Robertson, D., Andersson, Y., and Lunn, P. (2023). How consumer and provider responses to nutritional labelling interact: An online shopping experiment with implications for policy, *Food Policy,* Volume 121, November 2023, https://doi.org/10.1016/j.foodpol.2023.102563

How consumer and provider responses to nutritional labelling interact: an online shopping experiment with implications for policy

Deirdre A. Robertson,[a,b*] Ylva Andersson[a] & Peter D. Lunn[a,c]

a. Behavioural Research Unit, Economic and Social Research Institute, Dublin, Ireland.
b. School of Psychology, Trinity College Dublin, Ireland.
c. Department of Economics, Trinity College Dublin, Ireland.

*Address for Correspondence
Dr. Deirdre A. Robertson
Economic and Social Research Institute
Whitaker Square
Sir John Rogerson Quay
Dublin
Ireland
Email: Deirdre.robertson@esri.ie

Acknowledgements: We thank members of the Research Services and Policy Unit in the Department of Health and members of the steering group of the Department of Health/ESRI Research Programme in Healthcare Reform for feedback. We also thank other members of the Behavioural Research Unit for input during the development of this research.

Funding: This research was funded by the Government of Ireland Department of Health/ESRI Research Programme in Healthcare Reform.

Disclosure Statement: The authors report there are no competing interests to declare.

Abstract

Food labelling policies are usually conceptualised as a way to inform consumers about nutritional content of food. Although often unstated, a secondary aim is to encourage industry to reformulate recipes or introduce healthier alternatives. Parallel bodies of research examine how consumers and industry respond to food labelling policies. In this study we explored the interaction between provider and consumer responses by recording purchases under different assumptions about the impact of a label on product ranges. We simulated different online food markets and tested the effects of a food label, Nutri-Score, on incentivised consumer decisions. Consumers who were exposed to Nutri-Scores applied to snack products made healthier purchases, on average, than consumers who were not. Consumers who shopped in a market adapted to provide more healthy options made healthier purchases than those who shopped in the current market. These effects were additive: consumers who were exposed to Nutri-Scores on products when shopping in the adapted market made the healthiest choices. In a subsequent choice task, a market that simulated reformulation had a stronger effect on choices than one that merely added healthier options. The findings hence offer insight into the benefits of labelling and may be useful for informing both policy and the dialogue between policymakers and industry.

Keywords: Food choice; Nutritional Labels; Nutri-Score; Industry; Policy; Availability; Consumer Behaviour

1. Introduction

Food labelling policies are often billed as a way to help consumers make informed choices.[1] A large body of evidence suggests that food labels do inform consumers and do influence choice, mostly guiding consumers towards lower-energy or more nutritional products (Roberto et al., 2021). However, a secondary and sometimes unstated aim of policy is to encourage food businesses to modify the options available to consumers. To attain a better rating on labels, businesses can either reformulate existing products or introduce new products to a range. A recent meta-analysis concluded that food labelling interventions reduced trans-fat by 64% and sodium by 9% through reformulation of the options available (Shangguan et al., 2019). Policy evaluations in many countries have shown that when voluntary labelling policies are implemented, many of the products that are awarded the label have been reformulated (Roberto et al., 2021). Although the response can vary, food labelling policies can and do incentivise an industry response (De Marchi et al., 2023; Vandevijvere and Vanderlee, 2019).

Research tracks investigating the effect of food labels on consumer choice and on industry behaviour largely proceed in parallel, without considering the potential interactions between the two. Most consumer research tests how a food label would influence consumer behaviour if it were to be applied in the existing market. However, by the time most food labelling policies are in place, the market may have changed due to provider response to the forthcoming policy. This is important when evaluating the effect of food labels on choice because research from psychology shows that choice is not stable; consumer choices are influenced by changes in contexts. Influential contextual factors include the number of available options, the salience of some products and the attributes of alternative options (Dai et al., 2020; Fasolo et al., 2009; Grech and Allman-Farinelli, 2015; Lurie, 2004). For example, when new and healthier products are added to an existing range, consumers are more likely to purchase them, even though the original options remain available (Grech and Allman-Farinelli, 2015; Pechey et al., 2019; Pechey and Marteau, 2018). An industry response to a food label may change the product range in three ways (Pechey et al., 2020):

(1) new products are added or old products are removed so that the absolute number of products changes but the proportion of healthier products stays the same. For example a new food provider could enter the market offering both healthy and unhealthy products;
(2) old products are removed or replaced with new products so that the proportion of healthier options changes but the absolute number of products stays the same. For example, an existing food provider could reformulate their products or replace older products with new healthier products;
(3) new products are added to an existing range so that both the absolute number of products and the proportion of healthier options changes. For example, a provider could create new products with better ratings on the food label but continue to offer the original products.

There is some evidence that all three changes to a range can influence consumers to make healthier choices, but no studies that we are aware of have directly compared the influence on choice (Pechey et al., 2020). Understanding how consumers respond to different industry responses is important for informing conversations between industry and policymakers.

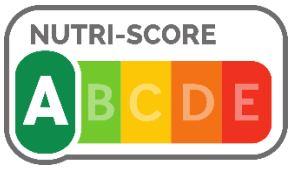
[1] Examples include the USA Food and Drink Administration's added sugar label and the European Union's regulation on food information provision.

In sum, a strand of research shows that good nutritional labelling affects consumer behaviour, a second strand demonstrates a relationship between nutritional labelling policies and industry response, and a third shows how changes to product ranges influence consumer behaviour. There are not, to our knowledge, investigations of how the impact of food labels on consumer behaviour is affected by how industry alters the product range. It is possible that food labels are particularly useful for signalling the nutritional content of new or reformulated products, but it is also possible that food labels may no longer influence consumers once an industry response has provided healthier alternatives. The aim of this paper is to investigate how the influence of food labels may differ depending on the choices available to consumers.

Nutri-Score

The food label we tested is the Nutri-Score (Figure 1). Nutri-Score is a nutritional label based on the British Food Standards Agency (FSA) nutritional score. It is a colour-coded 5-point scale (ranging from A to E) that is intended to indicate the healthfulness of a food product. The underlying FSA score ranges from -15 to 40 with lower scores indicating a healthier product. Points are added when products contain sugar, salt and saturated fat. Points are taken away when products contain fruits, vegetables, nuts, rapeseed oil, walnut oil or olive oil, and more fibre and protein in general. There is high consistency between Nutri-Score and nutritional recommendations (Dréano-Trécant et al., 2020; Hafner and Pravst, 2021; Szabo de Edelenyi et al., 2019).

Figure 1. Example of the Nutri-Score showing an 'A' rating.



Previous work on consumers across 12 countries suggests that Nutri-Score is understood better than similar labels such as traffic light labels, reference intake labels, the Health Star Rating system, and the Warning symbols system (Egnell et al., 2018). Laboratory trials and field trials have shown that putting Nutri-Score labels on foods changes both intended and actual purchases with consumers being more likely to purchase foods with lower (i.e. healthier) Nutri-Score ratings (Crosetto et al., 2019; Dubois et al., 2021; Ducrot et al., 2016; Julia and Hercberg, 2017). However, Nutri-Score is controversial in the public sphere in some European countries (Fialon et al., 2022). Opponents say that Nutri-Score discriminates against traditional Mediterranean foods that cannot be reformulated and that it does not factor in portion size (Follis, 2020; FoodNavigator, 2020).

This study investigated Nutri-Score in Ireland, a market where it had not yet been implemented but was being considered as a voluntary labelling policy. The research was funded by the Government's Department of Health. The primary research questions were: (1) Does Nutri-Score influence purchasing decisions?; (2) Does a simulated industry response that changes the availability of options influence purchasing decisions?; (3) Does adding new options or replacing old options have a greater influence on choice?; (4) Are effects of Nutri-Score and availability additive, i.e. does Nutri-Score influence purchasing decisions regardless of the availability of different options? A full list of specific, pre-registered hypotheses is provided in supplementary material (Table S2). Both the pre-registration and data are available on the Open Science Framework.[2]

[2] https://osf.io/nqs3f/?view_only=47b69ab6421445809ecc30116279885c

2. Materials and Methods

2.1. Participants

Participants (N = 800) were recruited by a market research company to take part in a 20-minute study on consumer behaviour in an online shopping environment. The sample was designed to be approximately representative of the national population of Ireland based on quotas of observable characteristics. A socio-demographic breakdown is provided in Section 3. Participants were paid €4 for taking part. To incentivise participants to choose products they actually wanted, they were told at the start that they may be selected to receive the products they chose. Fifty participants were randomly selected to receive their products. The products were shipped after the study period ended.

2.2. Materials

We created a database of the prices and nutritional content of all products within eight categories of snack foods from one of the largest supermarket chains in Ireland. The eight categories were sweet biscuits/cookies, crackers, cereal bars/protein bars, chocolate, sweets/mints, popcorn, nuts, and crisps/chips. We used the nutritional information provided by the supermarket chain to calculate the British Food Standards Agency nutrient profiling score (FSA score) which was then translated into the corresponding Nutri-Score. The components necessary to calculate the score were kilojoules, saturated fatty acids, sugars, proteins, fibres, sodium, and fruits, vegetables, pulses, nuts, rapeseed, walnut and olive oils per 100g of product. Information on each of these components is available on the nutritional labels or ingredient lists of products in Ireland. We used this database to generate distributions of Nutri-Scores for products currently in the Irish market. We then compared these distributions to the distributions in one of the largest French supermarket chains, where Nutri-Score has been in place as a voluntary label since 2017. The French distributions were generally shifted to the left of the distributions in the Irish market, indicating that the French market consists of products with lower Nutri-Scores (see Figures S1 and S2 for more detail). While there may be multiple reasons for this international difference in the distributions of scores, matching distributions to the same product categories in the French supermarket chain allowed us to test distributions that arise in a market in which Nutri-Score already exists and where some manufacturers may have either already reformulated products or introduced new products with lower Nutri-Scores.

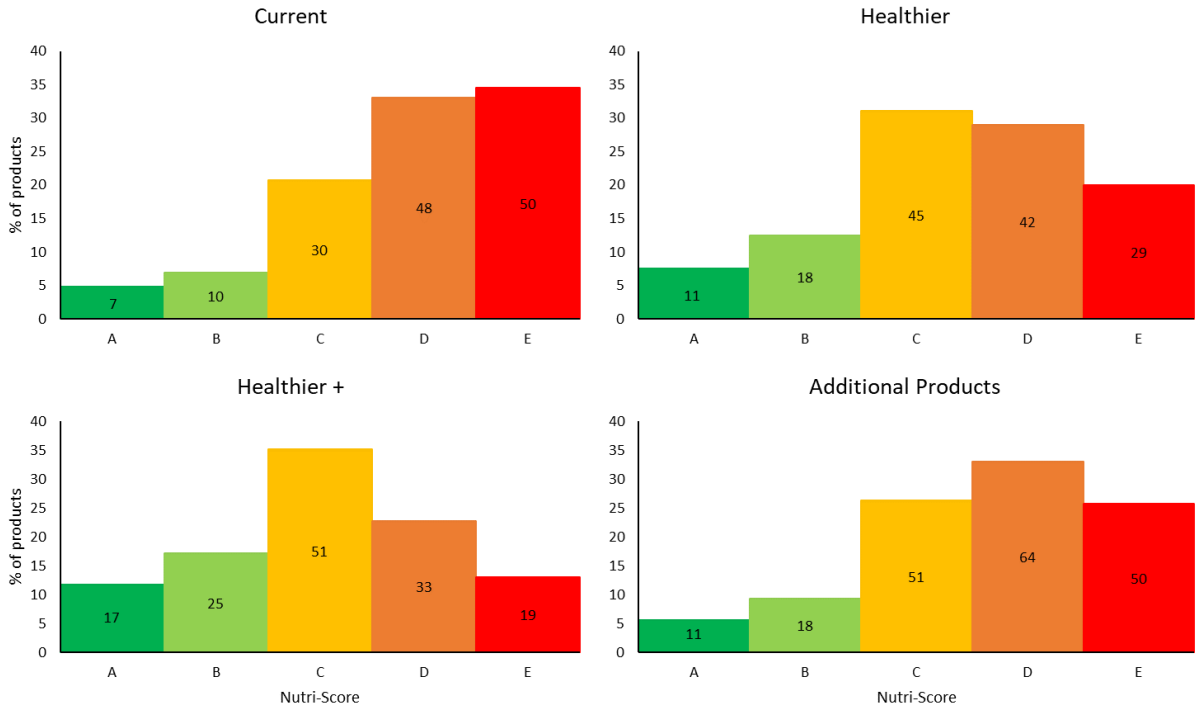
We then selected subsets of products from the Irish supermarket chain to create four different simulated markets (See Table 1):

1) A market representative of the current Irish market (“current” market);
2) A market representative of the current French market; i.e. with a higher proportion of products with lower FSA scores than in the Irish market (“healthier” market);
3) A market with a higher proportion of products with lower FSA scores than currently exists in the Irish or French market (“healthier+” market);
4) A market created by adding new healthier products to the current distribution (“additional products” market).

Figure 1 shows the percentage of Nutri-Scores across all food categories in each of the four simulated markets. There was significant overlap (66%) of products shown in the “current” and “healthier” market, which were the two markets shown in the incentivised shopping task. The “healthier+” market had 21% of the same products as in the “current” market. We created the markets using only products already sold in the Irish supermarket to avoid introducing products that

consumers were not familiar with. Markets 1-3 had 145 products each to choose from across 8 categories. Market 4 had 194 products.

Figure 1. Percentage of Nutri-Score category across all products in the four simulated markets.



Note. A-E indicate the Nutri-Score of the product. A is considered to have the most preferable nutritional score, and E the least. The y axis shows percentage of products and the values inside the bars show number of products.

2.3 Online shop

The online shop was programmed using Gorilla (www.gorilla.sc) (Anwyl-Irvine et al., 2020). Participants carried out two shopping tasks. One that was incentivised and one that was hypothetical.

2.3.1 Incentivised shop

Participants were told in advance that they had been given a €10 voucher to spend on anything in the shop. They could spend as much or as little of the voucher as they wished but anything unspent was not redeemable for cash. They were also told that they may be selected to receive the products they purchased and that they should select products they would like. Figure 2 shows a screenshot of the online shop. The eight categories of products were shown at the top of the page. Participants could move freely between categories. The order of categories and the order of products within categories were randomised between participants. All participants saw a picture of each product along with price, weight and, if they were in the Nutri-Score condition, Nutri-Score. Further information on each product was available by clicking on the individual product. This brought up a pop-up window showing a longer description of the product, a full ingredients list, the nutritional information table and, if they were in the Nutri-Score condition, the Nutri-Score and an explanation of Nutri-Score. We chose to include information on Nutri-Score because this label has never been introduced to the Irish market and we reasoned that, were it to be introduced, consumers would be given information on it. This information was not made salient during the shop and was only available if consumers clicked on the product to get more information. This layout was modelled on

the online shops of the big supermarket chains in Ireland in which such boxes contain information from the manufacturers and occasionally warnings. A box on the side of the screen showed participants the content of their basket and the total price. They could add and remove items from their basket as many times as they chose. They were able to checkout at any time after selecting at least one product. The prices displayed were the prices of products in the supermarket at the time of data collection. The prices of the replacement products in the healthier markets were matched to those in the current market. The average price of products was €2.08. There was no relationship between price and FSA score in either the current shop (Spearman's Rho = -0.0004, p = .99) or in the healthier shop (Spearman's Rho = 0.01, p = 0.87)

2.3.2 Hypothetical shop

In the hypothetical shopping task, participants again saw a version of the online shop, but this time were asked to choose one product they would buy from each of the eight categories, if they had to choose one. Participants saw one category at a time and could only move on from that category once they had selected one product.

2.4 Experimental Manipulations

Before the incentivised shop began, participants were randomised to a control condition (no Nutri-Score) in which they saw only the normal nutritional information table on products, or to an intervention condition (Nutri-Score) in which Nutri-Score was also shown on the products (see Figure 2). Participants were not aware that they had been assigned to an experimental condition, nor that others had been assigned to a different one. In the first shopping task, they were randomised to shop from the current market or the healthier market (Market 1 or 2, Table 1). Participants in all conditions chose from the same number of products.

In the hypothetical shop, participants remained in the same control or Nutri-Score condition but the market they shopped from was randomised within-person at the category level. In this task, participants were asked to choose their favourite product from each of the eight categories and each product category had one of the four possible distributions (Table 1): (1) current market; (2) healthier market; (3) healthier+ and (4) additional products market. As well as seeking to replicate any effects from the incentivised shopping task, this within-person experimental design allowed us to test for differences in the effects of markets (3) and (4) while retaining statistical power.

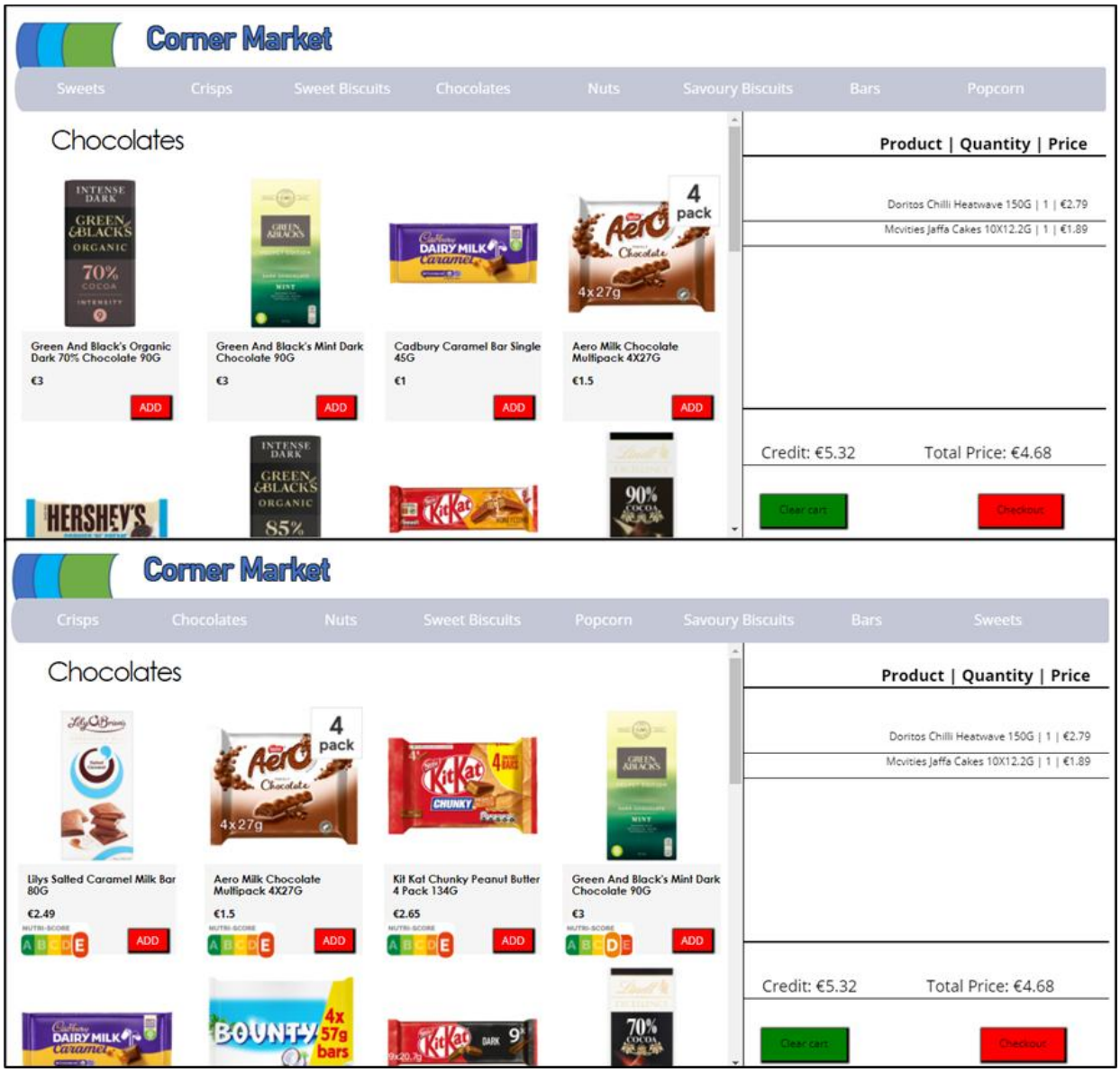
Table 1. Descriptions of the four markets.

Market	Description	Type of change from current market
1. Current	Subset of products currently available in the market.	
2. Healthier	As "current" market, but some products replaced with healthier alternatives to match distribution of Nutri-Scores in the French market.	Relative increase in healthier products.
3. Healthier+	As "healthier" market, but even more products replaced with healthier alternatives.	Larger relative increase in healthier products.
4. Additional products	As current market, but replacement products from "healthier" market added.	Absolute and relative increase in healthier products.

2.5 Additional Measures

Before the first shopping task, participants were asked how often they shop for groceries online. After the two shopping tasks, participants responded to some additional questions. The first question asked participants to say what they thought the study aims were. Participants were then asked how often they purchase food from the eight categories in the shop and to rate how much attention they had paid to the nutritional information while doing the shopping task. They were then shown a list of 16 products that had been randomly selected from the 96 products that were common across markets, i.e. products they had previously seen. For each product, they were asked to guess the correct Nutri-Score. After this, they were asked to rate, on a scale of 1-7, how much attention they normally pay to taste, value, healthfulness and familiarity with the brand of products when shopping for groceries like the ones they had seen in the shop. Finally, we collected socio-demographic information, including age, gender, educational attainment, employment status, social grade and whether they had a child under the age of 18 in their household.

Figure 2. Screenshots taken from the online shop for the no Nutri-Score (top) and Nutri-Score (bottom) conditions.



2.6 Procedure

Participants took part online. They were told that the purpose of the study was to assess consumer behaviour in an online shopping environment, that they would be given €10 to spend in an online shop and that they may be selected to have their products delivered to them. Participants were asked to choose at least one product but could choose as few or as many other products as they liked within the budget.

After completing the incentivised shop, participants were asked to do the second shopping task, told that their choices were hypothetical and that they should choose the product from each category that they would buy if they had to buy one. Participants made 8 decisions, two from each of the 4 possible markets.

Participants then completed the questionnaire. They were told if they had been chosen to receive the products they had selected and were asked for their contact details if they wished to receive them. The study complied with institutional ethics policy, including data protection procedures.

2.7 Data Analysis Plan

We pre-registered our analysis plan prior to data collection. To test the effect of experimental condition on the mean FSA score of baskets in the incentivised shop, we ran the following linear regression:

Equation 1. $$avFSA_i = \alpha + \beta_1 NS_i + \beta_2 Market_i + \beta_3 X_i + \varepsilon_i$$

avFSA_i is the mean FSA score of individual *i*'s basket, NS_i denotes individual *i*'s assignment to the Nutri-Score or control condition, and $Market_i$ denotes individual *i*'s assignment to the current market or the healthier market. Following initial estimation, X_i, a vector of control variables for individual *i* was added to the model.

In a third step, to test for mediating effects, we added two variables to the model: self-reported attention paid to nutritional information during the shop and the number of correct guesses of the Nutri-Score categories of products (out of 16). As an additional exploratory analysis, we ran a structural equation model to test formally for mediation effects (see below).

The hypothetical shopping task was analysed using a multilevel linear regression:

Equation 2. $$FSA_{ij} = \alpha + \beta_1 NS_i + \beta_2 Market_{ij} + \beta_3 Category_{ij} + \beta_4 X_i + \mu_i + \varepsilon_i$$

where FSA_{ij} is the FSA score of the item chosen by individual *i* on their *j*th decision, NS_i is the Nutri-Score condition individual *i* was assigned to, $Market_{ij}$ is the market condition individual *i* was assigned to on their *j*th decision and X_i is a vector of control variables for individual *i*. $Category_{ij}$ is a set of dummy variables denoting which of the eight food categories the item was chosen from for individual *i* on their *j*th decision and μ_i is a random effect for individual *i*.

To test for an effect of experimental condition on the number of items bought in the incentivised shop we ran the following ordinal logistic regression:

Equation 3. $$logit\big(P(no.of\ items \leq k)\big) = \alpha_k + \beta_1 NS_i + \beta_2 Market_i + \beta_3 X_i + \varepsilon_i$$

where $P(no.of\ items \leq k)$ is the probability that the number of items bought by individual *i* is less than or equal to k *and NS*, *Market*, and *X* are as in Equation 1.

To test the effect of experimental condition on the weight of products bought in the incentivised shop, we used identical models to those specified in Equations 1 and 2, except that the dependent variable was the mean weight of items bought or chosen respectively.

3. Results

Socio-demographic characteristics by condition are shown in Table 3. There were no statistically significant differences between conditions. Of the 50 people who were given the opportunity to receive their products, only 4 chose not to.

Table 3. Sociodemographic characteristics of conditions[3]

	No Nutri-Score + Control Market	**Nutri-Score + Control Market**	**No Nutri-Score + Healthier Market**	**Nutri-Score + Healthier Market**
Age (Mean, SD)	47.85 (15.72)	51.43 (16.07)	52.22 (15.02)	49.58 (15.52)
Gender (% male)	54.7%	62.5%	54.0%	51.4%
Employed	57.0%	55.9%	59.4%	60.6%
Degree +	41.6%	36.6%	36.9%	41.3%
Social Grade (ABC1)[a]	45.8%	50%	47.6%	48.8%
Child under 18	29.0%	25.3%	23.5%	29.6%
N	214	186	187	213

a. Social grade is a standard socio-economic classification that has six categories (A, B, C1, C2, D and E) and is based on the occupation of the chief income earner of the household.

3.1. Incentivised Purchasing Decisions

Our main dependent variable was the average FSA score of each person's basket. Thus, lower scores indicate baskets with better Nutri-Scores on average.

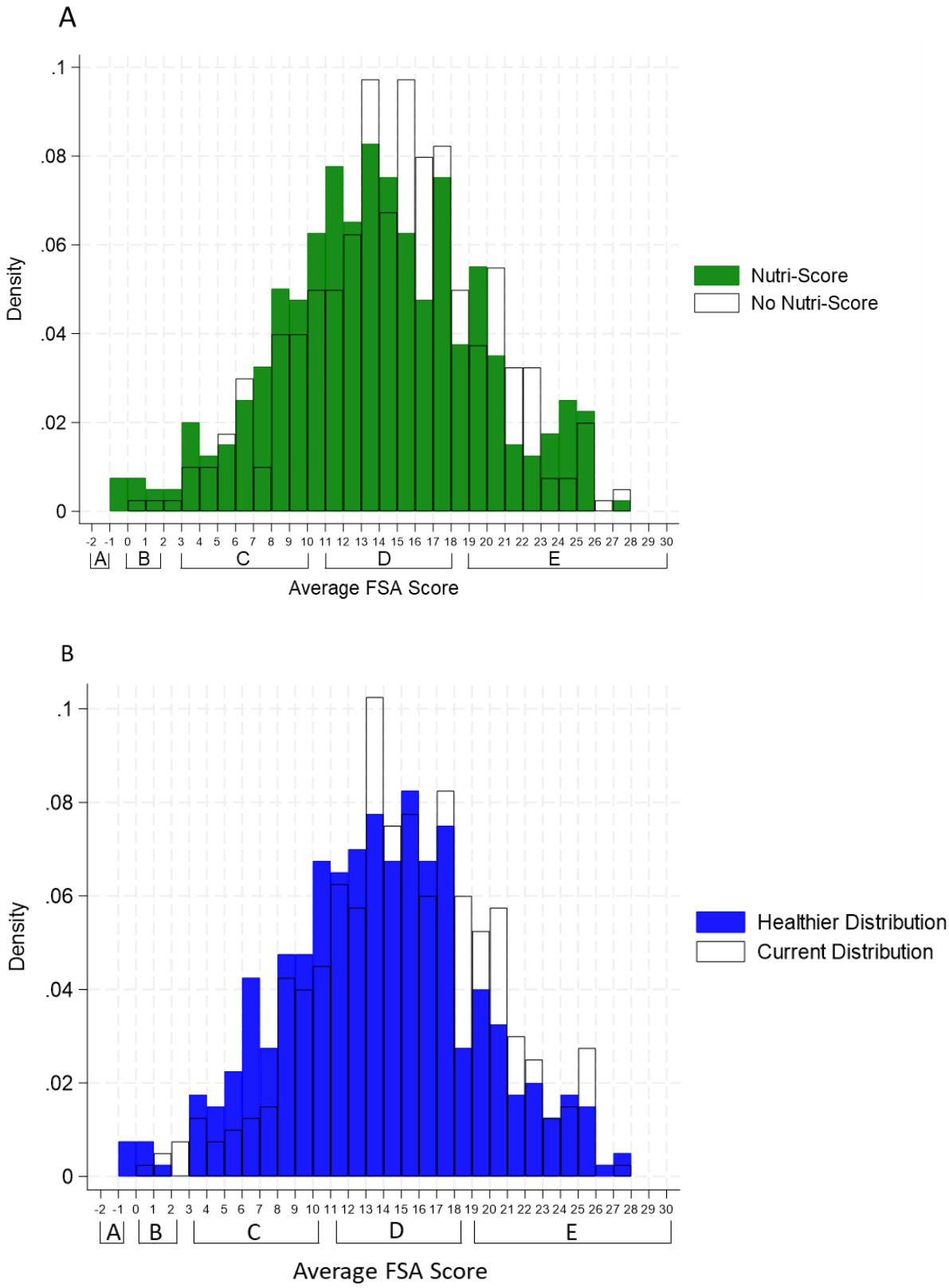
The mean FSA score of shopping baskets was 14.33 (SD = 5.26) with a range from -0.5 to 28. Participants bought an average of 5.42 products (SD = 1.56) with a range of 1-17. Most spent almost the maximum of €10 (M = 9.34, SD = 1.34, range 1-10).

3.1.1 *Does Nutri-Score or the distribution of products in the market influence the nutritional content of baskets?*

Figure 3 shows the distribution of mean FSA scores by condition. Participants who were exposed to Nutri-Scores had lower (healthier) FSA scores relative to those who were not exposed to Nutri-Scores. Similarly, participants who shopped from the healthier market compared to the current market had shopping baskets with lower mean FSA scores.

[3] See Table S1 in supplementary material for a breakdown of additional variables by condition.

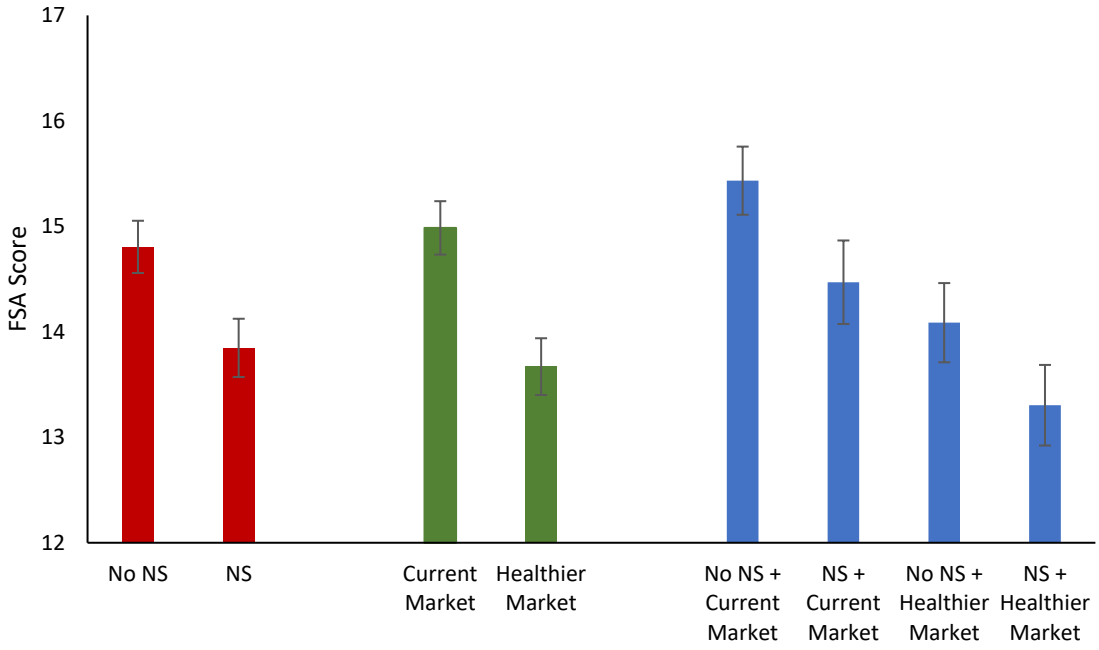
Figure 3. Distribution of mean FSA score of baskets for a) Nutri-Score vs no Nutri-Score and b) current market vs. healthier market.



Note. The letters underneath show Nutri-Scores, which are a categorisation of the continuous FSA score.

Figure 4 shows the mean FSA score of baskets across conditions. The final four bars of Figure 4 show that the healthiest purchases were made by participants who both were exposed to Nutri-Score labels and shopped from the healthier market.

Figure 4. Mean FSA score of baskets for Nutri-Score vs no Nutri-Score, current market compared to the healthier market, and all four conditions separately.



Note. No NS = No Nutri-Score shown, NS = Nutri-Score shown. Error bars represent standard error of the mean. Vertical axis scales to approximately 1 SD (5.26) in line with the recommendation from Witt (2019).

We tested for differences in FSA score between conditions using linear regression. In line with our hypotheses, participants who were exposed to Nutri-Score and participants who shopped in the healthier market had baskets with lower FSA scores (Table 2, Model 1). These main effects were statistically significant. Compared to the control group who shopped in the current market without Nutri-Scores, the three other groups each had significantly lower scores (Table 2, Model 2). We undertook equivalence tests of coefficients. Participants who shopped in a healthier market with Nutri-Scores had baskets with lower FSA scores than participants who shopped in the current market with Nutri-Scores, $F(1, 796) = 4.96$, $p = .01$, one-tailed, Cohen's $d = 0.21$. Participants who were exposed to Nutri-Scores in the healthier market chose baskets with lower FSA scores compared to participants who shopped in the same market but were not exposed to Nutri-Scores, $F(1, 796) = 2.24$, $p = 0.07$, one-tailed, Cohen's $d = 0.15$. Cohen's d effect size was 0.41 for the difference between those who shopped in the current market without Nutri-Scores compared to those who shopped in the healthier market with Nutri-Scores.[4] There was no evidence of an interaction between being exposed to Nutri-Score and the type of market (Table 2, Model 3). Taken together, this pattern of results suggests that the effect of the healthier market and Nutri-Scores on choices was additive.

[4] For a breakdown of the types of percentage of products purchased from each category by condition see Figure S5 in supplementary material.

Table 2. Linear regression analyses with the mean FSA scores of baskets as the dependent variable and condition as the independent variables.

	Model 1 B(SE)	Model 2 B (SE)	Model 3 B (SE)	Model 4 B (SE)
Nutri-Score (ref. No Nutri-Score)	**-0.87 (0.37)****		**-0.96 (0.52)***	**-0.44 (0.36)**
Healthier Market (ref. Current Market)	**-1.26 (0.37)*****		**-1.35 (0.52)****	**-1.27 (0.35)*****
Nutri-Score*Healthier Market			0.18 (0.74)	
Market + NS (ref. No NS + Current Market)				
NS + Current Market		**-0.96 (0.52)***		
No NS + Healthier Market		**-1.35 (0.52)****		
NS + Healthier Market		**-2.13 (0.50)*****		
Health consciousness				-0.53 (0.12)***
Paid attention to nutritional information during shop				-0.41 (0.11)***
Knowledge of Nutri-Scores				-0.23 (0.08)**
Socio-demographic controls?	No	No	No	Yes
N	800	800	800	797
R^2	0.02	0.02	0.02	0.12

Note. No NS = No Nutri-Score shown, NS = Nutri-Score shown. Model 4 and 5 controlled for Age, Gender, Degree, Employment, Child under 18, Social Grade (ABC1, C2DE, F). *p < .05, ** p < .01, *** p < .001, one-tailed in **bold**.

We ran additional pre-registered checks. We controlled for how much participants said they *normally* pay attention to health when shopping on a 1-7 scale, henceforth referred to as health consciousness. Participants who were more health conscious had baskets with lower FSA scores. The effect of exposure to Nutri-Score and the market remained statistically significant after controlling for health consciousness. We ran an additional model to check whether health consciousness moderated the effect of exposure to Nutri-Score on purchases. We found a statistically significant interaction such that exposure to Nutri-Score mainly had an effect on the decisions of those who were more health conscious (Table S3, Figure S3).

We then tested whether exposure to Nutri-Score increased knowledge of the nutritional content of products and attention paid to nutritional information during the shop.

When participants guessed the Nutri-Score of 16 products after the shopping tasks, the mean number of correct guesses was just 5.13 (SD = 2.17) products, with no score above 12/16. Of the 12,800 total guesses, 41% corresponded to a lower (healthier) Nutri-Score than in reality, while 27% corresponded to an underestimation (higher Nutri-Score). As hypothesised, participants who were exposed to Nutri-Score in the online shopping tasks provided more correct estimates, averaging 5.32 compared to 4.95 for those in the control condition, Z(-2.55), p = .005, one-tailed.

Also as hypothesised, participants who were exposed to Nutri-Score reported paying more attention to nutritional information during the shop on a 7-point scale (M = 3.57, SD = 1.98) compared to participants in the control condition (M = 2.93, SD = 1.88), Z(-4.66), p < .001 one-tailed. This difference was unaffected by excluding those who guessed that nutritional labelling was part of the experiment aims.

Next, we assessed whether the influence of being exposed to Nutri-Score labels on FSA scores of products purchased was mediated by knowledge of Nutri-Scores, as assessed by correct guesses of

Nutri-Scores, or by paying attention to nutritional information while shopping. Participants with better knowledge and those who reported paying more attention to nutritional information during the shop had lower mean FSA scores. When attention and knowledge were included in the model, the effect of seeing Nutri-Score on the mean FSA score of the basket was reduced and no longer statistically significant (Table 2, Model 4). We confirmed this mediation effect using a structural equation model (for full model see supplementary material, Table S4). This shows that participants who were exposed to Nutri-Scores paid more attention to nutritional information during the shop and had a better knowledge of Nutri-Scores later, and this influenced their purchasing decisions. In numerical terms, about 51% of the effect of Nutri-Score on average FSA scores was mediated by the combined effect of increased attention paid to nutritional information and knowledge of Nutri-Scores (see notes under Table S4). Most of the effect was due to increased attention paid to nutritional information (43%) compared to knowledge of Nutri-Scores (8%).

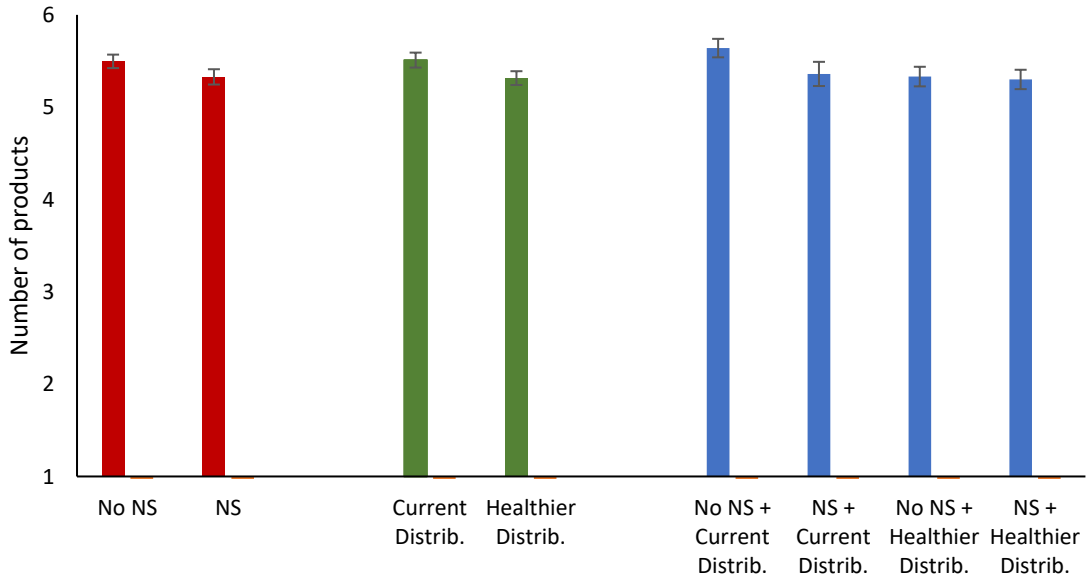
By contrast, participants who were randomly selected to shop in a healthier market did not pay more attention or display better knowledge than participants in the less healthy market. Our two main experimental manipulations therefore acted through different mechanisms.

Adding socio-demographic controls for age, gender, education, employment, social class and having a child in the house in all models produced no effects. We tested for hypothesised interactions and found that the Nutri-Score condition did not have a differential effect on either men or those with children (Interaction effect: B = -0.02, SE = 0.75, p = .98; B = 0.71, SE = 0.83, p = .39). As a final sensitivity check, we re-ran the analyses excluding 60 participants who guessed that nutritional labelling was one of the aims of the experiment. This did not change the results.

3.1.2 *Does Nutri-Score or the distribution of products in the market influence how much people buy?*

Figure 5 shows the total number of products purchased by participants in each condition. There was a very small difference between groups. Participants who were exposed to Nutri-Scores or who shopped in a healthier market purchased one third less of a product on average.

Figure 5. Number of products purchased for Nutri-Score vs No Nutri-Score; Current Distribution vs. Healthier Distribution and all four conditions separately.



Note: No NS = No Nutri-Score shown, NS = Nutri-Score shown. Error bars represent standard error of the mean.

The number of products purchased overall ranged from 1 to 17 but few participants purchased less than 4 or more than 8 products, so we transformed this dependent variables into a 7-level variable that grouped these extreme values.

Contrary to our hypotheses, ordinal logistic regression revealed no significant reduction in the number of items purchased by participants who were exposed to Nutri-Scores (versus not), nor a change in the number of items purchased by participants who shopped in the healthier compared to the current market (Table 3, Model 1). The interaction between seeing Nutri-Scores and the healthier market was non-significant (Table 3, Model 2). However, participants who shopped in the current market without seeing Nutri-Scores bought slightly more products compared to those in the other three conditions (Table 3, Model 3). The size of this effect was less than one product. Recall that there was no relationship between the FSA score of the products available and the price of the products.

There was no effect of socio-demographic variables, knowledge of Nutri-Scores, or attention to nutritional information during the shop (Table 3, Model 4) on the number of products purchased. People who reported being more health conscious tended to purchase fewer products (Table 3, Model 4).

These results were unchanged by excluding participants who guessed that nutritional labelling was one of the aims of the experiment or by using the original (untransformed) total count variable.

Table 3. Ordinal logistic regression analyses with the number of products purchased as the dependent variable and condition as the independent variables.

	Model 1 B(SE)	Model 2 B (SE)	Model 3 B (SE)	Model 4 B (SE)
Nutri-Score (ref. No Nutri-Score)	**-0.19 (0.13)**	**-0.35 (0.18)***		**-0.19 (0.13)**
Healthier distribution (ref. current distribution)	-0.15 (0.13)	-0.32 (0.18)‡		-0.16 (0.13)
Nutri-Score*Healthier Distribution		0.33 (0.25)		
Market + NS (ref. No NS + Current Market)				
NS + Current Market			**-0.36 (0.18)***	
No NS + Healthier Market			**-0.32 (0.18)***	
NS + Healthier Market			**-0.34 (0.18)***	
Health consciousness				-0.13 (0.04)**
Paid attention to nutritional information during shop				0.01 (0.04)
Knowledge of Nutri-Scores				0.004 (0.03)
Socio-demographic controls?	No	No	No	Yes
N	800	800	800	797

Note. No NS = No Nutri-Score shown, NS = Nutri-Score shown. Model 4 and 5 controlled for Age, Gender, Degree, Employment, Child under 18, Social Grade (ABC1 vs C2DE and F). ‡ $p = 0.07$, *$p < .05$, ** $p < .01$, *** $p < .001$, one-tailed in **bold**.

We did not find any evidence that participants in the Nutri-Score condition bought fewer total grams or smaller products on average compared to participants in the control condition (Figure S4).

3.2 Hypothetical Purchasing Decisions

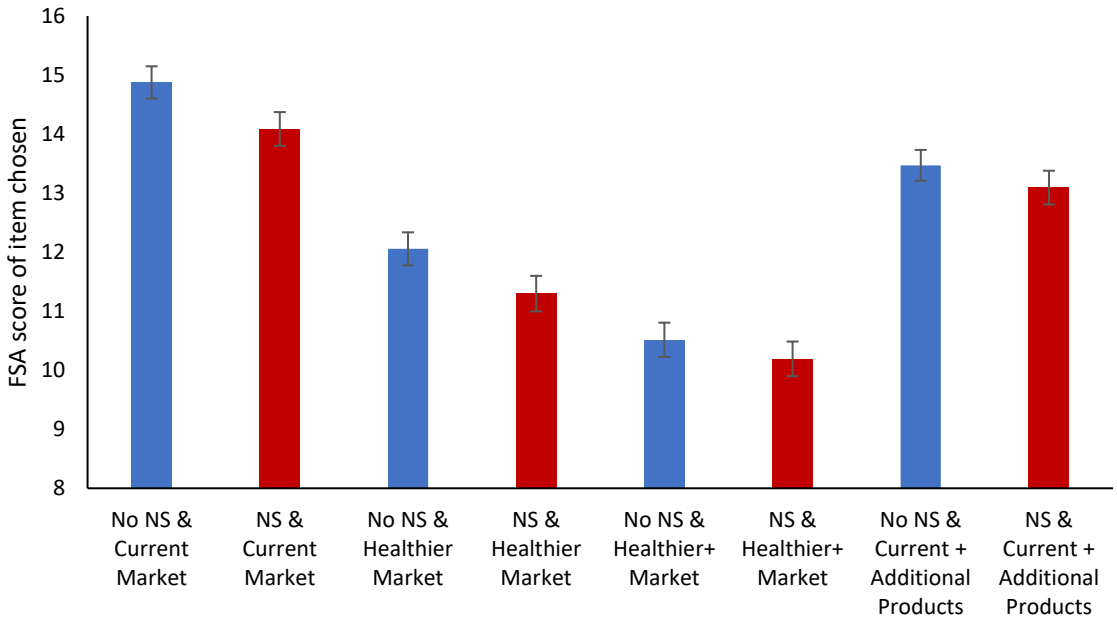
3.2.1 *Does Nutri-Score or the distribution of products in the market influence the nutritional content of products chosen?*

The second shopping task replicated effects observed in the incentivised shop. Participants who were exposed to Nutri-Scores and who shopped in healthier markets chose products with lower (healthier) FSA Scores (see Figure 6).

We ran a multi-level linear regression with random effects for individuals and fixed effects for category. In support of our hypotheses, participants who were exposed to Nutri-Score chose products with lower FSA scores than participants who were not exposed to Nutri-Score, as did participants who saw any of the three healthier markets than the current market (Table S5, Model 1). Similar to the incentivised shopping task, attention paid to nutritional information while shopping and knowledge of Nutri-Scores mediated the relationship between Nutri-Score and healthier FSA scores, but not between the market and healthier FSA scores. (Table S5, Model 2).

We did not find any difference in the size of the products chosen by condition (Table S5, Model 3).

Figure 6. FSA scores of the item chosen for each of the four simulated markets and whether Nutri-Score was shown or not.



Note. No NS = No Nutri-Score shown, NS = Nutri-Score shown. The market manipulations were: Current Market (distribution of Nutri-Scores that exists in the current market), Healthier Market (increased relative availability of healthier products), Healthier+ Market (greater increased relative availability of healthier products), Current + Additional Products (the existing market supplemented by additional healthier products). Error bars represent standard error of the mean. We have scaled the Y-axis to be approximately 1 SD (8.15) in line with the recommendation from Witt (2019).

3.2.1 *Does reducing the availability of unhealthier options or adding new healthier options have a bigger effect on choice?*

We tested for differences between the two types of market manipulations – a relative increase in availability (healthier market) and relative and absolute increase in availability (additional products

market) – using tests of equivalent coefficients. Both manipulations had a significant effect on the FSA score of the product chosen but the effect of shopping in an environment where unhealthier options had been replaced by slightly healthier ones was greater than shopping in an environment where the current market was supplemented with additional healthier products, $\chi^2(1)=67.63$, $p < .001$.

4. Discussion

In this study, consumers purchased healthier (lower FSA score) snack foods when Nutri-Score labels were applied to products and when they shopped in a healthier market. The data are consistent with these effects being additive. This is the first study to our knowledge that has examined these combined effects of labelling and market distribution on choice by simulating potential industry responses to the introduction of a front-of-package labelling policy.

4.1 Policy Implications

We discuss three aspects of our results and associated policy implications: 1) effect sizes seem small but could be meaningful at a population level; 2) consumers are willing to purchase healthier alternatives when the option is made available to them, but reformulating existing products may be more powerful than adding new healthier products; 3) nutritional labelling continues to have an influence on decisions, even when the market changes.

4.1.1 Nutri-Score draws consumer attention towards nutritional information and influences choices.

Previous lab and field trials have shown that consumers tend to purchase foods with lower FSA scores when Nutri-Score labels are applied to products (Crosetto et al., 2019; Dubois et al., 2021; Julia et al., 2016). The sizes of these effects vary between studies: two lab trials recorded a difference of -1 and -2.5 FSA points respectively, while a field trial found a difference of -0.1 FSA points (Crosetto et al., 2019; Dubois et al., 2021; Julia et al., 2016). Our difference was approximately -1 FSA points when comparing the purchases of those who were exposed to Nutri-Score to those who were not, and -2 FSA points when comparing those who were exposed to Nutri-Score while shopping from a healthier market compared to those who did not while shopping in the current market. Our effect-sizes thus broadly concur with previous lab studies. Importantly, we studied purchasing decisions only of snack foods, while most other studies have investigated purchases across a whole range of grocery products. Although nutritionists recommend only irregular consumption of the types of snack foods in our range, most of our participants reported regularly purchasing them. That we find an effect of applying Nutri-Score labels to these categories is important as it may be easier to shift consumers to healthier versions of snack products than to shift them away from consuming them entirely. Nevertheless, it should be noted that Nutri-Score labels appeared to be more effective in those who were already somewhat health conscious when purchasing groceries. Although the absolute differences in FSA scores are small, they were observed for a one-time purchase. If the choices we observed were repeated over time, the cumulative effect could become substantive. Indeed, there is some evidence that small changes in FSA scores can generate meaningful impacts on health. For instance, one study found that a 1-point increase in FSA score across the whole diet was associated with a 16% higher risk of obesity in men (Julia et al., 2015). More research on longer-term impacts is needed.

4.1.2 Reformulation may have a stronger impact on consumer choice than adding choice.

When consumers shop in markets offering a relatively higher proportion of healthier products, they tend to make healthier purchases (Grech and Allman-Farinelli, 2015; Pechey et al., 2019; Pechey et al., 2020; Pechey and Marteau, 2018). We observed a 1 point difference in the average FSA score of purchases when participants shopped in a healthier market than the current one. Crucially, we manipulated the relative availability of healthier options without removing the opportunity to make unhealthier choices. Previous work has theorised that there may be a different response depending on whether a market increases choice by adding healthier products or removes choice by replacing unhealthy products with healthier alternatives (Pechey et al., 2020). However no work to date had compared them directly. The question is important for food policy because the two manipulations simulate a different type of industry response; adding choice to a range is akin to industry responding to food labels by offering new products while replacing unhealthier products with healthier alternatives is akin to industry reformulating certain products. We find that in both cases consumers made healthier choices, but the effect was substantially larger when unhealthier products were replaced with healthier products. When devising labelling policies and negotiating with industry, policymakers may therefore want to incentivise reformulation of existing products over developing new ones.

4.1.3 Combining reformulation and food labelling policies will likely lead to bigger effects on food choice.

This is the first study to our knowledge to test whether the effect of Nutri-Score differs if the provider response is to change product ranges. We find an additive effect. Our data suggests that Nutri-Score attracts more attention to nutritional information, increasing the weight it receives in the decision process, while changed availability operates via other mechanisms, perhaps influencing the relative internal psychological scaling of product attributes (including healthiness) (see e.g. Fasolo et al., 2009; Mellers and Cooke, 1994; Pechey et al., 2020). As attention mediates the effect, a policy seeking to improve nutrition through labelling should ensure that both changes are salient enough to capture attention. Displaying information about Nutri-Score in additional places as well as on the front of packaging (e.g., on information leaflets, store displays, etc.) could increase attention drawn to the label and increase the effect.

4.2 Strengths and Limitations

The design of this study strove to prioritise external validity within the confines of a controlled experiment. It involved incentivised decision-making, by selecting a proportion of participants to receive the products they chose. This should have limited potential effects of experimenter demand. The online shopping environment was programmed to closely resemble those in the market. By using a database of nutritional information for all products in each of eight categories from one of the country's largest supermarket chains, we ensured that the control condition was representative of products currently available to consumers. Moreover, we created a realistic healthier market by matching the distributions of Nutri-Scores in one of the largest supermarket chains in a country with Nutri-Score. The products that we chose all existed in the current market, were available in the largest supermarket chains in the country and were therefore familiar to participants.

The study also has limitations. Participants were aware that they were part of a study, which may have influenced choices. A minority (7.5%) indicated correctly that a study aim was to examine the effect of nutritional labelling on decisions, although our results are unaffected by excluding these participants. Additionally, while the online study context may be representative of the context when people shop online (e.g., similar likelihood of interruption and desire to complete the task and do something else, etc.), it may result in different effect sizes relative to field settings, such as

supermarkets, where other influential contextual factors are present (e.g. a larger number of visible products, the presence of other shoppers, etc.). While we recorded relative differences between conditions, effect sizes may not be predictive of population level impacts. Nevertheless, small changes to diet can make big differences across a population and over a lifespan. The study also considered only dry pre-packaged snack foods. This was in part because we were interested in the effects of Nutri-Score on snack foods, but also because there are practical issues surrounding delivering products to participants. Our effects do not necessarily translate to other food products and the effectiveness of nutritional labels can vary depending on food category (Ikonen et al., 2020). Snack foods have high Nutri-Scores, which may make it easier for shoppers to switch to lower Nutri-Score alternatives. Hence the overall effect of Nutri-Score might have been less had we used an online shop that included all product categories. It is similarly unclear how a once-off exposure to Nutri-Score translates to repeated exposures. People could form a habit of paying attention to the label, strengthening the impact, or they could stop paying attention after a period of novelty, diminishing the effect. It would also be important to check whether people repurchase the lower Nutri-Score alternatives having tasted them or if they return to the original products. Studies that check repeated purchasing patterns through longitudinal data would be useful in future to assess this.

The previous paragraph considers limitations that might affect generalisability, but our outcome variable itself also merits consideration. Throughout this paper we have described products with lower FSA scores as being 'healthier', but we are aware that assessing the healthfulness of diet and of different products is complicated by factors such as portion sizes and frequency of consumption. Our aim was not to test whether Nutri-Score is nutritionally the most appropriate label, but to illustrate its potential impact on behaviour. Future work could analyse big data to assess whether suppliers alter specific nutritional components of their products in response to Nutri-Score labelling policies.

4.3 Conclusion

In conclusion, Nutri-Score labelling is likely to influence consumers to choose snack products with lower FSA scores by increasing the attention they pay to nutritional information. If the label increases the availability of snack products with lower FSA scores, our data imply an additional benefit in terms of lower FSA scores. Thus, if a policy aim is to encourage consumers to choose snack products with lower FSA scores, then our findings support parallel targeting of both labelling and changes to current market offerings.

References

Anwyl-Irvine, A.L., Massonnié, J., Flitton, A., Kirkham, N., Evershed, J.K., 2020. Gorilla in our midst: An online behavioral experiment builder. Behav. Res. Methods 52, 388-407.
Crosetto, P., Lacroix, A., Muller, L., Ruffieux, B., 2019. Nutritional and economic impact of five alternative front-of-pack nutritional labels: experimental evidence. Europ. Rev. Agr. Econ. 47, 785-818.
Dai, J., Cone, J., Moher, J., 2020. Perceptual salience influences food choices independently of health and taste preferences. Cognitive Research: Principles and Implications 5, 2.
De Marchi, E., Cavaliere, A., Pucillo, F., Banterle, A., Nayga, R.M., Jr, 2023. Dynamics of demand-side and supply-side responses to front-of-pack nutrition labels: a narrative review. Europ. Rev. Agr. Econ.
Dréano-Trécant, L., Egnell, M., Hercberg, S., Galan, P., Soudon, J., Fialon, M., Touvier, M., Kesse-Guyot, E., Julia, C., 2020. Performance of the Front-of-Pack Nutrition Label Nutri-Score to Discriminate the Nutritional Quality of Foods Products: A Comparative Study across 8 European Countries. Nutrients 12, 1303.
Dubois, P., Albuquerque, P., Allais, O., Bonnet, C., Bertail, P., Combris, P., Lahlou, S., Rigal, N., Ruffieux, B., Chandon, P., 2021. Effects of front-of-pack labels on the nutritional quality of supermarket food purchases: evidence from a large-scale randomized controlled trial. Journal of the Academy of Marketing Science 49, 119-138.
Ducrot, P., Julia, C., Méjean, C., Kesse-Guyot, E., Touvier, M., Fezeu, L.K., Hercberg, S., Péneau, S., 2016. Impact of Different Front-of-Pack Nutrition Labels on Consumer Purchasing Intentions: A Randomized Controlled Trial. Am. J. Prev. Med. 50, 627-636.
Egnell, M., Talati, Z., Hercberg, S., Pettigrew, S., Julia, C., 2018. Objective Understanding of Front-of-Package Nutrition Labels: An International Comparative Experimental Study across 12 Countries. Nutrients 10, 1542.
Fasolo, B., Hertwig, R., Huber, M., Ludwig, M., 2009. Size, entropy, and density: What is the difference that makes the difference between small and large real-world assortments? Psychology & Marketing 26, 254-279.
Fialon, M., Nabec, L., Julia, C., 2022. Legitimacy of front-of-pack nutrition labels: controversy over the deployment of the Nutri-Score in Italy. International journal of health policy and management 11, 2574.
Follis, A., Sept pays de l'UE s'opposent au système d'étiquetage des aliments Nutri-Score. September 2020. https://www.euractiv.fr/section/alimentation/news/sept-pays-de-lue-sopposent-ausysteme-detiquetage-des-aliments-nutri-score/. [Date Accessed: August 14, 2023]
FoodNavigator, Does Nutri-Score discriminate against traditional foods? December 11, 2020. https://www.foodnavigator.com/Article/2020/12/11/Does-Nutri-Score-discriminate-against-traditional-foods. [Date Accessed: August 14, 2023]
Grech, A., Allman-Farinelli, M., 2015. A systematic literature review of nutrition interventions in vending machines that encourage consumers to make healthier choices. Obes. Rev. 16, 1030-1041.
Hafner, E., Pravst, I., 2021. Evaluation of the Ability of Nutri-Score to Discriminate the Nutritional Quality of Prepacked Foods Using a Sale-Weighting Approach. Foods 10.
Ikonen, I., Sotgiu, F., Aydinli, A., Verlegh, P.W.J., 2020. Consumer effects of front-of-package nutrition labeling: an interdisciplinary meta-analysis. Journal of the Academy of Marketing Science 48, 360-383.
Julia, C., Blanchet, O., Méjean, C., Péneau, S., Ducrot, P., Allès, B., Fezeu, L.K., Touvier, M., Kesse-Guyot, E., Singler, E., Hercberg, S., 2016. Impact of the front-of-pack 5-colour nutrition label (5-CNL) on the nutritional quality of purchases: an experimental study. Int. J. Behav. Nutr. 13, 101.
Julia, C., Ducrot, P., Lassale, C., Fézeu, L., Méjean, C., Péneau, S., Touvier, M., Hercberg, S., Kesse-Guyot, E., 2015. Prospective associations between a dietary index based on the British Food Standard Agency nutrient profiling system and 13-year weight gain in the SU.VI.MAX cohort. Preventive Medicine 81, 189-194.

Julia, C., Hercberg, S., 2017. Nutri-Score: Evidence of the effectiveness of the French front-of-pack nutrition label. Ernahrungs Umschau 64, 181-187.
Lurie, N.H., 2004. Decision Making in Information-Rich Environments: The Role of Information Structure. Journal of Consumer Research 30, 473-486.
Mellers, B.A., Cooke, A.D., 1994. Trade-offs depend on attribute range. Journal of Experimental Psychology: Human Perception and Performance 20, 1055.
Pechey, R., Cartwright, E., Pilling, M., Hollands, G.J., Vasiljevic, M., Jebb, S.A., Marteau, T.M., 2019. Impact of increasing the proportion of healthier foods available on energy purchased in worksite cafeterias: A stepped wedge randomized controlled pilot trial. Appetite 133, 286-296.
Pechey, R., Hollands, G.J., Carter, P., Marteau, T.M., 2020. Altering the availability of products within physical micro-environments: a conceptual framework. BMC Public Health 20, 986.
Pechey, R., Marteau, T.M., 2018. Availability of healthier vs. less healthy food and food choice: an online experiment. BMC Public Health 18, 1296.
Roberto, C.A., Ng, S.W., Ganderats-Fuentes, M., Hammond, D., Barquera, S., Jauregui, A., Taillie, L.S., 2021. The Influence of Front-of-Package Nutrition Labeling on Consumer Behavior and Product Reformulation. Annu. Rev. Nutr. 41, 529-550.
Shangguan, S., Afshin, A., Shulkin, M., Ma, W., Marsden, D., Smith, J., Saheb-Kashaf, M., Shi, P., Micha, R., Imamura, F., Mozaffarian, D., 2019. A Meta-Analysis of Food Labeling Effects on Consumer Diet Behaviors and Industry Practices. Am. J. Prev. Med. 56, 300-314.
Szabo de Edelenyi, F., Egnell, M., Galan, P., Druesne-Pecollo, N., Hercberg, S., Julia, C., 2019. Ability of the Nutri-Score front-of-pack nutrition label to discriminate the nutritional quality of foods in the German food market and consistency with nutritional recommendations. Archives of Public Health 77, 28.
Vandevijvere, S., Vanderlee, L., 2019. Effect of Formulation, Labelling, and Taxation Policies on the Nutritional Quality of the Food Supply. Curr Nutr Rep 8, 240-249.
Witt, J.K., 2019. Graph construction: an empirical investigation on setting the range of the y-axis. Meta-psychology 3.