventions on food choices is studied in the forthcoming manuscript. The current study contributes to the literature on novel techniques to study consumer digital footprint in online shopping platforms, such as page visits, product additions and removals, and interactions with information labels.

We use the information derived from consumer interaction with the web pages to create the following outcomes.

View time: The OSOG platform can monitor the time a participant spends viewing a product, starting from when a participant clicks on the product until they exit the product description. This detailed tracking allows for a comprehensive analysis of consumer interests and engagement. By examining the view times, we can better understand which product categories attract more attention and potentially drive purchasing decisions. We calculated the view time for all the products viewed by the participants, and the view time for plant- and animal-based products, separately.

Hover time: This variable captures the time participants hover their mouse over the GHG label, which provides insights into user interest and engagement with the information presented. We calculated the hover time for all the products viewed by the participants and for plant-based and animal-based products.

4 EMPIRICAL MODEL

As expected, the customer footprint variables have a high proportion of zero values (77% for page views and 22% for hover time). Hence to investigate the impact of the three interventions on respondents' digital footprint, we employ a type I Tobit model regression model at the level of each purchase. Specifically, the model is expressed as:

$$Footprints_i = \alpha + \beta\, Treatment_i + \gamma X_i + \mu_i \qquad (1)$$

$Footprints_i$ is the outcome variable measured by the total time for page views and label hover time for each participant i. $Treatment_i$ is a vector of indicator variables that identifies the mutually exclusive treatment to which each participant i belongs. We control for participant age, household size, income, marital status, education level, race, employment status, and region of residence as these factors can influence food choices (X_i). Standard errors are corrected for heteroskedasticity.

5 RESULTS

Table 2 presents the mean and standard deviation for view and hover times for different product categories. *Control* group participants spent the most time viewing products overall (4.28 seconds), specifically animal-based products (2.33 seconds). The *categorization* group spent relatively more time viewing plant-based products (1.85 seconds), particularly the products they purchased. Please note that only the *Labeling* and *Combined* groups were exposed to GHG labels. The *Combined* group was shown both GHG labels and the curated plant-based category and spent significantly more time hovering over GHG labels (82.05 seconds) than the *Labeling* group (63.14 seconds). For purchased products, the *Control* group again exhibited the longest view times. Findings suggest that the interventions influenced consumer engagement and decision-making processes in the online environment, with the combined group displaying heightened attention to environmental impact information.

Digital Footprint: Table 3 examines the impact of GHG labels on specific shopping actions. Of the 538 participants who hovered over a GHG label and subsequently added that product to their cart, 282 were in the labeling group, suggesting a potential influence of GHG labels on purchase decisions. However, it is essential to note that a significant number of participants added both plant-based and animal-based products to their carts before interacting with any GHG label, highlighting pre-existing preferences. Interestingly, 119 participants, primarily in the labeling (65) and combined (54) groups, hovered over a GHG label and then removed an animal-based product from their cart. This suggests that the label may have prompted a reconsideration of environmentally impactful choices. However, the overall number of animal-based product deletions was similar across all groups, indicating that other factors play a role in these decisions. The number of plant-based product deletions was comparable across groups, suggesting that the presence of GHG labels did not significantly alter the removal of plant-based products from carts. The recorded digital footprints by treatment groups indicate that while GHG labels may influence some purchasing decisions, they are not the sole determinant. Pre-existing preferences and other factors also play an important role in the dynamics of online grocery shopping behavior.

Regression results for product view time: Marginal effects of interventions and demographic factors on product view time are presented in Table 4. Exposure to GHG labels (*Labeling*) significantly reduced total view time (by 3.53 seconds) and animal-based product view time (by 6.33 seconds) compared to the control group, suggesting that the labels may have reduced cognitive taxation and accelerated decision-making, particularly for less sustainable products. The presence of a curated plant-based category (*Categorization*) led to increased view time for plant-based products (2.66 seconds), but a decrease in total view time (1.72 seconds) and animal-based product view time (7.97 seconds). As expected, displaying a curated product category encouraged a focus on plant-based options, potentially at the expense of overall browsing. The *Combined* intervention exhibited the most pronounced effects, significantly reducing total view time (5.99 seconds), plant-based view time (1.34 seconds), and animal-based view time (14.67 seconds), suggesting more efficient decision-making across all products. The view time for products ultimately purchased, the patterns largely mirrored the overall effects.

Demographic factors also significantly influenced view time. Females spent less time viewing all product categories compared to males. Older participants tended to have longer view times across all product categories. Married individuals spent less time viewing all product categories. Larger households spent less time viewing all product categories. Participants with a college degree spent more time viewing all product categories, and income did not have a significant effect on view time. Employed individuals spent considerably less time viewing all product categories.

Regression results for hover time: Table 5 presents the marginal effects of interventions and demographic factors on GHG label hover time (the time spent hovering the cursor over a label image). Unlike view time, hover time specifically captures active engagement with product information. *Labeling* and *Combined* were the only two groups exposed to the GHG labels. The *Combined* intervention had no significant effect on total hover time for any product categories compared to the *Labeling* intervention. For the products purchased, the *Combined* treatment significantly decreased hover time compared to the labeling treatment. This suggests that product categorization assisted participants' decision-making by providing easier navigation to the preferred products, regardless of the product type.

Similar to the view time analysis, demographic factors significantly affected the product and label hover time. Females tended to hover significantly less than males. Older participants generally hovered longer over the labels, suggesting more deliberate consideration. Married individuals hovered less than unmarried participants, potentially indicating different shopping styles. Higher education levels were associated with increased hover times, and employed individuals hovered significantly less. Lastly, income and race did not significantly affect the hover time.

6 DISCUSSION

The ongoing integration of new technologies in the consumer retail environment has significantly influenced consumer behavior. As these technologies advance, there is a need for advanced analytical methods to effectively study and understand consumer behavior in new contexts, to encourage healthier and more sustainable consumption. The literature review explores the emergence of advanced technologies such as machine learning, big data analytics, social media information extraction, and facial recognition, which significantly enhance researchers' ability to study the complexities of modern consumer behavior.

The paper also provides an original analysis of consumer footprint data from an online grocery shopping environment to examine the decision-making pathways of consumers as they navigate the online platform. Specifically, the empirical analysis focuses on exploring the consumers' digital footprints, such as page visits, product additions and removals, and interactions with information labels (hover time). This information helps identify patterns and interests in consumer responses to healthy and sustainable consumption. Findings provide insight into consumer decision-making and the influence of factors available in the environment (e.g.: labeling) on consumption. Consumer footprint information allowed for a comprehensive analysis of consumer interest and engagement, providing valuable insights into preferences and behaviors related to product selection.

Our results reinforce and extend previous research using page views and mouse hovers to study consumer decision-making processes. As established in prior studies(Huang et al., 2009; Moe, 2003), page or product view time is a reliable indicator of consumer interest, with longer view time often correlating with increased purchase intention. Mouse hover time, on the other hand, provides a unique angle that reflects the cognitive processes of consumers as they engage and evaluate different attributes of a product. Longer hover times generally signal a heightened level of interest or uncertainty, prompting further consideration (Pieters & Wedel, 2007).

Tracking and analyzing these digital footprints in online shopping environments, this study demonstrates the value of adopting such an approach in gaining additional insights into consumer behavior. Furthermore, our analysis of consumer attention toward information and labeling, such as carbon footprints, highlights the potential policy implications for motivating healthy and sustainable consumer choices. As technology continues to advance, the ability to track and analyze these metrics will become increasingly sophisticated, providing stakeholders, such as retailers, policymakers, and health advocates, with powerful tools to understand and influence purchasing decisions. By leveraging these insights, more engaging and effective shopping experiences can be created to encourage healthy and sustainable consumption habits.

These findings, combined with the advancements in consumer behavior research outlined in the literature review, point to a promising way for leveraging technology to promote healthier and more sustainable food choices. Integrating innovative tools, like machine learning, big data analytics, and virtual reality with carefully designed interventions like carbon footprint labeling and curated product categories, can empower consumers to make informed decisions aligned with their values and dietary needs. This integrated approach has the potential to transform the food system by shifting consumer preferences towards more sustainable and health-conscious options.

While new technologies offer powerful tools for studying consumer behavior, they also raise ethical concerns related to privacy and data security. The large amounts of personal data collected from various sources, including social media platforms, personal use devices, and online transactions, pose significant risks to individual privacy. The not-so-transparent methods of data collection, lack of informed consent, and lack of communication about the way the data will be used can lead to concerns about data misuse and mistrust. Data privacy is another important concern, as data breaches and unauthorized access to personal information can have harmful implications, such as identity theft. Additionally, Big data algorithms can perpetuate the biases present in the data, which can lead to discriminatory practices, such as targeted advertising and credit scoring (Baruh et al., 2017).

To fully harness the potential of these technologies while upholding ethical principles, researchers and practitioners should prioritize transparency, informed consent, and robust data security measures. Proactive efforts to mitigate algorithmic bias and ensure equitable access to the benefits of these advancements are equally crucial.

Table 1: Descriptive statistics of demographic variables

	All sample	Labeling	Categorization	Combined	Control
Female	0.49	0.56	0.51	0.50	0.48
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age (years)	43.85	42.96	43.92	44.33	43.79
	(16.38)	(16.26)	(16.46)	(15.72)	(16.31)
Married	0.70	0.73	0.70	0.69	0.70
	(0.46)	(0.45)	(0.46)	(0.46)	(0.46)
Household size	3.20	3.34	3.18	3.15	3.21
	(1.51)	(1.52)	(1.53)	(1.43)	(1.48)
College-educated	0.54	0.53	0.55	0.55	0.54
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Race = White	0.72	0.73	0.69	0.70	0.75
	(0.45)	(0.44)	(0.46)	(0.46)	(0.44)
Income (thousand USD)	87.09	90.39	89.54	89.78	84.66
	(44.55)	(46.10)	(45.80)	(46.82)	(43.17)
Employed	0.77	0.78	0.78	0.76	0.77
	(0.42)	(0.41)	(0.42)	(0.42)	(0.42)
N	2405	584	609	602	610

Note: Compared to the U.S. Census Bureau (2023), the gender ratio in the sample aligns with the national average (50%). However, the survey sample shows higher averages for age, household size, and percentage of married individuals (39%, 2.3%, and 47%, respectively). Additionally, the percentages of college-educated individuals, white respondents, employed persons, and those with higher incomes are 10-20% above the national averages (41%, 61%, and 60%, respectively) and exceed the U.S. median income of $74,580.

Table 2: Outcome variables derived from the consumer footprints in the OSOG platform

Variable (in seconds)	All	Labeling	Categorization	Combine	Control
Average total view time	2.17	1.97	2.34	1.76	2.59
	(5.25)	(5.02)	(5.49)	(4.58)	(5.78)
Average plant-based products view time	0.82	0.71	0.97	0.80	0.79
	(2.30)	(2.14)	(2.49)	(2.19)	(2.34)
Average animal-based products view time	0.82	0.82	0.83	0.55	1.09
	(2.54)	(2.55)	(2.60)	(2.04)	(2.87)
Average view time for products purchased	1.41	1.40	1.42	1.19	1.63
	(3.67)	(3.72)	(3.65)	(3.31)	(3.96)
Average view time for	0.45	0.42	0.48	0.49	0.39
plant-based products purchased	(1.44)	(1.40)	(1.48)	(1.49)	(1.40)
Average view time for	0.59	0.62	0.55	0.39	0.79
animal-based products purchased	(1.96)	(2.01)	(1.93)	(1.60)	(2.24)
Average hover time	22.77	44.89	.	47.40	.
over GHG labels	(54.79)	(68.68)	.	(72.79)	.
Average hover time over	14.21	28.03	.	29.57	.
GHG labels for plant-based products	(32.61)	(40.22)	.	(43.06)	.
Average hover time over	6.80	13.49	.	14.06	.
GHG labels for animal-based products	(20.64)	(27.68)	.	(27.77)	.
Average hover time over GHG labels	3.42	7.13	.	6.75	.
for products purchased	(8.10)	(10.52)	.	(10.33)	.
Average hover time over GHG labels	1.61	3.35	.	3.17	.
for plant-based products purchased	(4.05)	(5.31)	.	(5.26)	.
Average hover time over GHG labels	0.91	1.78	.	1.90	.
for animal-based products purchased	(3.10)	(4.14)	.	(4.28)	.
N	2405	584	609	602	610

Note: We show the average total view and hover time during a purchase session. Total view time is the average of the total time consumers spend viewing all the products. Total plant-based view time is the average of the total time consumers spend viewing plant-based products. Total animal-based view time is the average of the total time consumers spend viewing animal-based products. Rows 4-5 show the average of the total time consumers spend viewing products they purchased in each category. GHG hover time is the average of the total time consumers spend hovering over carbon footprint labels. Time is reported in seconds. Appendix Table A2 provides the results for t-test between the control and the treatment groups.

Table 3: The number of people who change actions impacted by GHG labels

	All	Labeling	Categorization	Combined	Control
Number of participants who hovered on the GHG label, and added the corresponding product to their cart within 30 minutes	538	282	-	256	-
Number of participants adding plant-based, before hover GHG label	367	182	-	185	-
Number of participants adding animal-based, before hover GHG label	629	324	-	305	-
Number of participants who hovered on the GHG label then deleted an animal-based product from the cart within 30 minutes	119	65	-	54	-
Number of participants who deleted an animal-based product	1036	258	274	226	278
Number of participants who deleted a plant-based product	1057	293	261	233	270
N	2405	584	609	602	610

Table 4: Marginal effects on product view duration

	(1) View time for all products	(2) View time for plant-based products	(3) View time for animal-based products	(4) View time for all products purchased	(5) View time for plant-based products purchased	(6) View time for animal-based products purchased
Labeling	-1.473*	0.449	-2.076***	-0.496	0.803	-1.247
	(0.994)	(0.766)	(0.976)	(0.925)	(0.721)	(0.954)
Categorization	-0.739	1.370**	-3.031***	-0.605	1.085*	-2.700***
	(1.000)	(0.750)	(1.010)	(0.933)	(0.714)	(1.005)
Combined	-2.378***	0.919	-4.861***	-1.586**	1.170**	-4.160***
	(1.014)	(0.747)	(1.045)	(0.947)	(0.705)	(1.043)
Female	-1.112*	-0.666	-0.539	-0.560	-0.145	-0.662
	(0.729)	(0.524)	(0.757)	(0.674)	(0.489)	(0.747)
Age	0.235***	0.0793***	0.304***	0.217***	0.0793***	0.258***
	(0.0277)	(0.0192)	(0.0356)	(0.0258)	(0.0182)	(0.0339)
Married	-2.056***	-0.866	-1.812**	-1.602**	-0.434	-2.190***
	(0.913)	(0.668)	(0.996)	(0.856)	(0.637)	(0.942)
Household	-1.384***	-0.409***	-1.982***	-1.129***	-0.279	-1.627***
	(0.304)	(0.204)	(0.388)	(0.288)	(0.196)	(0.385)
College	0.838	0.214	1.170	0.564	-0.139	1.103
	(0.775)	(0.554)	(0.836)	(0.714)	(0.519)	(0.826)
White	0.555	-0.145	0.974	0.857	-0.0218	1.553*
	(0.838)	(0.592)	(0.959)	(0.792)	(0.565)	(0.970)
Income	0.0136*	0.00716	0.0144*	0.00842	0.00334	0.0155**
	(0.00877	(0.00632)	(0.00899)	(0.00818)	(0.00608)	(0.00855)
Employed	-6.685***	-3.185***	-4.310***	-5.287***	-2.011***	-3.988***
	(0.966)	(0.716)	(0.984)	(0.875)	(0.653)	(0.961)
Constant	-9.605***	-8.364***	-18.22***	-12.93***	-10.81***	-18.34***
	(2.237)	(1.597)	(2.644)	(2.145)	(1.581)	(2.578)
N	2405	2405	2405	2405	2405	2405

Note: The table presents the marginal effects from the type I Tobit model estimation for product view time as the outcome variable. View time is the sum of time a participant spends viewing different product pages. Plant-based view time accounts for plant-based products. Animal-based view time only accounts for animal-based products. View time for purchased products only accounts for products that were ultimately purchased. Unit of time is second. The base category is Control. The specifications control for demographic variables including, age, gender, education, race, marital status, household size, income, employment status, and region of residence. The standard errors in the parenthesis are corrected for heteroskedasticity. ***p<0.05, **p<0.10, *p<0.15.

Table 5 Marginal effects for participant GHG label hover time

	(1) Hover time for all products	(2) Plant-based products hover time	(3) Animal-based products hover time	(4) Hover time for all products purchased	(5) Hover time for plant-based products purchased	(6) Hover time for animal-based products purchased
Combined	-4.741	-2.843	-2.622	-1.830***	-0.875*	-0.786
	(4.112)	(2.612)	(3.895)	(0.866)	(0.588)	(0.961)
Female	-9.282***	-5.552***	-8.558***	-1.095	-0.388	-1.994***
	(4.118)	(2.630)	(3.985)	(0.867)	(0.589)	(0.975)
Age	1.096***	0.595***	0.889***	0.111***	-0.00696	0.195***
	(0.155)	(0.0963)	(0.157)	(0.0319)	(0.0216)	(0.0384)
Married	-15.49***	-9.374***	-18.60***	-0.566	0.487	-3.731***
	(5.639)	(3.555)	(5.096)	(1.112)	(0.761)	(1.207)
Household	-7.394***	-3.873***	-8.478***	-1.299***	-0.272	-1.568***
	(1.583)	(1.011)	(1.816)	(0.341)	(0.227)	(0.448)
College	13.39***	6.220***	19.29***	1.941***	0.377	3.234***
	(4.404)	(2.773)	(4.355)	(0.913)	(0.625)	(1.057)
White	3.241	0.292	3.722	-0.232	-0.716	-0.492
	(4.449)	(2.841)	(4.677)	(0.947)	(0.644)	(1.136)
Income	0.0620	0.0761***	0.0556	0.00428	0.00868	0.0116
	(0.0567)	(0.0356)	(0.0483)	(0.0110)	(0.00732)	(0.0111)
Employed	-56.92***	-28.24***	-47.60***	-6.613***	0.648	-9.523***
	(7.580)	(4.651)	(5.318)	(1.362)	(0.972)	(1.260)
Constant	60.72***	29.75***	-3.045	6.673***	-1.084	-4.889*
	(11.59)	(7.303)	(11.96)	(2.469)	(1.736)	(3.025)
Obs.	1186	1186	1186	1186	1186	1186

Note: The table presents the marginal effects from the type I Tobit model estimation for hover time as the outcome variable. The dependent variable is the total hover time that a participant spends hovering on GHG labels. Plant-based hover time accounts for plant-based products. Animal-based hover time only accounts for animal-based products. Hover time for purchased products only accounts for products that were ultimately purchased. Unit of time is second. The base category is Labeling. The specifications control for demographic variables including, age, gender, education, race, marital status, household size, income, employment status, and region of residence. The standard errors in the parenthesis are corrected for heteroskedasticity.***p<0.05, **p<0.10, *p<0.15. The sample consists of participants from *Labeling* and *Combined* treatments only.

Appendix

Table A1: Plant-based and Animal-based Food Products

Product	Category
Turkey ground	Animal-based meat
Lean ground beef	Animal-based meat
Franks/hot dogs	Animal-based meat
Vegan hot dogs	Plant-based meat
Vegan meatballs	Plant-based meat
Turkey meatballs	Animal-based meat
Vegan sausage	Plant-based meat
Turkey sausage	Animal-based meat
Beef sausage	Animal-based meat
Oat milk	Plant-based milk
Coconut milk	Plant-based milk
Almond milk	Plant-based milk
Whole milk	Animal-based milk
Soy milk	Plant-based milk
Vegan mozzarella cheese	Plant-based cheese
Vegan cheddar cheese	Plant-based cheese
Mozzarella cheese	Animal-based cheese
Cheddar cheese	Animal-based cheese
Vegan yogurt	Plant-based yogurt
Milk yogurt	Animal-based yogurt

Appendix Table A2 T-statistics of total view and hover time of treatment groups compared with control group (N =2405)

	Labeling	Categorization	Combined
View time all products	-1.88**	-0.76	-3.31***
View time plant-based products	-1.51*	-0.19	-1.30
View time animal-based products	-1.45*	-0.88	-3.42***
Hover time total	11.08***	.	9.06***
Hover time plant-based products	10.17***	.	7.80***
Hover time animal-based products	9.31***	.	8.31***

Note: We show the t-statistics of view and hover time of treatment groups compared with the control group. *p < 0.15; **p < 0.10; ***p < 0.5. View time is the total time participants spend viewing different product pages in a category. Hover time is the total time participants spend hovering over the GHG label. The *Control* and *Categorization* groups didn't see labels, so the hover time of these groups isn't measured. Unit of time is second.

References

Amin, M. D., Badruddoza, S., & McCluskey, J. J. (2021). Predicting access to healthful food retailers with machine learning. *Food Policy*, *99*, 101985. https://doi.org/10.1016/j.foodpol.2020.101985

Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online Privacy Concerns and Privacy Management: A Meta-Analytical Review. *Journal of Communication*, *67*(1), 26–53. https://doi.org/10.1111/jcom.12276

Bazzaz Abkenar, S., Haghi Kashani, M., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and Informatics*, *57*, 101517. https://doi.org/10.1016/j.tele.2020.101517

Boland, M., Alam, F., & Bronlund, J. (2019). Chapter 8—Modern Technologies for Personalized Nutrition. In C. M. Galanakis (Ed.), *Trends in Personalized Nutrition* (pp. 195–222). Academic Press. https://doi.org/10.1016/B978-0-12-816403-7.00006-4

Bryant, C. J. (2022). Plant-based animal product alternatives are healthier and more environmentally sustainable than animal products. *Future Foods*, *6*, 100174. https://doi.org/10.1016/j.fufo.2022.100174

Burris, D. (2023). *Why even your local grocery store wants your digital data*. https://www.cnbc.com/video/2023/12/10/why-even-your-local-grocery-store-wants-your-digital-data.html

Caputo, V., & Lusk, J. L. (2022). The Basket-Based Choice Experiment: A Method for Food Demand Policy Analysis. *Food Policy*, *109*, 102252. https://doi.org/10.1016/j.foodpol.2022.102252

Caputo, V., & Scarpa, R. (2022). Methodological Advances in Food Choice Experiments and Modeling: Current Practices, Challenges, and Future Research Directions. *Annual*

Review of Resource Economics, *14*(Volume 14, 2022), 63–90. https://doi.org/10.1146/annurev-resource-111820-023242

Cash, S. B., Alhabash, S., Fretes, G., & Ma, M. (2022). Food, beverages and social media: trends and tools for economic research. In *A Modern Guide to Food Economics* (pp. 297-326). Edward Elgar *Publishing*.

Chandra, S., Verma, S., Lim, W. M., Kumar, S., & Donthu, N. (2022). Personalization in personalized marketing: *Trends* and ways forward. *Psychology & Marketing*, *39*(8), 1529-1562.

Chen, S., Brahma, S., Mackay, J., Cao, C., & Aliakbarian, B. (2020). The role of smart packaging system in food supply chain. *Journal of Food Science*, *85*(3), 517-525.

Choudhury, N., Mukherjee, R., Yadav, R., Liu, Y., & Wang, W. (2024). Can machine learning approaches predict green purchase intention? -A study from Indian consumer perspective. *Journal of Cleaner Production*, *456*, 142218. https://doi.org/10.1016/j.jclepro.2024.142218

Dennis, E. J., Tonsor, G. T., & Lusk, J. L. (2021). Choosing quantities impacts individuals choice, rationality, and willingness to pay estimates. *Agricultural Economics*, *52*(6), 945–962. https://doi.org/10.1111/agec.12678

Dong, Y., Sharma, C., Mehta, A., & Torrico, D. D. (2021). Application of Augmented Reality in the Sensory Evaluation of Yogurts. *Fermentation*, *7*(3), Article 3. https://doi.org/10.3390/fermentation7030147

du Plessis, C. (2022). A Scoping Review of the Effect of Content Marketing on Online Consumer Behavior. *Sage Open*, *12*(2), 21582440221093042. https://doi.org/10.1177/21582440221093042

Ekman, P. (2002). Facial Action Coding System (FACS). *A Human Face*. https://cir.nii.ac.jp/crid/1573105975118295296

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897–904. https://doi.org/10.1016/j.jbusres.2015.07.001

e Silva, B. V. R., Rad, M. G., Cui, J., McCabe, M., & Pan, K. (2018). A mobile-based diet monitoring system for obesity management. *Journal of health & medical informatics*, *9*(2).

Fang, D., Nayga Jr, R. M., West, G. H., Bazzani, C., Yang, W., Lok, B. C., Levy, C. E., & Snell, H. A. (2021). On the Use of Virtual Reality in Mitigating Hypothetical Bias in Choice Experiments. *American Journal of Agricultural Economics*, *103*(1), 142–161. https://doi.org/10.1111/ajae.12118

Fujiwara, M., Moriya, K., Sasaki, W., Fujimoto, M., Arakawa, Y., & Yasumoto, K. (2018). A Smart Fridge for Efficient Foodstuff Management with Weight Sensor and Voice Interface. *Workshop Proceedings of the 47th International Conference on Parallel Processing*, 1–7. https://doi.org/10.1145/3229710.3229727

Heller, J., Chylinski, M., de Ruyter, K., Mahr, D., & Keeling, D. I. (2019). Touching the Untouchable: Exploring Multi-Sensory Augmented Reality in the Context of Online Retailing. *Journal of Retailing*, *95*(4), 219–234. https://doi.org/10.1016/j.jretai.2019.10.008

Howe, H. S., Ubel, P. A., & Fitzsimons, G. J. (2022). Open Science Online Grocery: A Tool for Studying Choice Context and Food Choice. *Journal of the Association for Consumer Research*, *7*(4), 393–402. https://doi.org/10.1086/720449

Hu, Y., House, L. A., & Gao, Z. (2022). How do consumers respond to labels for crispr (gene-editing)? *Food Policy*, *112*, 102366. https://doi.org/10.1016/j.foodpol.2022.102366

Huang, P., Lurie, N. H., & Mitra, S. (2009). Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods. *Journal of Marketing*, *73*(2), 55–69. https://doi.org/10.1509/jmkg.73.2.55

Huang, Y.-Z., Liu, Y., Yu, X.-L., Li, K., Li, G.-D., Zhu, B.-W., & Dong, X.-P. (2023). Rapid screening of sensory attributes of mackerel using big data mining techniques and rapid sensory evaluation methods. *Journal of Texture Studies*, *54*(6), 872–884. https://doi.org/10.1111/jtxs.12776

Huang, Z., & Benyoucef, M. (2013). From e-commerce to social commerce: A close look at design features. *Electronic Commerce Research and Applications*, *12*(4), 246–259. https://doi.org/10.1016/j.elerap.2012.12.003

Innocenti, A. (2017). Virtual reality experiments in economics. *Journal of Behavioral and Experimental Economics*, *69*, 71–77. https://doi.org/10.1016/j.socec.2017.06.001

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

Kouarfaté, B. B., & Durif, F. (2023). Understanding Consumer Attitudes toward Cultured Meat: The Role of Online Media Framing. *Sustainability*, *15*(24), Article 24. https://doi.org/10.3390/su152416879

Li, F., Larimo, J., & Leonidou, L. C. (2021). Social media marketing strategy: Definition, conceptualization, taxonomy, validation, and future agenda. *Journal of the Academy of Marketing Science*, *49*(1), 51–70. https://doi.org/10.1007/s11747-020-00733-3

Li, J., & Hu, Q. (2021). Using culturomics and social media data to characterize wildlife consumption. *Conservation Biology*, *35*(2), 452–459. https://doi.org/10.1111/cobi.13703

Lin, W., Ortega, D. L., & Caputo, V. (2023). Experimental quantity, mental budgeting and food choice: A discrete choice experiment application. *European Review of Agricultural Economics*, *50*(2), 457–496. https://doi.org/10.1093/erae/jbac017

Liu, Y., & Lopez, R. A. (2016). The impact of social media conversations on consumer brand choices. *Marketing Letters*, *27*(1), 1–13. https://doi.org/10.1007/s11002-014-9321-2

Lizin, S., Rousseau, S., Kessels, R., Meulders, M., Pepermans, G., Speelman, S., Vandebroek, M., Van Den Broeck, G., Van Loo, E. J., & Verbeke, W. (2022). The state of the art of discrete choice experiments in food research. *Food Quality and Preference*, *102*, 104678. https://doi.org/10.1016/j.foodqual.2022.104678

Lu, H., Guo, L., Azimi, M., & Huang, K. (2019). Oil and Gas 4.0 era: A systematic review and outlook. *Computers in Industry*, *111*, 68–90. https://doi.org/10.1016/j.compind.2019.06.007

Ma, T., Wang, H., Wei, M., Lan, T., Wang, J., Bao, S., ... & Sun, X. (2022). Application of smart-phone use in rapid food detection, food traceability systems, and personalized diet guidance, making our diet more health. *Food Research International*, *152*, 110918.

Marston, J. (2017). Quick-service restaurants are quickly turning to facial recognition. *The Spoon*.

Mellouk, W., & Handouzi, W. (2020). Facial emotion recognition using deep learning: Review and insights. *Procedia Computer Science*, *175*, 689–694. https://doi.org/10.1016/j.procs.2020.07.101

G. Mitchell, E., M. Heitkemper, E., Burgermaster, M., E. Levine, M., Miao, Y., L. Hwang, M., ... & Mamykina, L. (2021, May). From reflection to action: Combining machine learning with expert knowledge for nutrition goal recommendations. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1-17).

Mordor Intelligence, (2020). Facial recognition market–growth, trends, and forecasts (2020–2025). https://www.mordorintelligence.com/industry-reports/facial-recognition-market Accessed on 20-June-2024

Moe, W. W. (2003). Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream. *Journal of Consumer Psychology*, *13*(1), 29–39. https://doi.org/10.1207/S15327663JCP13-1&2_03

Mogaji, E., Balakrishnan, J., & Kieu, T. A. (2021). Examining consumer behaviour in the UK Energy sector through the sentimental and thematic analysis of tweets. *Journal of Consumer Behaviour*, *20*(2), 218–230. https://doi.org/10.1002/cb.1820

Mokas, I., Lizin, S., Brijs, T., Witters, N., & Malina, R. (2021). Can immersive virtual reality increase respondents' certainty in discrete choice experiments? A comparison with traditional presentation formats. *Journal of Environmental Economics and Management*, *109*, 102509. https://doi.org/10.1016/j.jeem.2021.102509

Mol, J. M. (2019). Goggles in the lab: Economic experiments in immersive virtual environments. *Journal of Behavioral and Experimental Economics*, *79*, 155–164. https://doi.org/10.1016/j.socec.2019.02.007

Pancer, E., Philp, M., & Noseworthy, T. J. (2022). Boosting engagement with healthy food on social media. *European Journal of Marketing*, *56*(11), 3007–3031. https://doi.org/10.1108/EJM-07-2021-0565

Philp, M., Jacobson, J., & Pancer, E. (2022). Predicting social media engagement with computer vision: An examination of food marketing on Instagram. *Journal of Business Research*, *149*, 736–747. https://doi.org/10.1016/j.jbusres.2022.05.078

Pieters, R., & Wedel, M. (2007). Goal Control of Attention to Advertising: The Yarbus Implication. *Journal of Consumer Research*, *34*(2), 224–233. https://doi.org/10.1086/519150

Rajagopal, A. (2023). *Half of all grocers are upping their shopper data game: Report*. https://www.supermarketnews.com/online-retail/half-all-grocers-are-upping-their-shopper-data-game-report

Rankin, B. (2017). What it's like to order poke using facial-recognition software. *Dallas Observer*.

Redman, R. (2021). *Amazon to nearly double online edible grocery sales by 2026*. https://www.supermarketnews.com/online-retail/amazon-nearly-double-online-edible-grocery-sales-2026

Rini, L., Schouteten, J. J., Faber, I., Bom Frøst, M., J A Perez-Cueto, F., & De Steur, H. (2024). Social media and food consumer behavior: A systematic review. *Trends in Food Science & Technology*, *143*, 104290. https://doi.org/10.1016/j.tifs.2023.104290

Samoggia, A., Riedel, B., & Ruggeri, A. (2020). Social media exploration for understanding food product attributes perception: The case of coffee and health with Twitter data. *British Food Journal*, *122*(12), 3815–3835. https://doi.org/10.1108/BFJ-03-2019-0172

Santini, F. de O., Ladeira, W. J., Pinto, D. C., Herter, M. M., Sampaio, C. H., & Babin, B. J. (2020). Customer engagement in social media: A framework and meta-analysis. *Journal of the Academy of Marketing Science*, *48*(6), 1211–1228. https://doi.org/10.1007/s11747-020-00731-5

Shen, Y., Hamm, J. A., Gao, F., Ryser, E. T., & Zhang, W. (2021). Assessing consumer buy and pay preferences for labeled food products with statistical and machine learning methods. *Journal of Food Protection*, *84*(9), 1560-1566.

Singh, A., & Glińska-Neweś, A. (2022). Modeling the public attitude towards organic foods: A big data and text mining approach. *Journal of Big Data*, *9*(1), 2. https://doi.org/10.1186/s40537-021-00551-6

Sundaravadivel, P., Kesavan, K., Kesavan, L., Mohanty, S. P., & Kougianos, E. (2018). Smart-Log: A Deep-Learning Based Automated Nutrition Monitoring System in the IoT. *IEEE Transactions on Consumer Electronics*, *64*(3), 390–398. https://doi.org/10.1109/TCE.2018.2867802

Titova, J., Cottis, G., & Allman-Farinelli, M. (2023). Using social media analysis to study population dietary behaviours: A scoping review. *Journal of Human Nutrition and Dietetics*, *36*(3), 875–904. https://doi.org/10.1111/jhn.13077

Tumasjan, A. (2024). The many faces of social media in business and economics research: Taking stock of the literature and looking into the future. *Journal of Economic Surveys*, *38*(2), 389–426. https://doi.org/10.1111/joes.12570

U.S. Census Bureau. (2023). *Census Bureau Data*. https://data.census.gov/

Vrontis, D., Makrides, A., Christofi, M., & Thrassou, A. (2021). Social media influencer marketing: A systematic review, integrative framework and future research agenda. *International Journal of Consumer Studies*, *45*(4), 617–644. https://doi.org/10.1111/ijcs.12647

Wang, Q., Yang, S., Liu, M., Cao, Z., & Ma, Q. (2014). An eye-tracking study of website complexity from cognitive load perspective. *Decision Support Systems*, *62*, 1–10. https://doi.org/10.1016/j.dss.2014.02.007

Wang, R. J.-H., Malthouse, E. C., & Krishnamurthi, L. (2015). On the Go: How Mobile Shopping Affects Customer Purchase Behavior. Journal *of Retailing*, *91*(2), 217–234. https://doi.org/10.1016/j.jretai.2015.01.002

Wang, Z., Xue, M., Wang, Y., Song, M., Li, S., Daziano, R. A., ... & Zhang, B. (2019). Big data: New tend to sustainable consumption research. *Journal of Cleaner Production*, *236*, 117499.

Wedel, M., & Kannan, P. K. (2016). Marketing Analytics for Data-Rich Environments. *Journal of Marketing*, *80*(6), 97–121. https://doi.org/10.1509/jm.15.0413

Xi, N., & Hamari, J. (2021). Shopping in virtual reality: A literature review and future agenda. *Journal of Business Research*, *134*, 37–58. https://doi.org/10.1016/j.jbusres.2021.04.075

Xu, Q. A., Chang, V., & Jayne, C. (2022). A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decision Analytics Journal*, *3*, 100073. https://doi.org/10.1016/j.dajour.2022.100073

Yang, K., & Banamah, A. (2014). Quota Sampling as an Alternative to Probability Sampling? An Experimental Study. *Sociological Research Online*, *19*(1), 56–66. https://doi.org/10.5153/sro.3199

Zachlod, C., Samuel, O., Ochsner, A., & Werthmüller, S. (2022). Analytics of social media data – State of characteristics and application. *Journal of Business Research*, *144*, 1064–1076. https://doi.org/10.1016/j.jbusres.2022.02.016

Zhao, S., & Yue, C. (2020). Risk preferences of commodity crop producers and specialty crop producers: An application of prospect theory. *Agricultural Economics*, *51*(3), 359–372. https://doi.org/10.1111/agec.12559

Zhou, M., Chen, G. H., Ferreira, P., & Smith, M. D. (2021). Consumer behavior in the online classroom: Using video analytics and machine learning to understand the consumption of video courseware. *Journal of Marketing Research*, *58*(6), 1079-1100.

Zuo, J., Feng, J., Gameiro, M. G., Tian, Y., Liang, J., Wang, Y., ... & He, Q. (2022). RFID-based sensing in smart packaging for food applications: A review. *Future Foods*, *6*, 100198.