London's Global University

University College London
Institute of Child Health

Front of Pack Food Labels and UK Policy

A thesis submitted for the degree of Doctor of Philosophy

Jessica Packer

2023

Declaration

I, Jessica Packer, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Jessica Packer

23rd of July 2023

Abstract

Front of pack labels (FOPLs) aim to enable consumers to make healthier choices and encourage product reformulation. At the time of writing, the United Kingdom (UK) Government was reviewing the current FOPL policy to ensure that it is based on high quality, relevant evidence.

I conducted an online randomised controlled trial to evaluate the effectiveness of FOPL schemes (Multiple Traffic Light, MTL; Nutri-Score, N-S; Warning Label; Positive Choice tick) on objective understanding of product healthiness in a British population and to assess subjective factors likely to impact on engagement, informing UK policy development. Additionally, I conducted latent class analysis to understand how participant characteristics may impact on FOPL engagement and held a patient and public involvement/engagement (PPIE) session to understand the views of children and young people (CYP) on FOPLs and policy options.

I found that all FOPL schemes improved participants' ability to correctly rank products according to their healthiness, with N-S, followed by MTL showing the largest effects. Time to correctly rank the products was fastest using N-S; MTL and N-S were most favourably perceived; and descriptive analyses indicated N-S was the most stable FOPL across education and income groups. Latent class analysis indicated that regardless of FOPL engagement and motivation factors, FOPLs improved participants' ability to correctly understand product healthiness, but perceptions of labels and confidence using labels varied. The PPIE session showed CYP supported mandatory implementation of FOPL, finding N-S and MTL as the easiest to use and understand.

I found FOPLs were effective at improving the ability of consumers to understand product healthiness, with evidence suggesting a summary indicator FOPL may work best across all population groups. The evidence supports the mandatory implementation of FOPLs in the UK, which will be integral to their effectiveness, improving diet quality through consumer understanding and product reformulation.

Impact Statement

My findings regarding the effectiveness of front of pack labels (FOPLs) in enabling consumers to understand the healthiness of products addressed several research areas. These included a research gap with evidence lacking in representative British samples and areas where there was no evidence at all (including confidence in using FOPLs, speed of ranking). This evidence will be central to the development and implementation of UK policy in this area. This work was commissioned by the Department of Health and Social Care (DHSC) and is a significant and key contribution to the current policy development of FOPL options in the UK. The findings will be the key evidence that informs development and consultation with stakeholders, including industry, regarding the details of which scheme and how it should be implemented. Any updated policy could improve the food environment and the nutritional quality of diets of the UK population.

This research has had an impact at the national level, as I have led policy briefings to DHSC, Public Health England, Food Standards Agency in Northern Ireland, Food Standards Scotland, and the Welsh Government in the form of two briefing papers and multiple presentations, written for non-academic audiences. At the invitation of national policy makers and key stakeholders, I have also been involved with international FOPL policy development through the Global Better Health Programme funded by UK Government. For this work, I peer-reviewed a protocol for a FOPL experiment in Thailand written by Australian researchers. I was also invited to present at an international knowledge exchange session facilitated by UK Government to inform and provide learning opportunities for the Thai team, including researchers from Australia and South Africa.

My findings that FOPLs can successfully convey the nutritional quality of products and assist consumers identify healthier options will benefit public health researchers specialising in this field by substantially developing the evidence base and providing a foundation for further work. For policymakers, this work provided high quality, timely, and cost effective evidence. This adds to a growing international evidence base for the effectiveness and adoption of interpretive FOPLs.

My latent class analysis and descriptive analysis of the impact of education and income will be of interest to policymakers as it identifies the components of labels that work effectively across all socioeconomic status groups. This will be useful in the development and implementation of a policy that is effective across the population and does not widen health inequalities.

I have disseminated these findings with the wider research community in several ways: as an oral presentation at the 2022 European Congress on Obesity in Maastricht, as a poster presentation at the 2023 European Congress on Obesity in Dublin, and as an invited speaker at the Nutritional Epidemiology group at Leeds University. I have also published my work in three peer-reviewed open access scientific journals (Nutrients, Journal of Human Nutrition, and Dietetics). This work has strengthened the evidence base and will be valuable in international policy discussions.

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

   a) **What is the title of the manuscript?**

   Assessing the Effectiveness of Front of Pack Labels: Findings from an Online Randomised-Controlled Experiment in a Representative British Sample

   b) **Please include a link to or doi for the work**

   https://doi.org/10.3390%2Fnu13030900

   c) **Where was the work published?**

   Nutrients

   d) **Who published the work?** (e.g. OUP)

   MDPI

   e) **When was the work published?**

   March 10th 2021

   f) **List the manuscript's authors in the order they appear on the publication**

   Jessica Packer, Simon J. Russell, Deborah Ridout, Steven Hope, Anne Conolly, Curtis Jessop, Oliver J. Robinson, Sandro T. Stoffel, Russell M. Viner, and Helen Croker

   g) **Was the work peer reviewed?**

   Yes

   h) **Have you retained the copyright?**

   Yes

   i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

   No

For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

The study was conceived and designed by J.P., R.M.V., H.C., S.J.R., D.R. The experiment was designed by J.P., H.C., A.C., C.J.; O.J.R., S.T.S. advised on the experimental design of the study. The data collection was led by C.J. and A.C. The analysis was designed and undertaken by D.R., J.P.; H.C., J.P., R.M.V., D.R., S.J.R., S.H. contributed to the interpretation of the data. The first draft was written by H.C. and J.P. All authors critically reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

2. In which chapter(s) of your thesis can this material be found?

Chapter five

3. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Jessica Packer

Date:

29th of June, 2023

Supervisor/ Senior Author (where appropriate)

Russell Viner

Date

5th of July, 2023

UCL Research Paper Declaration Form

referencing the doctoral candidate's own published work(s)

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

   j) **What is the title of the manuscript?**

   Secondary Outcomes of a Front-of-Pack-Labelling Randomised Controlled Experiment in a Representative British Sample: Understanding, Ranking Speed and Perceptions

   k) **Please include a link to or doi for the work**

   https://doi.org/10.3390%2Fnu14112188

   l) **Where was the work published?**

   Nutrients

   m) **Who published the work?** (e.g. OUP)

   MDPI

   n) **When was the work published?**

   May 24th 2022

   o) **List the manuscript's authors in the order they appear on the publication**

   Jessica Packer, Simon J. Russell, Deborah Ridout, Steven Hope, Anne Conolly, Curtis Jessop, Oliver J. Robinson, Sandro T. Stoffel, Russell M. Viner, and Helen Croker

   p) **Was the work peer reviewed?**

   Yes

   q) **Have you retained the copyright?**

   Yes

   r) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

   No

2. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

The study was conceived of and designed by J.P., R.M.V., H.C., S.J.R. and D.R. The experiment was designed by J.P., H.C., A.C. and C.J. The data collection was led by C.J. and A.C. The analysis was designed and undertaken by D.R.; H.C., J.P., R.M.V., D.R., S.J.R. contributed to the interpretation of the data. The first draft was written by H.C. and J.P. All authors critically reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

3. **In which chapter(s) of your thesis can this material be found?**

Chapter six

4. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Jessica Packer

Date:

29th of June, 2023

Supervisor/ Senior Author (where appropriate)

Russell Viner

Date

5th of July, 2023

UCL Research Paper Declaration Form

referencing the doctoral candidate's own published work(s)

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

   s) **What is the title of the manuscript?**

   Front of pack nutritional labelling schemes: a systematic review and meta-analysis of recent evidence relating to objectively measured consumption and purchasing

   t) **Please include a link to or doi for the work**

   https://doi.org/10.1111/jhn.12758

   u) **Where was the work published?**

   Journal of Human Nutrition and Dietetics

   v) **Who published the work?** (e.g. OUP)

   John Wiley & Sons Ltd

   w) **When was the work published?**

   May 4th 2020

   x) **List the manuscript's authors in the order they appear on the publication**

   Jessica Packer, Simon J. Russell, Deborah Ridout, Steven Hope, Anne Conolly, Curtis Jessop, Oliver J. Robinson, Sandro T. Stoffel, Russell M. Viner, and Helen Croker

   y) **Was the work peer reviewed?**

   Yes

   z) **Have you retained the copyright?**

   Yes

   aa) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

   No

2. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

All authors participated in the protocol development. CS and JP conducted the literature searches. JP and HC screened and extracted the data from the literature. JP conducted the meta-analysis and created all figures and tables. HC and JP wrote the first draft and all authors reviewed and commented on subsequent drafts and approved the final version submitted for publication.

3. **In which chapter(s) of your thesis can this material be found?**

Chapter three

4. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Jessica Packer

Date:

29th of June, 2023

Supervisor/ Senior Author (where appropriate)

Russell Viner

Date

5th of July, 2023

Table of Contents

List of Tables

List of Figures

List of Frequently Used Acronyms And Abbreviations

BMI	Body Mass Index
CI	Confidence Interval
CVD	Cardiovascular Disease
CYP	Children and Young People
DALYs	Disability-Adjusted Life Years
DHSC	Department of Health and Social Care
DIG	Daily Intake Guide
ES	Effect Size
EU	European Union
FOPL	Front of Pack Label
FSA	Food Standards Agency
FuF	Facts Up Front
g	Gram
GDA	Guideline Daily Amount
HFSS	High in Fat, Salt and Sugar
ITS	Interrupted Time Series
kcal	Kilocalories
kJ	kilojoules
LCA	Latent Class Analysis
MTL	Multiple Traffic Lights
N-S	Nutri-Score
NCB	National Children's Bureau
NCD	Non-Communicable Diseases
NHS	National Health Service
OPRU	Obesity Policy Research Unit
PC	Positive Choice Tick
PHE	Public Health England
PPIE	Patient and Public Involvement and Engagement
RI	Reference Intakes
RM	Relative Mean
RR	Relative Risk
RRR	Relative Risk Ratio
SD	Standard Deviation
SDIL	Soft Drinks Industry Levy
SES	Socioeconomic Status
TL	Traffic Light
UK	United Kingdom
UPF	Ultra-Processed Food
WHO	World Health Organization
WL	Warning Label
WMD	Weighted Mean Difference

List of Publications Related to this Thesis

Peer-reviewed Journal Articles

Packer J, Russell SJ, Ridout D, Hope S, Conolly A, Jessop C, Robinson OJ, Stoffel ST, Viner RM, Croker H (2021) Assessing the Effectiveness of Front of Pack Labels: Findings from an Online Randomised-Controlled Experiment in a Representative British Sample. Nutrients. 13(3):900. https://doi.org/10.3390%2Fnu13030900

Packer J, Russell SJ, Ridout D, Conolly A, Jessop C, Viner RM, Croker H. (2022) Secondary Outcomes of a Front-of-Pack-Labelling Randomised Controlled Experiment in a Representative British Sample: Understanding, Ranking Speed and Perceptions. Nutrients, 14, 2188. https://doi.org/10.3390/nu14112188

Croker H, **Packer J**, Russell S, Stansfield C, Viner R. (2019) Front of pack nutritional labelling schemes: a review and meta-analysis of recent evidence relating to objectively measured consumption and purchasing. London: Great Ormond Street Institute of Child Health, University College London. p. 1-29. https://doi.org/10.1111/jhn.12758

Policy Briefing Papers

Packer J, Russell S, Ridout D, Hope S, Connolly A, Jessop C, Viner R, Croker H. (2021) Preliminary results from an experimental study examining the impact of front of pack labelling skills on knowledge. London: Great Ormond Street Institute of Child Health, University College London. p. 1-24 (Internal policy paper)

Packer J, Russell S, Ridout D, Hope S, Connolly A, Jessop C, Viner R, Croker H. (2021) Front of Pack Labelling Experiment – Final Report. London: Great Ormond Street Institute of Child Health, University College London. p. 1-24 (Internal policy paper)

Conference Presentations

Packer J, Russell S, Ridout D, Hope S, Connolly A, Jessop C, Viner R, Croker H. (2022) Oral presentation: Front of pack labels and objective understanding of product healthiness: Findings from an Online Randomised-Controlled Experiment in a Representative British Sample. Obes Facts 15 (Suppl. 1): 1–240. https://doi.org/10.1159/000524469

Packer J, Russell S, Viner R. (2023) PO2.084 Front of pack labels, who engages with them and why: a latent class analysis from an online experiment testing label effectiveness. Obes Facts 16 (Suppl. 1): 1–351. https://doi.org/10.1159/000530456

List of Other Publications Completed During PhD

Packer J, Croker H, Goddings AL, Boyland EJ, Stansfield C, Russell SJ, Viner RM (2022) Advertising and Young People's Critical Reasoning Abilities: Systematic Review and Meta-analysis. Pediatrics e2022057780. https://doi.org10.1542/peds.2022-057780

Packer J, Russell SJ, Siovolgyi, G.; McLaren, K.; Stansfield, C.; Viner, R.M.; Croker, H. The Impact on Dietary Outcomes of Celebrities and Influencers in Marketing Unhealthy Foods to Children: A Systematic Review and Meta-Analysis. Nutrients 2022, 14, 434. https://doi.org/10.3390/nu14030434

Packer J, Russell SJ, McLaren K, et al. The impact on dietary outcomes of licensed and brand equity characters in marketing unhealthy foods to children: A systematic review and meta-analysis. Obesity Reviews. 2022; 23(7):e13443. https://doi:10.1111/obr.13443

Viner R, Russell S, Saulle R,.. **Packer, J**. et al. (2022) School Closures During Social Lockdown and Mental Health, Health Behaviors, and Well-being Among Children and Adolescents During the First COVID-19 Wave: A Systematic Review. JAMA Pediatr 176(4):400–409. https://doi.org/10.1001/jamapediatrics.2021.5840

Russell SJ, Hope S, Croker H, **Packer J**, Viner RM (2022) Is it possible to model the impact of calorie-reduction interventions on childhood obesity at a population level and across the range of deprivation: Evidence from the Avon Longitudinal Study of Parents and Children (ALSPAC). PLOS ONE 17(1): e0263043. https://doi.org/10.1371/journal.pone.0263043

Russell, SJ, Hope, S, Croker, H,…**Packer, J** et al. (2021) Modeling the impact of calorie-reduction interventions on population prevalence and inequalities in childhood obesity in the Southampton Women's Survey. Obes Sci Pract 7(5): 545- 554. https://doi.org/10.1002/osp4.520

Viner R, Russell S, Croker H, **Packer J**, et al. (2020). School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. The Lancet Child & Adolescent Health. 4(5):397 - 404 https://doi.org/10.1016/S2352-4642(20)30095-X

Acknowledgements

Firstly, I would like to thank and acknowledge my supervisors, Russell Viner, Helen Croker and Simon Russell. Thank you for your support and belief that I could complete this PhD while working full-time. I appreciate the guidance, feedback and the time you all invested in this PhD. I have learnt a lot from you all. Thank you to Deborah Ridout for your statistical expertise and all of the time and support you offered me, it was invaluable. Thank you also to UCL and the Great Ormond Institute of Child Health who supported me as a UCL staff member to complete this study.

Secondly, this work would not have been possible without NatCen and the NatCen panel. Special thank you to Curtis Jessop and Anne Conolly. Thank you to the Young Research Advisors for their input and the facilitators from the National Children's Bureau.

A huge thank you to my friends. To the other PhD students in the Population, Policy and Practice department, thank you for the inspiration, camaraderie and the laughs! To my friends in London, Australia, all over, thank you for being understanding when I was slow to respond to messages!

Lastly and most importantly, my family. Thank you for all of the support and encouragement, for reading chapters or listening to presentations, it means a lot. To Sara and Mike, thank you for both always being there and always selflessly supporting me, no matter what. Thank you for sitting next to me while I was working, for the cuppas, for editing sections and for the hugs. I could not have done it without you both and I cannot wait to celebrate soon!

Chapter 1: Introduction

The aims of this thesis were to address a research gap relating to front of pack labels (FOPLs) and their effectiveness in the United Kingdom (UK). This work was requested by the Department of Health and Social Care (DHSC) and was undertaken as part of the Obesity Policy Research Unit (OPRU) work programme to inform policy development. This thesis is divided into nine Chapters: Chapter 1 will introduce the research area and the UK context, Chapter 2 will outline the evidence for FOPLs on understanding in a literature review, Chapter 3 is a systematic review examining the impact of FOPLs on consumer behaviour, Chapter 4-8 relate to a series of studies reporting the findings from an online experimental study, and Chapter 9 discusses the findings, the original contribution of the work and the academic and policy impact.

Nutrition and Public Health

Non-communicable diseases (NCDs) are the leading cause of death worldwide (74%), including cardiovascular disease (CVD), cancer and Type 2 diabetes.[1, 2] In the UK, over a quarter of all deaths are caused by CVD (160,000 deaths per year) and in 2019, 89% of all deaths were attributable to NCDs.[3, 4] Unhealthy diets are a key behavioural risk factor for NCDs, contributing to excess weight gain and an increased risk of obesity.[2] The latest figures from the Health Survey for England 2021 show that 26% of adults are living with obesity (defined as having a body mass index (BMI) of 30 kg/m^2 or above) and a further 38% are living with overweight (defined as having a BMI between 25-30 kg/m^2).[5, 6] Diet-related ill health in the UK is estimated to have cost the National Health Service (NHS) £6.1 billion in 2014-15; this is projected to reach £9.7 billion by 2050.[7]

In England, health inequalities, which are defined as inequitable and avoidable differences in health across the population, exist and are increasing.[8-10] Health inequalities are a result of social inequalities; with broad social determinants of health including income, occupation and education, all factors which are often used as measures of socioeconomic status (SES).[9] Social gradients exist across many of these determinants, with people in England living in the most deprived areas having lower life expectancies, than those in the least deprived areas (9.7 years fewer for males and 7.9 years fewer for females).[8, 11] Inequalities in obesity are prevalent, with rates of obesity in childhood and adulthood twice as high in the most deprived areas compared to the least deprived (14% vs 6% for reception aged children; 34% vs 20% for adults) and by ethnicity, with the highest prevalence for black children compared to white children (16% vs 10% for reception aged children).[6, 12] Adult hospital admissions directly attributable to obesity are also over three times more likely in the most deprived compared to the least deprived areas.[13] Differences in obesity rates by inequalities are thought to be driven by differential access to the resources required for healthy diets and adequate physical activity, including financial, social, physical and cognitive resources.[14]

UK dietary recommendations were published by the former Public Health England (PHE), following recommendations from the Committee on Medical Aspects of Food Policy (COMA) and the Scientific Advisory Committee on Nutrition (SACN).[15] Table 1 shows the summary of the current guidelines for adults (aged 19 years and over). Across the UK, the average diet does not meet these nutritional recommendations; exceeding limits for energy intake, sugar, saturated fat and salt, and not meeting consumption guidelines for fibre, or fruit and vegetables.[16-18] A modelling study projected that moving UK diets in line with these recommendations could reduce deaths by approximately 33,000 per year.[19]

Table 1: Summary of the daily UK dietary recommendations for adults aged 19 years+[15]

	Males (intake per day)	**Females (intake per day)**
Energy*	2500kcal	2000kcal
Protein	55.5g	45g
Fat*	Less than 97g, no more than 35% total food energy	Less than 78g, no more than 35% total food energy
Saturated fat*	31g, no more than 11% total food energy	24g, no more than 11% total food energy
Polyunsaturated fat*	18g, no more than 6.5% total food energy	14g, no more than 6.5% total food energy
Monounsaturated fat*	36g, no more than 13% total food energy	29g, no more than 13% total food energy
Carbohydrate*	At least 333g, at least 50% total food energy	At least 276g, at least 50% total food energy
Free sugars*	33g, no more than 5% total food energy	27g, no more than 5% total food energy
Salt	Less than 6g	Less than 6g
Dietary fibre	30g	30g

*For adults aged 65 years and older, there are slight differences to the nutritional requirements, with a decreasing requirement for energy, all fats, carbohydrate and free sugar. The protein requirement for males decreases with age (to 53.5g/day) but requirement for females increases with age (to 46.5g/day). G = gram, kcal = kilocalories.

To improve health in the UK, the government has announced policies focused on reducing the intake of food and drink products that are high in fat, salt and sugar (HFSS).[20] Current food environments are obesogenic, i.e., obesity is promoted through surroundings, opportunities and conditions, which encourage poor diets and overconsumption. Reshaping food environments to support healthy decisions has been identified as a key policy focus.[21, 22] Some of the specific measures introduced by DHSC focus on this rebalancing, with an aspiration to improve the food environment (described as “anti-obesity measures”), including fiscal policies such as the Soft Drinks Industry Levy (SDIL), and advertising restrictions for HFSS products.[20, 23] Policy interventions aimed at the whole population can be effective prevention measures (such as media campaigns to target health behaviours or changing the distribution of a nutrient to avoid), with evidence that small changes can lead to large impacts.[24] Excess energy intake of just 30 kilojoules (kJ) or 7 kilocalories (kcals) per day has been calculated as the underlying energy imbalance leading to the observed average weight gain, indicating that nutrition policies eliciting small changes could be effective at preventing excess weight gain.[25]

The mechanisms of policy actions can be described as upstream, which focus on changing social factors, or downstream, which focus on individual action/behaviour change.[26] In relation to health inequalities, public health experts and evidence suggests that for nutrition policies and interventions aimed at promoting healthy eating, upstream interventions are likely to be more equitable and effective, as they do not rely on individual agency (e.g., changing the price rather than the person, i.e., through dietary counselling).[27-30] But interventions that rely on high individual agency tend to be favoured by governments globally, with a repeated focus on individual responsibility.[27] A recent report quantified the impact and cost/benefit of four upstream obesity prevention policies targeting environmental factors that have been implemented or are scheduled in the UK, including restrictions for HFSS products based on location or volume, pre-watershed advertising ban for TV and online and the SDIL.[31] They were found to have a predicted substantial benefit to UK in terms or health and economic outcomes, around £76 billion (mid-range best estimate), outweighing the estimated cost to industry (£6.3 billion), but the benefits and costs did vary. Another measure, and the focus of this thesis, is FOPLs.

Front of Pack Nutritional Labels

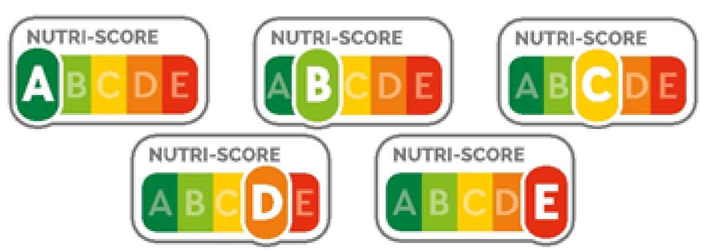
Aims of FOPLs

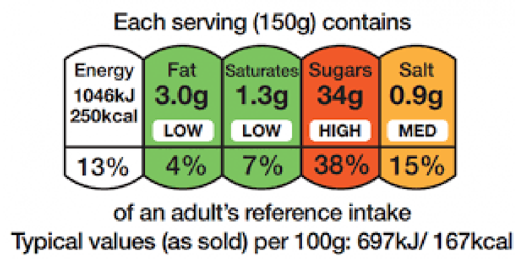
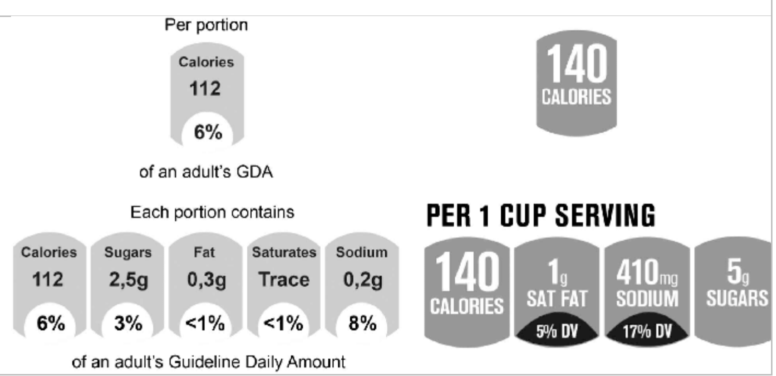
FOPLs are simplified nutritional labels displayed on the front of food and drink product packaging to communicate the nutritional composition/healthiness of products.[32] They are intended to be easy to understand and useful for comparing products at-a-glance; with the aim to aid healthier choices and to drive product reformulation.[32, 33] FOPLs are separate and in addition to the mandatory nutritional information on the back of packaging or shelf labelling. The information provided on the back of the package nutrition declaration includes energy, fat, saturates, carbohydrate, sugars, protein and salt and is stipulated by European Union Food Information for Consumers (EUFIC) Article 30(1) and (3) (which the UK retained after Brexit).[34]

Different FOPL Schemes

FOPLs vary in their design, both functionally and visually, in the content of information (nutrients of interest, per serving size or package), how they provide the information (text, numbers, percentages), and the design (size, colour, shape, placement).[35] FOPLs can be interpretive (providing a judgement on product healthiness or nutrient content using nutrient profiling algorithms, or cut-off points) or non-interpretive (where only information and no judgement or specific indication of how these relate to a healthy diet is provided).[33, 36] Interpretive FOPLs vary and have been broadly categorised as: A) summary indicators (an overall assessment/summary based on the combined nutrient information, which can be positive or negative e.g., French Nutri-Score, N-S or Australian/New Zealand Health Star Rating, HSR; B) nutrient-specific interpretive (assessment provided per nutrient of interest and presented separately, which can be positive or negative e.g., UK Multiple Traffic Light, MTL); C) nutrient-specific warning (a warning is provided for each nutrient that has exceeded a specified threshold, is negative, e.g., Chilean Warning labels); and D) endorsement logos (overall product endorsement/assessment based on combined nutrient information, is positive, indicating a product has met a pre-specified nutritional standard and is 'healthy', e.g. Healthy Choice Logo). Lastly, E) non-interpretive FOPLs include the Daily Intake Guide (DIG), Facts Up Front (FuF), Guideline Daily Amount (GDA), Reference Intake (RI), which are all comparable and used in the EU, US, UK, Australia, respectively, see Figure 1 for real-world FOPL examples.[33]

Figure 1: Examples of front of pack labels and label type

Label type	**Label details**	**Example**
Summary indicator	**A)** Nutri-Score, developed in France. Presents a coloured scale of A (green, higher quality) to E (red, lower quality); Health Star Rating	

	developed in Australia and New Zealand. Presents a rating score from 0.5 (least healthy) - 5 stars (most healthy)	
Nutrient-specific interpretative	**B)** Multiple Traffic Lights, developed in the United Kingdom. Indicates red (high), amber (medium), or green (low) levels of fats, sugars, and salt	
Nutrient-specific warning	**C)** Warning Labels, developed in Chile. Marks products as high in saturated fats, salt, sugar or calories.	
Endorsement logo	**D)** Healthy Choice Logo, developed by the food industry	
Non-interpretive	**E)** Guideline Daily Amount, used in Europe and Fact's Up Front, used in the US	

Theoretical Frameworks

For this thesis, two main theoretical frameworks and logic models are used to underpin the understanding of how FOPLs work and were used for the development of the study materials. A logic model by Crockett et al., proposes that FOPLs can reduce mortality and morbidity from diet-related diseases by improving diet quality via two main action pathways/processes: improving consumer understanding of product healthiness (consumers must engage with labels) and driving product reformulation by industry, see Figure 2.[32] A theoretical framework developed by Grunert and Wills focuses on the first pathway, improving consumer understanding, and describes the factors involved in the consumer response to FOPLs, see Figure 3.[37] This framework outlines that consumers first have to be exposed to the label, then see and perceive the information (consciously or subconsciously). This perception can lead to understanding, which is impacted by individual consumers' liking of the label, and finally the label can be used to make a decision. The framework differentiates between subjective and objective understanding; where objective understanding is if the consumer correctly understands the label and subjective understanding is if the consumer *thinks* that they have correctly understood the label. The framework also outlines factors that can impact this pathway (i.e., label effectiveness), including interest in nutrition, knowledge in nutrition, demographics (including sex, age, ethnicity, SES) and format of the label. The effectiveness of FOPLs at equipping consumers with information to guide healthier choices relies on the ability of consumers to accurately understand the information and then to use that information when making purchasing or consumption decisions (i.e., reliant on consumer engagement and their ability to act on the provided information, such as being able to afford the healthier option).[38] These models are useful and broadly accepted, although both do not encompass all factors that may impact on behaviour change/label use (for example, interest and knowledge in nutrition does not equate to consumer motivation). The Capability-Opportunity-Motivation-Behaviour (COM-B) model is a useful framework to understand how FOPLs could improve diet quality/impact on behaviour change/healthier food purchase and consumption decisions using broader behavioural theory/general behaviour change principles. FOPLs could improve the *capability* of consumers (i.e., improve understanding of product healthiness) and having consistent labelling would provide the initial *opportunity* to use them but other factors can be limiting (i.e., cost and access to healthier options), and finally, *motivation* to use FOPLs is still needed.[39] This thesis addresses the use and feasibility of FOPLs through the following factors from the models: objective and subjective understanding of FOPLs, perception and liking of

FOPLs, demographics, interest and knowledge in healthy eating; motivations for label use. Contextual factors that impact on label use are discussed in section 'Factors that impact engagement with labels and purchasing' below. The UNICEF FOPL policy brief included a pathway that outlined how FOPLs could reduce obesity rates and ultimately decrease government spending on healthcare costs, see Figure 4.[40] The impact on obesity and the reduction in morbidity and mortality remains a big question, which is not addressed in this thesis.

Figure 2: Logic model of the process by which nutritional labelling may have an impact on diets and health (extracted from Crockett et al., 2018)[32]

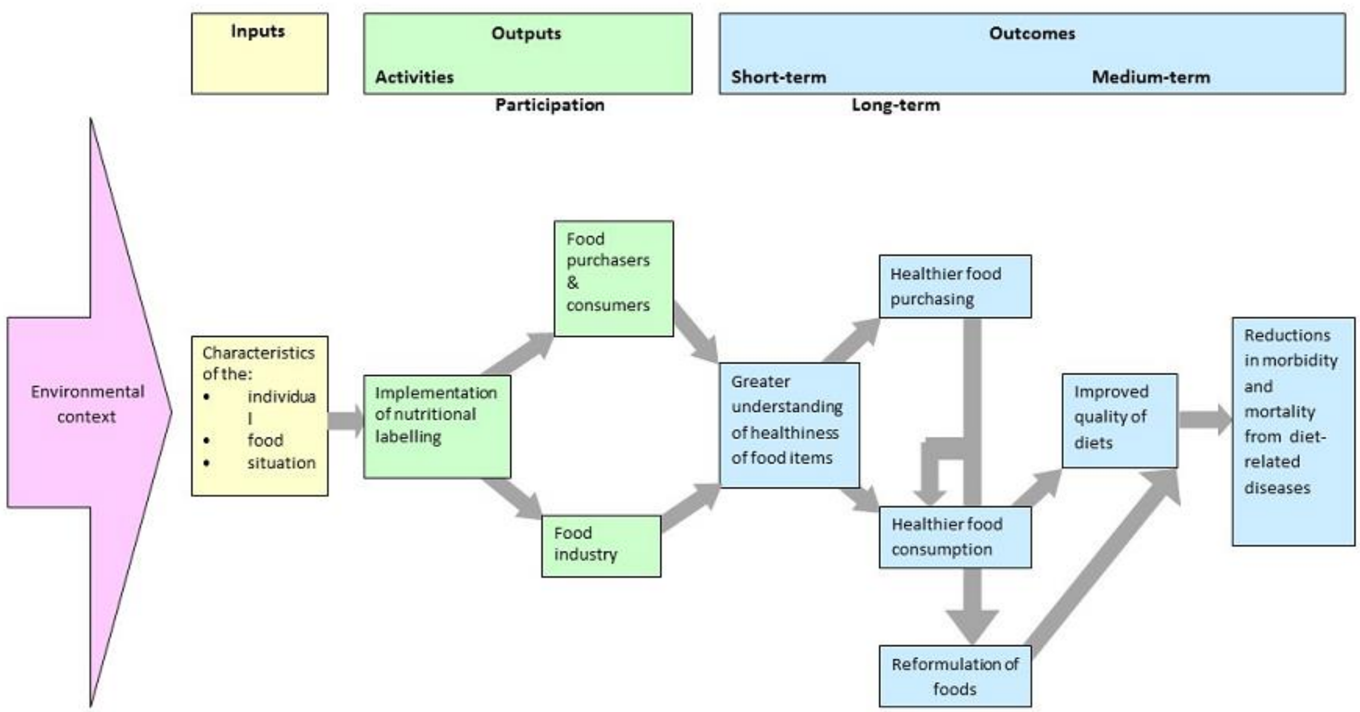


Figure 3: Theoretical model for response to nutritional information on food labels (extracted from Grunert & Wills, 2007) [37]

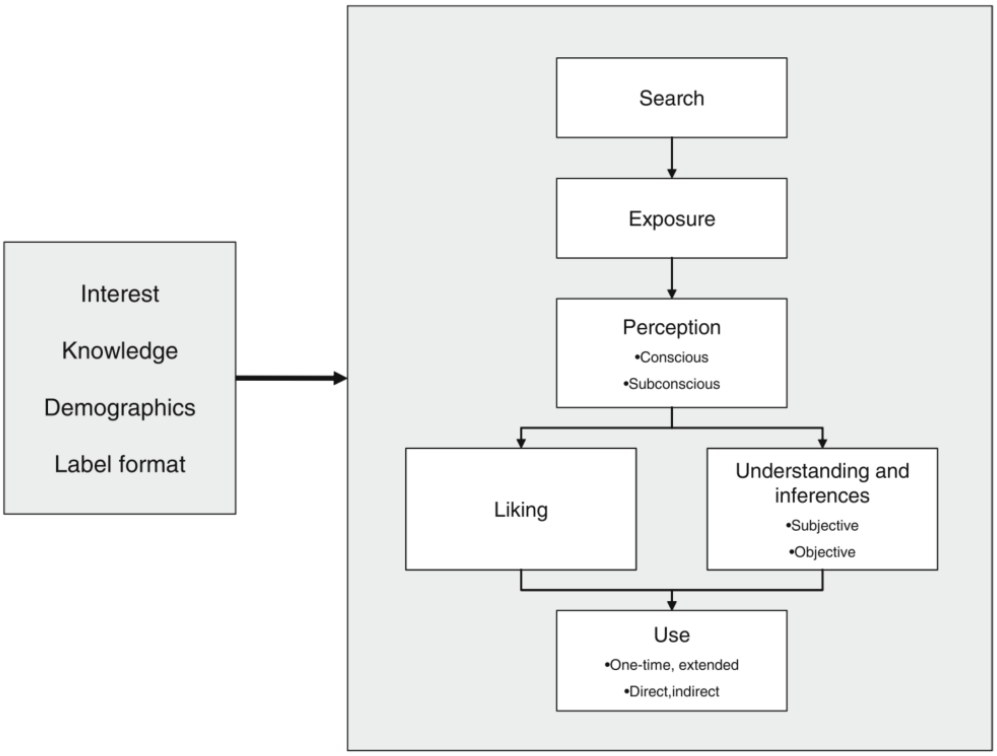
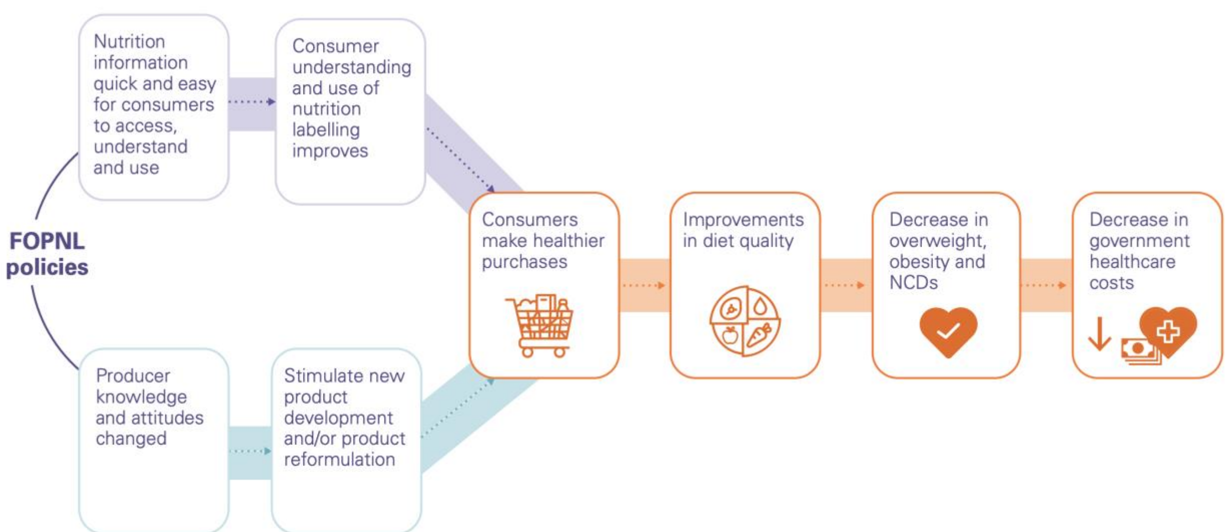
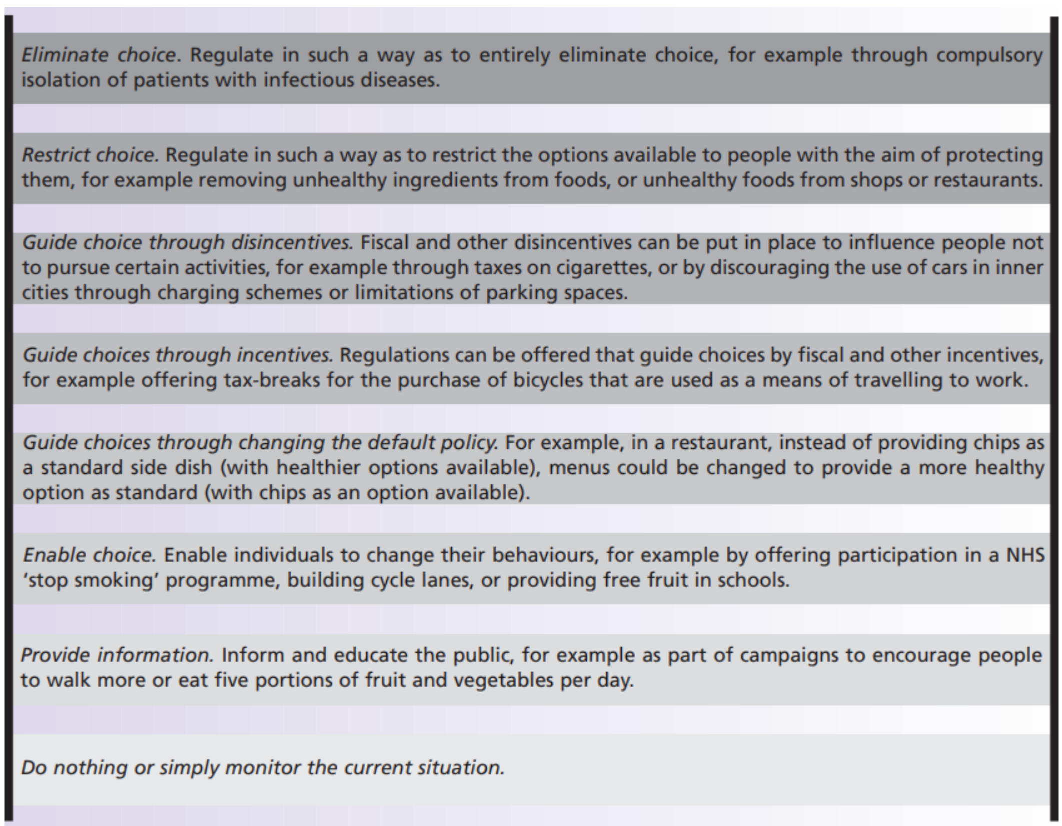


Figure 4: Pathway from the UNICEF policy brief on FOPL outlining how FOPL policies impact on healthier diets and population health outcomes [40]



The Nuffield Intervention Ladder presented in Figure 5 indicates that FOPL policy will impact level 3 (enabling choice), but could also impact level 4 through reformulation (guiding choice through changing the default by making healthier choices the default option).[41] The impact of FOPLs on product reformulation may be key to the overall effectiveness to improving diet quality (to reduce the reliance on individual agency). A study that assessed government strategies to improve nutrition and health, found point of purchase labelling could exacerbate health disparities, as smaller effects may be seen with more disadvantaged/marginalised groups.[30]

Figure 5: Nuffield Intervention Ladder (of possible government options) [41]



Factors that Impact Engagement with Labels and Purchasing

Consumer engagement with FOPLS at the point of purchase or consumption is essential for FOPLs to have an impact on improving the healthiness of choices. International research indicates that purchasing behaviours and engagement with labels are impacted by individual, social and environmental contextual factors, see Table 2 for the direction of the impact.[33]

Individual factors include comprehension of labels (e.g., nutritional knowledge), motivations for purchasing or consuming foods (e.g., interest in healthy eating, health goals, trying to lose weight), sociodemographic characteristics (socioeconomic status (SES), age, sex, ethnicity), household factors (e.g., children in household, family members with specific dietary requirements), health status, BMI, time available to make decisions and hunger at time of shopping.[33, 38, 42-51] Social and environmental contextual factors include product and label factors such as price, previous purchase or product familiarity, product or brand preference, size and position of label, consistency in placement of label, product type and taste (vice vs virtue, hedonic pleasure), other information presented on packaging (e.g., product or health claims), and broader contextual factors such as the business of shops, price promotions, trust and credibility of the label.[33, 38, 43, 44, 50, 52-54]

Table 2: Factors that impact on purchasing behaviours and FOPL engagement

Factors	**Direction of pattern**
Individual level factors (comprehension of labels, nutritional knowledge, motivations when purchasing or consuming foods, socio-demographic characteristics)	More likely: • Greater nutritional knowledge/objective understanding[33, 49] • Interested in healthy eating/health conscious[43, 49, 55, 56] • Nutrient-specific health goals[48] • Trying to lose weight[49, 55] • Females[44, 51] • As age increases[57] • As education, income, SES increases (social gradient)[43, 51, 57] • White ethnicity[51] • Special dietary requirements or need for health-related information, also if household members[43, 51, 56, 57] • Children in household[51] • Larger household[51] Less likely: • Less nutrition conscious[43] • Children, adolescents and older adults[51] • Ethnic minority[43, 47] • Higher BMI score[43] • Children in household[43]

Social and environmental factors (label and product factors)	More likely: • More time available[38, 46, 50, 53] • Lower emphasis on price[51] • Larger label size[43, 54] • Top right position[43, 54] • Consistent position[43, 54] • Nutrition is unclear e.g., soup compared to crisps[50] • Virtue vs vice e.g., for a party[44] • Taste (mixed – positively related in 3/5)[51] Less likely: • Price focused[33, 43, 50] • Price promotions[33, 38] • Hungry at the time of shopping[50] • Shops are busy/stressful[50] • Lack of understanding of terms or concerns of accuracy[53, 54] • Size of print is too small[53] • Taste (mixed – negatively related in 2/5)[33, 38] • Health claims[52] • A product is familiar/ previously purchased[44, 50] • Brand or product preference[33]

Current Evidence for Front of Pack Labels

The pathways illustrated above in Figure 2 and Figure 4 outline how FOPLs could impact on healthier diets and population health outcomes, either through improving consumer understanding and impacting on consumer behaviours, or through product reformulation. There is evidence for each of these factors, which I outline in this Chapter and the following two Chapters. In this Chapter, I narratively synthesise the findings and impact of FOPLs on consumer acceptability, understanding and use of FOPLs, and reformulation. I present this across the literature and also present evidence specifically focusing on the UK, since the impact of policies is context specific (see section 'Policy context of front of pack labels in the UK' below). Due to the importance of objectively measured understanding of FOPLs, I present a rapid review specifically addressing this in Chapter 2. In Chapter 3, I present a systematic review that examined the impact of FOPLs on objectively measured consumer behaviours, including both purchasing and consumption outcomes.

International evidence for FOPLs shows that consumer acceptability is high and that FOPLs are more likely to be used and easier to understand than the nutrition information on the back of pack.[33, 37, 49] Systematic reviews have shown that FOPLs improve the healthiness of product selection and purchases, and improve knowledge and ability to identify healthier products, in both experimental and real-life settings.[43, 52, 58-60] There is limited direct evidence that FOPLs change intake.[32, 58] A Cochrane review on nutritional labelling for healthier food or non-alcoholic drink purchasing and consumption found no effect in a meta-analysis on consumption from FOPLs from six (poor quality) experimental studies.[32] Additionally, another systematic review and meta-analysis showed food labelling could decrease calorie intake by 3.6%, although this was also not statistically signficant.[58] Research indicates that interpretive labels have the largest impact on consumer behaviour, especially simpler versions that provide an overall evaluation or use colour, compared to non-interpretive labels.[33, 44, 60, 61] Evidence further suggests that summary indicator labels are easier for consumers to understand and interpret compared to nutrient-specific labels, especially for at-risk populations, such as consumers with lower levels of education and income.[49, 62]

One aim of FOPLs is to encourage product reformulation and improve the availability of healthier foods in the food system and there is some evidence to show this can be achieved.[55, 63, 64] Evidence from a modelling study suggests that mandatory implementation could decrease obesity rates in Europe by 2.5% and save 15 million disability-adjusted life years (DALYs), compared to a no labelling scenario.[65] A policy analysis of FOPL regulation conducted in 2019, identified 32 governments who endorse FOPLs, of which ten are mandatory.[66] One example, is Chile, who introduced the WL FOPL in conjunction with marketing and school sale restrictions in 2016.[67] Policy evaluations indicate that when examining the impact on reformulation, the proportion of products classified as ‘high in’ for at least one nutrient reduced from 52% to 42% (pre vs post implementation).[67, 68]

The influence of FOPLs on understanding is central to their effects and recent experimental studies have examined this using ability to rank foods in order of 'healthiness' as a marker of knowledge and understanding. An online experimental study in 12 countries tested the ability of participants to rank product healthiness, when randomised to one of five FOPL conditions (HSR; MTL; N-S; RI; WL).[69] All FOPLs improved the ability of participants to correctly rank products, with the largest improvements seen for N-S, followed by the MTL, HSR, WL, and RI labels. An extension of this study with an additional six countries showed consistent results.[70] An online experimental study in Canada tested the ability of participants to rank food product healthiness, comparing MTL, HSR, WL and a no FOPL control.[71] All labels improved the ability of participants to identify and rank unhealthy products, with the HSR and MTL groups performing the best.

Since MTL was introduced in 2013, new labels have been developed and implemented, including N-S and WL, and research indicates they may work better in disadvantaged groups. The Chilean WL has been shown to reduce purchases of 'high in' sugar, sodium or saturated fat beverage purchases equally across higher and lower educated households.[72] The N-S labels appears to be more appealing to subjects with lower adherence to nutritional recommendations, compared to labels including MTL.[73]

Policy Context of Front of Pack Labels in the UK

Historical UK Context

Since 2014, the World Health Organization (WHO) has recommended the adoption of consumer-friendly FOPLs that are interpretative and easy to understand as a policy priority.[74] This recommendation was reiterated in 2016, as part of the WHO High Level Commission on Ending Childhood Obesity.[75] The UK government endorsed MTL on a voluntary basis in 2013, following research conducted by the Food Standards Agency (FSA), when the UK was still a member of the EU.[47, 76-78] Countries in the EU are constrained from implementing mandatory FOPLs, due to EU legislation (1169/2011).[79] Since the decision to leave the EU, the UK government committed to reviewing the current FOPL policy and assess if MTL was still 'best for Britain', by ensuring that the labelling scheme remains up to date with the evidence.[20, 80]

The evidence review included evaluating if the usage should be extended (i.e., maintain and extend the use of FOPL across appropriate products) and to examine the nature of implementation (i.e., voluntary vs mandatory). Before the MTL label was endorsed, there was a greater range of different FOPLs used by businesses voluntarily, including MTL and GDA.[81] Due to the importance of having a consistent label (for consumers to notice, become familiar with and gain confidence in using), research on behalf of the UK government was undertaken focusing on variations of MTL and GDA, to understand which option would work best in the UK market (including modelling, stakeholder, public consultation and focus groups).[47, 49] The variation that worked best at that time was traffic lights (TL) combined with text and percentage Reference Intakes.[82] The UK government and devolved administrations released specific industry guidance for the creation of the MTL label, relating to the content and appearance.[34, 83] The basic requirements, as outlined in Figure 6, are demonstrated by the label shown in Figure 1. The two elements required by EUFIC regulations are the location (provided in the 'principal field of vision', most likely to be seen at first glance) and font size (x-height must be at least 1.2mm).[34] Other specifications for the formatting of MTL are only provided as guidance, with '*should be'* used in relation to colour (*should be* vibrant, pastel colours *should be* avoided, *should* aim to use the specified colours), and position (determined by brand position, additional information on pack, pack size and shape, space, legibility).[34]

Figure 6: Basic requirements of the Multiple Traffic Light label, from FOPL Guidance (DHSC, 2017)[34]

The basic elements of a FoP nutrition label

A FoP label developed in accordance with this guidance will contain:

- Information on the energy value in kilojoules (kJ) and kilocalories (kcal) per 100g/ml and in a specified portion of the product.
- Information on the amounts in grams of fat, saturated fat ("saturates")[2], (total) sugars and salt in grams, in a specified portion of the product.
- Portion size information expressed in a way that is easily recognisable by, and meaningful to the consumer. For example, ¼ of a pie or 1 burger.
- % RI information based on the amount of each nutrient and energy value in a portion of the food.
- Colour coding of the nutrient content of the food.

Companies may additionally include the descriptors "High", "Medium" or "Low" (HML) together with the colours red, amber or green respectively to reinforce their meaning.

The FoP label design must not mislead or confuse the consumer.

Guidance for how consumers should use MTL to improve the healthiness of diets is also provided in the specification document, stating: [83]

- "If you see a red on the front of the pack, it does not mean that you should not or cannot eat it, but that you should try to keep an eye on how often you choose these foods, or how much of them you eat. A diet with fewer reds can help you achieve a healthier diet"
- "Green means the food is low in that specific nutrient that you may wish to avoid overconsuming to improve your diet. The more green colours, the healthier the choice, but you don't have to eat only green colour coded foods, including a few ambers and reds can be part of your balanced diet and will help you to get all the beneficial nutrients you need."

Use of MTL is also incorporated into the Eatwell Guide, a tool created by the former PHE, see

Figure 7 and in NHS online guidance outlining how MTL can be useful when making quick decisions in a supermarket, see Figure 8.[84, 85]

Figure 7: Eatwell Guide, with front of pack label guidance in the top left corner [84]

Figure 8: NHS guidance on how to use Multiple Traffic Light label when in a supermarket [85]

Food shopping tips

You're standing in the supermarket aisle looking at 2 similar products, trying to decide which to choose. You want to make the healthier choice, but you're in a hurry.

If you're buying pre-packaged or ready meals, check to see if there's a nutrition label on the front of the pack, and then see how your choices stack up when it comes to the amount of energy, fat, saturated fat, sugars and salt.

If the nutrition labels use colour coding, you'll often find a mixture of red, amber and green.

So when you're choosing between similar products, try to go for more greens and ambers, and fewer reds, if you want to make a healthier choice.

Evaluations of MTL in the UK and Context of this Thesis

I have examined the impact of FOPLs across the literature and here I specifically focus on UK evidence, since the impact of policies is context specific. The use and impact of the voluntary MTL scheme in the UK has been assessed as an "overall success" by the UK government, even though there is no legal duty for companies to implement it.[81] Following the announcement of the preferred system in October 2012, by the rollout in 2013, the UK government stated that all major supermarkets and other major food companies had voluntarily signed up, indicating MTLs would be found on 60% of the food sold in the UK.[86] Recent research on the prevalence of labels in UK supermarkets shows that this figure is consistent, with adoption found to be 59% (when examining all products from these three categories: chilled beef burger, Greek yoghurt and muesli categories across five UK online supermarkets).[87] Of the products with MTLs (59%), they were mostly seen on own brand products (64% vs 19% for branded products) and on healthier products with less than or equal to ≤ 1 red traffic lights (96% vs 62% for less healthy products, with 2 or more red traffic lights), compared to branded (19%) and less healthy products (≥2 red traffic lights, 1%). Compliance with one or more elements of the specification guidance was also just 62%, highlighting various issues with voluntary implementation. Another aspect of implementation is the display of FOPLs in online supermarkets/food retailers. In 2018, £12.3 billion was spent on online grocery shopping (a 7% share of total grocery retail sales), a figure that has increased since COVID-19.[88, 89] Currently, there are no specifications for how FOPLs need to be presented online (e.g., supermarket websites). A content analysis comparing consistency of FOPLs between five online supermarket retailers compared to physical supermarkets, found information was inconsistent within retailers, across retailers and with in-store packaging (for example, inconsistencies between online product information and in-store products ranged from 25% Tesco - 90% Morrisons).[90] It was possible to purchase all items without a visible FOPL and consumers commonly needed to scroll to the product details in order to view the FOPL on the online stores.

In terms of how successfully the MTL label is understood and used by consumers in the UK, Kantar research commissioned by DHSC in 2016 found: most people looked at MTL labels (80%) but a substantial proportion never do (20%); labels were seen as helpful for a healthier diet; people that reported looking at MTL had healthier shopping baskets (less sugar, fat, salt, energy and more fibre); motivations for shopping impacted use (e.g., people shopping for their families were most likely to look at FOPL); and lower SES groups paid less attention to FOPLs.[81] In terms of MTL components, the traffic light colours and number of calories were seen as useful; sugar was viewed as the most important nutrient; but RI information, recommended portion size and grams of nutrients were viewed as least useful or difficult to interpret. Recent data from wave seven of the Northern Ireland 'Eating Well Choosing Better' tracking survey showed that in 600 adults: 87% recognised MTL, 79% reported knowing what MTL is for, 64% chose foods with healthier colours always/most times, males and low SES groups were less likely to use labels/seek healthier options, and the main nutrients of interest were sugar, salt and then saturated fat.[91] Decision making with traffic light colours has been shown to be complex when products have contradictory levels of nutrients to avoid e.g., high sugar and low fat; and the detailed information could be particularly discouraging for consumers with lower education.[90] Qualitative research provides further insight into these difficulties, with participants reporting that "some nutrients were high on one label and other nutrients high on the other label;" "to put it literally, it gives me a headache, and I just put it down. Perhaps I'd just go for something I'm used to … that's what I'd do;" and "I would get annoyed … because it should be easy".[92]

Experimental research in representative UK samples is limited, so DHSC commissioned OPRU (an independent research body) to evaluate which FOPLs are most effective for identifying healthier choices, in a nationally representative sample and across sociodemographic groups. This work was complementary to a four-nation public consultation on FOPLs in the UK, which together will form the evidence base to support any changes to FOPL policy, to maximise effectiveness.[81]

Project Plan

This thesis will seek to evaluate the effectiveness of different FOPL schemes in a UK population and to assess other factors that are likely to impact on engagement with FOPLs, to inform UK policy development. Using an online randomised controlled experiment in a nationally representative British sample I will: investigate whether FOPLs improve participants' ability to identify the healthiness of foods and drink products, between schemes and compared to no-label control; explore if the ability to understand the FOPLs is associated with sociodemographic factors or engagement characteristics (time to use, motivations, food habits, label perceptions); and assess which FOPL scheme may work best across the entire UK population/all sociodemographic groups.

In Chapter 2, I present a literature review outlining the evidence for FOPLs and the impact on understanding of product healthiness. In Chapter 3, I present findings from a systematic review and meta-analysis of evidence that assessed impact of FOPLs on objectively measured consumer behaviours (purchasing and consumption). In Chapter 4, I describe the design and methodology of the online randomised controlled experiment, including the rationale, ranking task logistics, participant and statistical information. I assess the ability of participants to correctly understand product healthiness in Chapter 5, with and without FOPLs. I then assess other factors that impact on engagement with FOPLs in Chapter 6, including time needed to correctly use labels, perceptions of their ability, perceptions of the FOPL schemes and examine the association with sociodemographic factors. In Chapter 7, I examine if the outcomes from the experiment differ by sociodemographic and engagement characteristics, using latent class analysis. In Chapter 8, I describe a patient and public involvement session conducted with children and young people, to assess their ability to use FOPLs and understand their views and suggestions for FOPL policy.

PhD thesis aim:

1. Explore the best FOPL options for the UK, in terms of label effectiveness for understanding healthiness and factors of engagement, across sociodemographic groups.

PhD thesis objectives:

1. To systematically review and describe the evidence on impact of FOPLs on consumer purchasing and consumption behaviours (Chapter 3).
2. To examine if FOPLs improve the ability of participants to rank products according to healthiness, compared to no-label control (Chapter 5).
3. To examine if specific FOPL schemes (MTL, N-S, WL, PC) improve the ability of participants to rank products according to healthiness (Chapter 5).
4. To examine if ability to determine product healthiness differs by participant characteristics (sociodemographic, contextual) (Chapter 5).
5. To examine if perceptions of label use, liking and subjective understanding differ by FOPL scheme (Chapter 6).
6. To examine if confidence of ranking varies by FOPL scheme, compared to control (Chapter 6)
7. To examine if time taken to complete rankings differs by FOPL scheme, compared to control (Chapter 6)
8. To understand if engagement characteristics (food habits/motivations) differ by sociodemographic characteristics and impact on effectiveness of FOPLs (Chapter 7)

Chapter 2: Literature Review on Front Of Pack Labels and Objective Understanding

Chapter 2 Background

In this Chapter, I will describe the literature on front of pack labels (FOPLs) relating to the impact on objective understanding and knowledge of the nutritional value of products. In Chapter 1, I outlined the overall aims and background on FOPLs, and briefly described the impact of FOPLs on consumer understanding. In Chapter 3, I present the findings of a systematic review assessing the impact on consumer behaviours (purchase and consumption). The literature review presented in this Chapter was conducted to aid in the design of the experiment that is presented in Chapters 4-7.

As discussed in Chapter 1, FOPLs aim to improve the diet quality of consumers through two pathways, the first is to improve consumer understanding of product healthiness, with the aim to then improve the healthiness of the products they then purchase and consume. This literature review aims to address this pathway, relating to the short term outcome of improved understanding of the healthiness of food items presented in the Crockett et al. logic model and the objective understanding presented in Grunert and Wills.[32, 37] This rapid literature review was conducted to guide the development of the experiment (presented in Chapter 4-7), as the policy request was for empirical research to test whether specific FOPLs help improve ability to identify healthier food and drink choices. Objective understanding is a more robust measure than subjective understanding and is an important factor for label use, as outlined in the conceptual framework by Grunert and Wills.[37] Therefore, the impact of FOPLs on the objective understanding of product healthiness was the focus of this rapid review.

Literature Review Aim

To rapidly summarise the evidence on the impact of FOPL schemes on objective understanding or knowledge outcomes of product healthiness/nutritional quality.

Chapter 2 Method

Searches were conducted for experimental studies with an exposure of any FOPL scheme on any food or drink products. The included outcomes were any objective measures of knowledge or understanding of product nutrition/healthiness, and included comparisons were no FOPL control or to another FOPL scheme. Studies were excluded if they were observational, qualitative, reviews, dissertations, conferences and if the exposures were nutrition claims (e.g., natural, 'low in fat', vegan), allergens (gluten free), or labelling at restaurants. Searches were conducted in July 2020 on PubMed before the experiment was designed and re-ran on the 12th of April 2021 before the submission of the final briefing paper to the Department of Health and Social Care (DHSC). I also conducted additional citation searching of relevant systematic reviews (see Table A1 in Appendix A for full search strategy).

Chapter 2 Results

The search resulted in 263 articles, of which 28 articles were included on title and abstract and then 17 were included on full-text (of the 11 excluded on full-text, 6 were due to exposure and 5 due to outcome). An additional 10 articles were found through citation searching, leading to a total of 27 articles (29 studies) included on full-text. The descriptive characteristics of the included studies are presented in Table 3.

Table 3: Descriptive table of experimental studies with understanding or knowledge outcomes

Author, Year, Country	Design, FOPL scheme	Outcome measure	Finding
Borgmeier, 2009, Germany	Research lab, *n* = 420 Randomised to one of five FOPL conditions: -no FOPL control -simple "healthy choice" tick -MTL -monochrome GDA -coloured GDA	Healthier product identification (28 pair-wise comparisons). Food portion selections for total daily intake.	Number of correctly identified healthy products differed significantly between the conditions. The no label condition had lowest number of correct healthier choices and TL had the highest number. Food portion selections did not differ significantly between the conditions.
Roberto, 2012, US	Online, *n* = 703 Randomised to one of five FOPL conditions: -no FOPL control -TL -TL plus information about protein and fibre -FUF -FUF plus information about "nutrients to encourage"	Pair-wise selection of product with higher/lower levels of nutrients (saturated fat, sugar, sodium, calories, protein, and fibre). Estimated levels of saturated fat, sugar, sodium, fibre, and protein in 8 products (low, medium, or high) and estimated calories per serving.	All FOPL conditions outperformed the control groups for the total nutrient comparison quiz scores and calorie and sodium estimation. FUF+ and TL+ groups had the highest nutrient comparison quiz scores, they did not differ from each other. FUF labels more likely to wrongly estimate nutrients (underestimated sat fat and sugar, overestimated fibre and protein) compared to control and TL labels.
Maubach, 2014, NZ	Online, *n* = 768 Randomised to 2 of 12 conditions: 4 label conditions: -NIP -DIG -MTL -HSR	Rated the perceived healthiness of cereals from 1 to 10 (1 = very unhealthy & 10 = very healthy) of two muesli products from 12 possible nutrition labels (four label types X three nutrition profiles (good, moderate, poor).	Healthiness ratings differed most between the three nutrition profiles with the MTL labels, with significant differences between the scores. Rating differences were smaller but still significant for DIG and NIP. HSR had smallest difference in ratings between.

	X 3 product nutrition profiles: -good -moderate -poor		the products, no significant difference between good and moderate).
Julia, 2016, France	Shopping lab, *n* = 901 Randomised to 1 of 3 conditions: -no FOPL control -5-CNL -5-CNL + consumer information on use and understanding of the label	Rated the healthiness of two products with labels after 2 second exposure (Among the two products I showed you, which do you believe is the healthier one?).	Objective understanding was significantly higher in the intervention groups, and more importantly in the label +communication group.
Egnell, 2018, 12 countries*	Online experiment, *n* = 12,015 Randomised to 1 of 5 conditions: -RI -MTL -WL -N-S -HSR	Ranking of products by healthiness (3 categories with 3 products).	All FOPL conditions improved compared to baseline ranking. N-S performed best, followed by MTL, HSR, WL and then RI.
Goodman, 2018, Canada, US, UK, Australia	Online experiment, *n* = 11,617 Randomized to 1 of 11 conditions: -no FOPL control -red circle -red 'stop sign' -magnifying glass	Rated the levels of nutrient content (saturated fat/sugar) of a cereal product as low/ moderate/ high/ don't know/ refuse.	FOPLs with 'high in' text led to significantly greater likelihood of correct identification of products 'high' in sat fat or sugar, compared to control.

	-magnifying glass + exclamation mark -'caution' triangle + exclamation mark -plus, each of these 5 conditions with a 'high in' text descriptor		
Liu, 2018, UK	Online, *n* = 812 Randomised to 1 of 4 conditions: -calorie text -activity text -calorie image -activity image	Estimation of food product (milk/chocolate) contribution to recommended daily intake of energy on 0-100 percent scale.	Overall, contribution of the food products to the daily recommended intake was overestimated. There was no difference between the FOPLs.
Roseman, 2018, USA	Online, *n* = 161 Randomised to 1 of 4 conditions: -no FOPL control -FUF -FUF Extended -a binary symbol	Selected healthier product from food product pairs (healthy and less healthy option) from three product categories (cereal, dairy, and snacks).	Greater proportion of participants selected the healthier (more nutrient dense) option in the snack food category, compared to control. In cereal and dairy categories most participants, across all conditions, correctly identified the products.
Khandpur, 2018, Brazil	Online, *n* = 1,607 Baseline (no label) and then randomised to 1 of 2 conditions: -TL -WL	Selection of healthier product (out of two or three products); identification of excessive nutrients (sugar/ sodium/ saturated fat/ 'None of these nutrients are in excess') in three products (snack, biscuit, lemonade); identification of products with larger quantities of nutrients (sugar/ sodium/ saturated fat).	Both labels led to improved understanding of excessive content of nutrients and ability to identify the healthier product, compared to baseline with no labels. WL led to largest improvement in understanding of excessive content of nutrients and ability to identify the healthier product, compared to TL labels.

Vargas-Meza, 2019, Mexico	Online, *n* = 2,105 Randomly assigned to 1 of 3 conditions: -Mexican GDA -Ecuadorian MTL -Chilean WL in red	Selected the product with the lowest nutritional quality among three products.	Higher odds of correctly identifying the least healthy product with both WL and MTL label, compared to the GDA.
Khandpur, 2019, Brazil	Online, *n* = 2,419 Randomly assigned: -no FOPL control -Triangular WL 'A lot of /Muito' (Muito) -Triangular WL 'High in/Alto em' (Alto) -Chilean WL 'High in/Alto em' (Chile)	To assess understanding of the nutritional content of nine products ('None'- 'A lot' of total fat/sugar/ sodium/ sat fat/ trans-fat/ artificial sweeteners) and product healthfulness (which is healthier out of products/ which of pair has more nutrients), identification of excessive nutrients (sugar/ sodium/ saturated fat/ 'None of these nutrients are in excess').	WLs significantly improved understanding and perceptions of the nutrient profile of products, compared to control.
Andreeva, 2020, Bulgaria	Online, *n* = 1,010 Randomised to 1 of 5 conditions: -RI -MTL -WL -N-S -HSR	Ranking of products by healthiness (3 products of varying nutritional quality in 3 categories- cake, pizza, cereal, coded as correct/incorrect) at baseline and then randomised FOPL condition.	N-S, followed by HSR, led to significant improvements in ranking ability across all products, compared to RI. MTL and WL did not significantly improve ranking ability, compared to Reference Intakes.
Egnell, 2020, 12 EU countries	Online, *n* = 12,391 Randomised to 1 of 5 conditions: -RI -MTL -WL -N-S -HSR	Ranking of products by healthiness (3 products of varying nutritional quality in 3 categories- cake, pizza, cereal, coded as correct/incorrect) at baseline and then randomised FOPL condition.	N-S led to largest improvement in ranking ability, compared to RI, then followed by WL, MTL and HSR.

Vandevijvere, 2020, Belgium	Online, *n* = 1,007 Randomised to 1 of 5 conditions: -GDA -MTL -WL -N-S -HSR	Ranking of products by healthiness (3 products of varying nutritional quality in 3 categories- cake, pizza, cereal, coded as correct/incorrect) at baseline and then randomised FOPL condition.	N-S, followed by MTL, led to significant improvements in ranking ability across all products, compared to GDA. HSR and WL led to significant improvements in cake category only, compared to GDA.
Fialon, 2020, Italy	Online, *n* = 1,032 Randomised to 1 of 5 conditions: -MTL -WL -N-S -HSR -RI	Ranking of products by healthiness (3 products of varying nutritional quality in 3 categories- cake, pizza, cereal, coded as correct/incorrect) at baseline and then randomised FOPL condition.	Proportion correct increased from baseline to FOPL condition for N-S, HSR and MTL. N-S, followed by HSR, led to significant improvements in ranking ability across all products, compared to reference intakes. MTL and WL led to no significant differences in ranking, compared to RIs.
Vanderlee, 2021 Canada	Online, *n* = 1,997 Randomised to 1 of 4 conditions: -no FOPL control -WL -MTL -HSR	Ranked 3 sets of 5 products from healthiest to least healthy -identify the healthiest product -identify least healthy product -rank products according to healthiness.	Across all three outcomes (identify the healthiest and least healthy products and rank products according to healthiness) participants in HSR were significantly more likely to be correct, compared to all other conditions. The MTL condition was significantly more likely compared to the WL and control conditions. WL was significantly more likely than control.
Ares, 2021, Uruguay	Online, *n* = 1,772 After-only design (before and after implementation of FOPL)	Three products with varying nutritional compositions from three categories (cookies, juice and yoghurt): -Identification of healthiest product -Identification of excessive content of sugar, fat, sat fat and sodium.	After implementation of WL, the ability to identify healthiest product significantly increased, compared to no label. Ability to identify excessive nutrient content was also significantly higher after the implementation of warnings.

	-no label baseline, before implementation of FOPL (baseline, May-June 2019) - WL, after implementation (follow-up, March 2020)		
Ducrot, 2015, France (Effectiveness)	Online, *n* = 13,578 Each participant saw each of these labels: -no FOPL control -GDA -MTL - N-S (5-Color Nutrition Label) -Green Tick	Ranking of three products according to their nutritional quality, five categories (prepared fish dishes, pizzas, dairy products, breakfast cereals, and appetizers).	All FOPL significantly increased the percentage of correct answers compared to control. 5-CNL/N-S performed best, followed by MTL, GDA and then tick. N-S was significantly higher than MTL in 2 categories, GDA in 5 categories and Tick in all. MTL was significantly higher than GDA in 2 categories, although GDA was higher in 1, and all tick.
Hodgkins, 2015, UK, Germany, Poland and Turkey	Online, *n* = 2,068 Randomised to 1 of 4 FOPLs: -GDA - TL -GDA+TL hybrid (MTL equivalent) -Health logos (positive tick equivalent) All used basic label first (BL) (TL format, just nutrient amount in grams)	Ranking of products by healthiness (3 categories with 3 products, with objective scores) on a 15-point scale: 1 = least healthy, 15 = most healthy.	Trend for improved objective understanding with one of four FOPLs, compared to basic, but this was not significant. All FOPL were similar.
Grunert, 2010, UK	At-home survey, n= 921 Each participant saw each of these labels: -GDA -TL	Understanding of FOPL and BOP information (3 products) -lowest in sat fat/serving -lowest in calories/ 100g -highest GDA for sugar	All FOPLs had high level of understanding (72-92% of participants correct. Results were similar and differences between the labels were not tested.

	-GDA+TL hybrid (MTL equivalent)	-more than half of GDA for fat -most salt	
Raats, 2015, Germany, UK, Spain, France, Poland and Sweden	Online, *n* = 13,117 Randomised to 1 of 3 reference amount conditions: 'per 100 g', 'typical' portion, 50% reduction on 'typical' portion	Healthiness rating of product pairs (three product categories: biscuits, sandwiches, yogurts). Without GDA and then with energy GDA (positioning of slider location between 'least healthy food you can think of' and 'most healthy food you can think of').	Participants correctly rated/ranked foods according to objective healthfulness with or without additional GDA (both had reference amounts).
Gorski, 2018, US	Online, *n* = 1,247 Randomised to 1 of 6 FOPLs: -no FOPL control -single TL -MTL -FUF -NuVal -0-3-star ranking	Healthier product identification (8 pair-wise comparisons); product with greater specific nutrient quiz (8 pair-wise comparison); and individual nutrient quizzes (7 products) estimate low/medium/high level of nutrients (saturated fat, sugar, sodium, protein, and fibre) and calories per serving estimation (7 products).	All FOPLs led to greater accuracy in healthier product identification compared to no label, except for STL; NuVal and MTL led to the highest scores. All FOPLs significantly increased ability to identify product with higher/lower nutrients compared to no label, except for MTL; STL was significantly greater than all other conditions. MTL significantly greater compared to all other labels for individual nutrient quizzes in sat fat, sugar and sodium. No differences for protein or fibre. All FOPLs significantly improved ability to accurately estimate calories/serving compared to no label, except for NuVal. STL, MTL and FuF led to significantly higher scores compared to NuVal and 3-star.
Feunkes (S1 + S2) UK, Germany, Italy, Netherlands	Online, *n* = 1,630 Randomly assigned 3 out of the 6 conditions:	<u>Study 1</u>	<u>Study 1</u> Smileys and Stars were significantly better at differentiating between

(Only UK, Italy for S2)	-HCT -HPF -Stars -Smileys -MTL -WoH Study 2, *n* = 776 Randomly assigned 2 of 4 labels: -HCT -Stars -GDA -MTL	Difference between perceived healthiness of product pairs rated on 1-5 scale (9 pairs of a healthier and less healthy product). Study 2 Change in perceived healthiness of product pairs (2 pairs of a healthier and less healthy product) = difference between baseline score and score after FOPL exposure.	healthier and less healthy products, compared to all other FOPLs. HCT, followed by Smileys and Stars, were the most consistent differentiator across product categories. The most inconsistent differentiator was MTL, followed by the HPF and WoH. Study 2 FOPLs increased the perceived healthiness of the healthier products and slightly decreased the healthiness of the less healthy products, compared to baseline. There was no consistent pattern between FOPLS.
Acton, 2018, Canada	Shopping mall, *n* = 675 Randomly assigned to 1 of 4 conditions: - no FOPL control - numeric rating - HSR - simplified TL (STL)	Rated the products healthiness of soda, chocolate milk and unflavoured milk (coded as correct if response = unhealthy, moderately healthy or healthy).	Significantly greater proportion of participants in FOPL conditions correctly perceived healthiness of chocolate milk as 'moderately healthy', compared to no-label control. HSR and STL conditions led to significantly greater proportion of participants correctly healthiness, compared to numeric rating. No differences across conditions for unflavoured milk or soda.
Arrúa, 2017, Uruguay	Online, *n* = 32 (study 1), *n* = 387 (study 2) Randomly assigned 1 of 3 conditions: -GDA -TL -WL	Study 1 Identification of products with high sodium (present/absent) across three products (chicken nuggets, hamburgers and instant soup) with medium or high sodium content and time taken to respond. Study 2	Study 1 Response times to identification of high sodium content to were significantly faster for warning systems, followed by TL and then GDA. Study 2

		Selection of the most healthful product (between 2 or 3 products).	Correct identification of most healthful product was significantly different, highest proportion in TL and WL, compared to the GDA condition. High rates of correct answers overall (95%).
Aschemann-Witzel, 2013, Germany + Poland	Shopping centre, *n* = 1,000 Randomised to one of ten conditions: -9 labels with combinations of: -GDA (4 with) -Colour coding (4 with TL, 2 with blue) -Text description of nutrient content low/ medium/high (4 with) -Healthy choice tick (basic label, not all products qualified, used as 'control')	Healthy choice task – chose which product they thought was most healthful from 10 and 20 products (healthiness based on SSAg/1). Healthy sorting task – sorted 10/20 products into three piles "healthful", "neither healthful nor un-healthful" and "not healthful".	No difference in healthy choice task between the basic label and other FOPLs. Text on labels increased accuracy in sorting task. Colours on TL labels, compared to blue and no colours significantly increased accuracy in German consumers.
Edge, 2014, US	Online, *n* = 7,363 Randomised to one of four conditions: -No label control -calorie -GDA (nutrients to limit -GDA+ (+nutrients to encourage)	Identification of absolute amounts and percent daily values of nutrient measurements for six products within two of four potential categories (breakfast cereals, frozen entrées, salad dressings and savoury snacks) of varying nutritional quality (high, medium, low).	Correct identification of nutritional content was more frequent in GDA+ and GDA, compared to the other conditions. This was significant in some of the tasks across different product categories.

*Argentina, Australia, Bulgaria, Canada, Denmark, France, Germany, Mexico, Singapore, Spain, UK, US. GDA: Guideline Daily Amount; MTL: Multiple Traffic Light; TL: Traffic light; WL: Warning Labels; HSR: Health Star Rating; HCT: Healthier Choice Tick; HPF: Health Protection Factor; WoH: Wheel of Health; FuF: Facts Up Front; NIP: Nutrition Information Panel; DIG: Daily Intake Guide

Study Characteristics

The setting for the majority of studies was online (*n* = 24), followed by shopping centres (*n* = 2), research laboratories (*n* = 2) and one at-home survey. The study designs included randomised controlled experiments (*n* = 13), labels randomly allocated with no control group (*n* = 13), participants exposed to all labels (n= 2), or before and after study, mirroring real-world (*n* = 1).[93] The research was conducted in a range of countries, including the US (*n* = 4), Uruguay (*n* = 3), UK (*n* = 2), France (*n* = 2), Canada (*n* = 2), Brazil (*n* = 2), Mexico (*n* = 1), NZ (*n* = 1), Germany (*n* = 1), mixed European countries (*n* = 9), mixed countries worldwide (*n* = 2). The sample sizes ranged from 32 to 13,578 participants. The FOPL exposures included N-S, Traffic light (TL, single, multiple), HSR, WL, RI (equivalent to Guideline Daily Amount, GDA and Facts Up Front, FuF), activity labels, smiley faces. The outcomes measures included the identification of the healthiest product from two or three options (*n* = 8); nutrient estimation or quiz (*n* = 10), ranking of products according to their healthiness (*n* = 9*), identification of the least healthy product (*n* = 2); rating of product healthiness (*n* = 5) (multiple outcomes so numbers exceed number of studies.

Chapter 2 Results

One experiment was published in two main articles[69, 70] (the second was combined with additional experiments) and three were country-specific articles.[94-96] Only the findings of the two main papers are presented, so data is not duplicated. Sixteen studies found that exposure/inclusion of a FOPL compared to a no label control group (*n* = 11) or no label baseline measure (*n* = 5), significantly improved the objectively measured outcome of understanding of product healthiness, and one showed a trend of improving (*n* = 1). Only one study found no difference between exposure to a FOPL and the no label control group; the FOPL was GDA, which is a non-interpretive FOPL.[97] The remaining studies did not include a control or baseline comparison, and only compared different FOPL groups (*n* = 8). There was high heterogeneity across the included studies for design type, country and exposure. The findings by FOPL scheme were inconsistent, therefore it is not possible to summarise the findings by FOPL scheme. No FOPL scheme performed best across all studies. N-S performed best in three studies.[69, 70, 98] There were mixed results when comparing TL and WL, TL performed better in one[99] similar in two[100, 101] and WL better in two.[100, 102] Six studies found no difference between FOPL conditions.[97, 103-107]

Chapter 2 Discussion

This rapid literature review set out to identify experimental studies that assessed the impact of FOPLs on objective measures of nutritional understanding/ability to determine product healthiness. A total of 29 studies were included (from 27 articles). The results showed that FOPLs can improve the objective understanding of product healthiness compared to baseline or no FOPL control. Examining the results by FOPL type/schemes indicates that interpretive systems (excluding endorsement logos) may be more effective at improving objective understanding than non-interpretive labels (n = 6).[69, 70, 98, 100, 101] Findings also indicate that interpretive aids such as summary text (e.g., 'high in') and colour may improve objective understanding, as discussed in Chapter 1.[108, 109] Interpretive FOPLs are recommended by the WHO, as the interpretation of product healthiness is more helpful and may be easier to see than the non-interpretive information.

This review was undertaken to assist with the design of the experiment that was commissioned by the DHSC, to test the ability of participants to correctly understand product healthiness. The results were unable to determine which FOPL scheme is the best choice generally, and for the UK. The summary of the results, study design and outcome measures were useful in the planning of the experiment, which is outlined in Chapter 4.

The limitations of the review include the rapid methodology, as a single database was used, although the citation searching did strengthen the search strategy. It was also conducted in 2020 and re-ran in 2021, so it is now out of date, although it was up to date for when the experiment was being designed and therefore it fulfilled the intended purpose. The outcome of interests were also very narrow, in order to aid in the design of the experiment, therefore the real-world impact on consumer behaviours and on subjective measures of understanding are not addressed. In Chapter 3, the real-world impact on consumer behaviours will be assessed with a systematic review.

Chapter 2 Conclusions

Evidence suggests that FOPLs may improve objective understanding and that this is aided by interpretive elements, but more research is needed. The impact of improved understanding on actually improving the nutritional quality of consumer behaviours (including purchasing and consumption), is unclear. In the next Chapter, I present the results of a systematic review that examined this, by assessing the impact of FOPLs on objectively measured consumer behaviours, the next step on the pathway to label engagement.

Chapter 3: Systematic Review Examining the Impact of Front of Pack Labels on Consumer Behaviours

Chapter 3 Background

This Chapter, reports the findings of a systematic review that assessed the evidence on front of pack labels (FOPLs) impact on consumer behaviours. This builds on the evidence presented in Chapter 2, where the impact of FOPLs on consumer understanding of product healthiness, was seen to be effective. The research question of this review relates to the outcomes of healthier food purchasing and consumption, which is next part of the Crockett et al. logic model pathway for label engagement, following on from improved understanding.[32]

As outlined in Chapter 1, FOPLs aim to assist consumers in making healthier food choices by informing them about the product healthiness. The literature review presented in Chapter 2 relates to the impact of FOPLs on understanding with evidence showing that they do increase understanding, but this does not always mean that this then impacts on consumer behaviours (purchasing or consumption). Two reviews have examined the impact of FOPLs, including consumption outcomes, with both finding limited studies and non-significant effects.[21, 110] Another review found that food labelling significantly reduced the consumption of energy and fat, but the included labels were not confined to FOPLs and included labels on menus and other point-of-purchase labelling.[111] Evidence for the impact of FOPLs on objectively measured purchasing is limited, although findings from one review suggested some impact on consumers choosing to purchase healthier products.[44]

The evidence on which FOPL scheme type is most effective at impacting the healthiness of consumer behaviours is also unknown. An experimental study across 12 countries with large samples, found that all five of the studied FOPL schemes (N-S, MTL, HSR, warning symbol, Reference Intakes) improved knowledge but their effectiveness varied considerably; N-S performed best, followed by the MTL, and then Reference Intakes.[69] Reviews indicate that FOPL schemes with interpretive elements (text or colour) are easiest to understand/ impact on knowledge, but again the impact on consumer behaviours were not addressed.[43, 44]

In 2018, a Cochrane review assessing the impact of nutritional labelling for food and non-alcoholic drinks on purchasing and consumption of healthier items was published.[32] This review identified six experimental studies, which examined the impact on consumption of labelling on pre-packaged foods but no impact on calorie intake was found and the studies were deemed to be of low quality. The majority of the evidence identified came from studies which considered the effect of nutritional labelling in 'out of house' food service settings (nutritional information on menus/ menu boards/ labels near food products in restaurants, cafeterias and coffee shops). This evidence supported labelling for encouraging healthier purchasing. The Cochrane review did not identify studies which measured the impact of FOPL on pre-packaged foods on objectively measured purchasing behaviours. Searches were conducted in April 2017 and, given that this is a highly active research area, it was likely that recently published studies could add to the evidence base and help to inform current policy. The aim of this systematic review was to examine studies published since the Cochrane review with a focus on FOPL on pre-packaged foods (not interested in any menu labelling or labelling placed next to a product rather than on the packaging), examining purchase and consumption, and using more inclusive eligibility criteria for purchasing outcomes (i.e., experimental purchasing outcomes were included if they were made with participants' money, allocated money or hypothetical purchases, providing the experiment was set up to reflect a realistic shopping experience). Meta-analyses were conducted with studies that used experimental conditions to compare FOPLs to a no-FOPL condition for multiple purchasing outcomes and separately by FOPL scheme.

Related publications: Croker, H., Packer, J., Russell, S., Stansfield, C.& Viner, R.M. (2020) Front of pack nutritional labelling schemes: a systematic review and meta-analysis of recent evidence relating to objectively measured consumption and purchasing. J Hum Nutr Diet. 33, 518– 537 https://doi.org/10.1111/jhn.12758

Chapter 3 Methods

The systematic review was conducted in collaboration with UCL Institute of Education and using EPPI-Reviewer 4 software.[112] The study was registered with PROSPERO (registration number

CRD42019135743, see footnote[1]), and the systematic review is reported in accordance with the PRISMA Checklist.[113] The protocol describes the work as a rapid review update, as machine learning was utilised to allow a full review to be conducted in a rapid timeframe.

Eligibility Criteria, Information Sources, and Search Strategy

Experimental and intervention studies were included, such as randomised or quasi-randomised controlled trials, controlled before-and-after studies, and interrupted time series (ITS) studies. Eligible for inclusion were participants of any age; studies from April 2017 onwards; intervention criteria of any FOPL on pre-packaged foods; and outcomes of objectively measured consumption (at individual level) and purchasing behaviour (either quantity of unhealthy/ healthier products or nutritional content of purchased products at individual or family level). The inclusion and exclusion criteria are shown in Table 1. Experimental purchasing outcomes were included if they were made with participants' money, allocated money or hypothetical purchases (providing the experiment was set up to reflect a realistic shopping experience). This was assessed from the instructions given to participants (e.g., directed to complete a weekly shop for their household), whether the environment was constructed based on real shopping environments (online or actual), and whether representative products and prices were presented. Experimental studies were required to have a no-FOPL control group.

[1] https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=135743

Table 4: Inclusion and exclusion criteria for systematic review

Inclusion criteria	
Exposure	Front of pack nutritional labels on pre-packaged foods, at individual/family level
Comparison	Randomised or quasi-randomised controlled trials (RCTs/Q-RCTs); Controlled before-and-after studies; Interrupted time series (ITS) studies; Compared a labelled product (with information on nutrients or energy) with the same product without a nutritional label
Outcome	Purchasing behaviour (individual/ family); Dietary intake (individual); Sales data (higher level)
Study design	Experimental; intervention (with a control group)
Other	All languages and countries, from 2017 onwards
Exclusion criteria	
Exposure	Any unpackaged food, at food supply level
Study design	Non-experimental studies or lack of appropriate comparison

The search strategy was adapted from the Cochrane review search, with assistance from an information scientist (CS) at the EPPI-Centre.[32] Changes were made to narrow the focus of the review to FOPL on pre-packaged foods (in retail settings or experimental contexts), excluding menu labelling in food service settings (such as menus or labels placed near foods in restaurants, cafeterias or coffee shops), and expanding the search by adding focused search terms for FOPL and named labelling schemes. Systematic searches of 13 bibliographic databases were conducted covering the research disciplines of medicine, psychology, science, social science and business: ASSIA (Proquest), ABI Inform Global (Proquest), CINAHL (EBSCO), Cochrane Central Database of Controlled Trials, Cochrane Database of Systematic Reviews, EMBASE (OVID), HMIC (OVID), Medline (OVID), PsycINFO, Sociological Abstracts (Proquest), SCOPUS, Trials Register of Promoting Health Interventions (TRoPHI), and Web of Science (Science Citation Index, Social Science Citation Index, Emerging Sources Citation Index). Further details about the search and the full search strategy for each database are included in Appendix B, see Table B1 and Table B4.

The searches were conducted on the 8th of April 2019 and records were imported into Endnote (a reference manager software) where duplicates were removed. Articles were then imported into EPPI-Reviewer 4 software where duplicate records were again assessed and removed (as EPPI has greater flexibility when setting up duplicate searches). A ‘cited by’ search was conducted using Google Scholar; studies included in the current review were used as the key papers and the 'cited by' function was used to identify any relevant articles published since the date of the main searches. The 'cited by' search was conducted on the 4th of June 2019 and this date was considered the cut-off point for inclusion in the review. A completed report of an ongoing trial identified in the Cochrane review was identified.[32]

Study Selection

Exclusion criteria were date (pre-2017), intervention (any non-nutritional FOPL including restaurant and menu labelling, shelf labelling, back of pack labelling etc.), study type (systematic reviews, dissertations, magazine articles, conference abstracts), and outcome measure (e.g., attitudes, liking, understanding, knowledge, self-reported purchasing or consumption intention). Articles were included if all other inclusion criteria were met. All studies were independently screened by two reviewers on title and abstract using the EPPI-Reviewer 4 software. All queries were reconciled by the reviewers and any outstanding queries resolved with the wider research team. Full-texts were retrieved using both web and library services. Full-text screening was independently completed by two reviewers (JP & DD) using EPPI-Reviewer 4 software and queries jointly reconciled.

Data Extraction

I extracted the descriptive data and a second reviewer (HC) double-checked, with any discrepancies mutually resolved. This included study descriptors (authors, country, publication year, and design), participant descriptors (sample size, age range and mean), comparison type, intervention type, outcome type and intake measure (if applicable). Data from experimental studies for inclusion in meta-analyses were independently extracted by two reviewers (JP & HC). Corresponding authors were contacted to provide raw data where necessary; six authors were contacted for seven studies, and all provided additional data.

Assessment of Quality

Risk of bias for the experimental studies was assessed independently by two reviewers (JP & HC) using the RoB 2 for randomised trials or ROBINS-I tool for non-randomised studies.[114, 115] The sources of potential bias evaluated were randomisation procedure, deviation from intended interventions, missing outcome data, selective reporting, confounding, selection bias and classification of interventions, as appropriate for study design. Risk of bias for the ITS studies was assessed by two reviewers (JP & HC) using a tool which was under development by researchers at the EPPI-Centre and was used in a review on standardised packaging.[116] The tool included a critique of the data sampling, data collection, measures, analysis and inferences made.[117]

Data Synthesis

Data from experimental studies were included in meta-analyses. To be included, studies were required to have compared the effect of a FOPL to a no-FOPL control (the latter included the back of pack nutrition information panel) on objectively measured purchase or consumption behaviour. I identified only two articles (comprising three studies) which reported consumption data, the outcome measures were inconsistent; therefore, these data were not suitable for meta-analysis. Studies measuring purchasing outcomes were required to report the nutritional content of purchased products and to provide mean values with standard deviations.

Chapter 3 Statistical Analysis

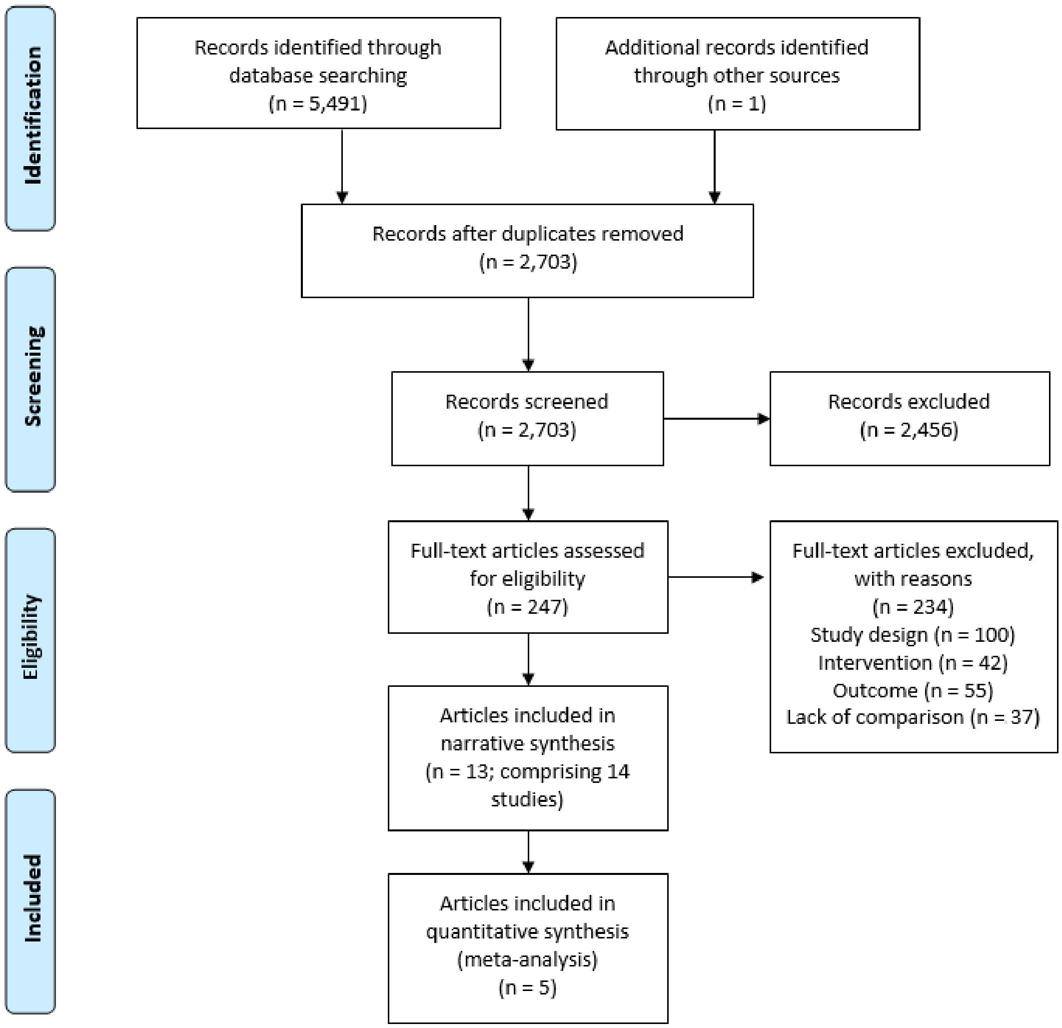
The DerSimonian-Laird random-effects model was used for meta-analysis due to the differences in the studies, including the settings (laboratory, shopping centre, online or supermarket) and measurement of purchase outcomes (‘real purchase’ task of a single product per category, simulated weekly purchase, actual purchases over four weeks). Energy outcomes were converted to kilocalories (kcals) if kilojoules were reported (4.184 kJ= 1 kcals)[118] and salt outcomes were converted to sodium (1mg sodium = 2.55mg salt).[119] All outcomes were also standardised to report the nutrient content per 100g (energy, kcal; sugar, g; saturated fat, g; sodium, mg). Four meta-analyses comparing any FOPL to a no-FOPL control were conducted, one per nutritional outcome (calories, sugar, saturated fat and sodium), with a single combined FOPL condition calculated for each study using methods consistent with the Cochrane review.[120] To allow for comparison between individual FOPL labels, meta-analyses were also conducted for the individual FOPL schemes presented as separate data points. Further information about the meta-analysis decisions, including the rationale for meta-analysis inclusions, the data points used, how experimental conditions were combined and how data were standardised to 100g is provided in Appendix B. Effect sizes (ES) are reported for each nutritional outcome (calories, sugar, saturated fat and sodium) per 100g. To assess publication bias, a funnel plot was created to assess asymmetry using Egger's regression test and trim and fill analysis.[121] Stata/SE 15.1 was used for meta-analyses.

Chapter 3 Results

Study Selection

The database searches resulted in 5,491 records, which included 2,702 unique records following duplicate removal. Screening on title and abstract resulted in 246 records which were screened on full text and assessed for eligibility. One additional record was identified through other sources (via the 'cited by' of included studies). A resulting 14 studies, from 13 articles, met the inclusion criteria (see Figure 9 for flowchart). Of these, 11 were experimental studies, three of which (from two articles) measured consumption, eight measured purchasing (three where participants used their own money and five where the study took place in a virtual shop, and the remaining three were ITS studies (i.e., 'real-world'). Five of the experimental studies reported purchasing outcomes and were suitable for meta-analysis.

Figure 9: Flow chart of searches, screening and inclusion process.



Study Description and Results

A summary of the study descriptions is provided in Table 5.

Table 5: Descriptive summary of included studies

Author, year	Country	Sample description	Design	FOPL condition	Control condition	Intervention	Outcome measures
Experimental studies reporting purchasing behaviour outcome							
Acton and Hammond, 2018	Canada	*n* = 675 Age range = 16+ Mean age = NS	Experimental (shopping centre), within-between subjects	HSR 'High sugar' symbol Text health warning	No label	Purchase of single beverage from 20 options. Given $5 CAD to complete purchase and kept change.	Per beverage: energy (kcal); free sugar (g)
Acton et al, 2019	Canada	*n* = 3,584 Age range = 13+ Mean age = NS	Experimental (shopping centre), within-between subjects	'High in' warnings MTL HSR Nutrition grade*	No label	Purchase of single beverage and a single snack, both from 20 options. Given $5 CAD to complete purchases and kept change.	Per beverage and per snack: energy (kcal); total sugar (g); sodium (mg); saturated fat (g)
Ang, Agrawal and Finkelstein, 2019	Singapore	*n* = 512 Age range = 21+ Mean age = 38.1	Experimental (online), RCT	'High in sugar' symbol (nearly identical to WL) Text-based warning	No label	Hypothetical online household grocery selection), with a compulsory spend of S$50-S$250. Following the shopping task, participants completed a brief survey and were compensated according to the web panel's in-house point system.	Total sugar (g) per shopping trip
Crosetto et al, 2017	France	*n* = 591 Age range = NS (50% 30-45)	Experimental (paper catalog + computer), pre + post and between-subjects design	NutriColors^ NutriMark+ NutriRepère# NutriScore SENS	No label	2-day household grocery selection (food and beverages), baseline and intervention shopping baskets. Required to buy	Mean FSA score of basket at follow up (per 100g)

						a subset of their selected products. Remunerated 32€ per session.	Mean FSA score decrease
Graham et al, 2017	United States	*n* = 153 Age range = NS (parent) Mean age = 38.1 (parent)	Experimental (laboratory grocery aisle), RCT	MTL Facts Up Front	Access to BOP	Laboratory grocery aisle, selection of 6 food items (2 chips/snacks, crackers/cookies, cereal). Instructed to behave as if a typical food shop and to select 6 items to take home. Compensated with the 6 selected products, a $US 20 gift card, parking remuneration and a small toy.	Mean amount per serving: energy (kcal); sugar (g); sodium (mg); saturated fat (g)
Machin et al, 2018	Uruguay	*n* = 437 Age range = 18-77 Mean age = NS	Experimental (online), between subjects	MTL system WL ('High in symbol')**	No label	Simulated weekly online food purchase (food and beverage), asked to select all products they would buy. Entered into a raffle for a voucher valued at $US 100.	Mean amount per 100g: energy (kcal); sugar (g); salt (mg); saturated fat (g)
Neal et al, 2017	Australia	*n* = 1,578 Age range = 18+ Mean age = 37.9	Experimental (online/smartphone), randomized parallel-group trial	HSR MTL Daily Intake Guide ***	NIP	Real purchases of packaged foods and beverages from the grocery store over 4 weeks, recorded by scanning barcodes, images of receipts. $AU 100 shopping voucher.	Mean amount per 100g from 4 weeks of purchases: energy (kJ); sugar (g); sodium (mg); saturated fat (g) Mean NPS score

Ni Mhurchu et al, 2017	New Zealand	*n* = 1,357 Age range = 18+ Mean age = 33	Experimental (online/smartphone), RCT	MTL**** HSR	NIP	Real purchases of packaged foods and beverages from the grocery store over 4 weeks, recorded by scanning barcodes, images of receipts. NZ $10 Voucher at run in phase, then NZ $NZ80 voucher after completion of final questionnaire.	Mean amount per 100g from 4 weeks of food purchases: energy (kJ); sugar (g); sodium (mg); saturated fat (g); fibre (g); protein (g); fruit + vegetable, nut, and legume points Mean NPSC of food and beverages
Experimental studies reporting consumption outcome							
Tangari et al, 2019, study 1	Australia	*n* = 140 Age range = 20-49 Mean age = 26	Experimental (lab), between-subjects design	Serving size only Serving size with calorie information	No label	Snack consumption (crackers), time ns.	Intake (g)
Tangari et al, 2019, study 5	Australia	*n* = 134 Age range = 18-52 Mean age = 23	Experimental (lab), between-subjects design	Standard serving information + calorie Double serving information + calorie	No label	Snack consumption (Pringles-style crisps), time ns, instructed to eat as many as they wanted.	Intake (g)

Zandstra, Willems and Lion, 2018	United Kingdom	*n* = 493 Age range = 24-65 Mean age = NS	Experimental (lab), between-subjects design	30% reduced in salt Same great taste, less salt More taste, less salt and kids will love it	No label	Served chicken noodle soup, with salt shaker.	Salt consumption (g)
Interrupted time series studies							
Araya et al, 2018	Chile	Customers at a major grocery store chain (125,485 consumers, 210,819 eligible transactions	Mandatory FOPL regulation vs. pre regulation period (6 stores) May- July 2016 vs. May-July 2015	WL ('High in' warnings) (Mandatory)	-	-	Consumer level point-of-sale data (price, date, time for purchase of fruit juice, breakfast cereal, chocolate & candy and cookies)
Elshiewy and Boztug, 2018	United Kingdom	Customers at a major grocery store chain, 188,062 loyalty card members, 4,131,570 purchases	In store brand products with FOPL vs. other products 2007 vs. 2006	GDA (Voluntary)	-	-	Energy content of purchases from scanner data for four-week aggregated periods (biscuits, breakfast cereal, and soft drink)

Peñaherrera, 2018	Ecuador	Random sample of 1646 households	Traffic light labelling mandatory vs. pre-regulation, Jan 2013-Aug 2014 vs. Sep 2014-Dec 2015	TL (Mandatory)	-	-	Purchase of soft drinks from Kantar panel and mean nutritional content (sugar, g) from ARCSA data.

FOPL, front of pack label; HSR, Health Star Rating; BOP, back of pack; NS, not stated; CAD, Canadian dollar; MTL, Multiple Traffic Light; WL, Warning Label; GDA, Guideline Daily Amounts; RCT, randomised controlled trial; cal, calories; kcal, kilocalories; g, gram; kJ, kilojoule; FSA, Food Standard Agency; NPI, Nutrient Profile Index; NIP, Nutrition information panel; NPS, Food Standards Australia New Zealand nutrient profile score; NPSC, Food Standards Australia New Zealand Nutrient Profiling Scoring Criterion; ARCSA, Agency National Regulation, Control and Health Surveillance of Ecuador. *Equivalent to Nutri-score; ** Chilean Warning system; ***Equivalent to Facts Up Front ****Text stated that the comparison group was a 'Traffic Light' label but assumed from reference provided that this was a Multiple Traffic Light label; ^Equivalent to Multiple Traffic Light, + Equivalent to Health Star Rating, # Equivalent to Facts Up Front.

Experimental Studies

Settings

Included studies were conducted in a range of countries: two in Canada by the same authors,[122, 123] two in Australia,[124, 125] and then the US,[126] Singapore,[127] France,[128] Uruguay, [129] New Zealand,[130] and the UK.[131] The study setting varied, some were conducted in shopping centres with shoppers invited to participate,[122, 123] online or using a smartphone,[124, 127, 129, 130] a paper catalogue for an experimental food store with online e-shopping environment,[128] an experimental food store,[126] or in a laboratory.[125, 131]

Participants

Participants in the majority of studies were adults (18 years old+), but three studies included adults and children; two studies had a minimum age of 16 or 13 years[122, 123] and the third included parent and child dyads, where children were aged 6-9 years.[126]

Comparisons

Experimental conditions varied between studies. For the three studies reporting consumption data, one used a nutritional label containing information about salt content[131] and the other two used serving size and calorie information.[125] Eight studies reported impacts on purchasing and all comprised more than one comparison group. Five studies included the HSR or similar;[122-124, 128, 130] six studies included a MTL group;[123, 124, 126, 128-130] two studies included N-S or similar;[123, 128] one study included the Chilean WL;[129] three studies included other 'high in' warnings (including 'high sugar' warning in a red circle);[122, 123, 127] one study included the Facts up Front label;[126] and one study included the DIG,[124] both label types were considered 'straight up' (or non-interpretive) nutritional information. Other comparisons included text-based warnings.[122, 127, 128] All of the experimental studies included a no FOPL control (as per the inclusion criteria), the majority used a no label information group as the comparison but two studies used a 'nutrition information panel' on the back of the product[124, 130] and in one study, the back of pack label was accessible to participants as they were able to pick up the products.[126]

Outcome Measures

The reported outcomes varied between studies. The studies examining consumption reported either food intake in grams[125] or salt intake in grams.[131] For the studies reporting purchasing, the outcomes were the sugar and energy content of a single drink product,[122] the sugar, energy, sodium and saturated fat content of a single food or drink product,[123] mean sugar, energy, sodium and saturated fat content of six items purchased,[126] sugar content of a single shopping trip,[127] or sugar, energy, sodium and saturated fat of a weekly shop[129] or grocery shopping over a four week period.[124, 130] One of these studies reported the nutritional content of both a purchased snack and beverage, these outcomes were included as separate data points.[123] Three studies reported the mean 'nutrition score', one for a basket of shopping using the 'Food Standard Agency' score[128] scale 1-100 (1 = least healthful and 100 = most healthful) and two for purchases over four weeks using the Food Standards Australia New Zealand Nutrient Profiling Scoring Criterion or calculator.[124, 130] For the purchasing outcomes, four studies were experiments in real-world settings,[122-124, 130] two studies were in labs[126, 128] and two studies were online.[127, 129] The purchasing outcomes also varied in size and contents of purchase, two studies were single purchases of a drink product +/- a food product settings,[122, 123] one study was a purchase of six food items,[126] two studies directed participants to complete 'a real household grocery shop' trip[127] or a 'a weekly food purchase for their household' with food and beverage options[129] and lastly, two studies recorded the household food and beverage purchases over four weeks.[124, 130]

Consumption Findings

Only two articles (comprising three studies) were identified which examined the impact of FOPL on food consumption. In one of these, the presence of serving size information on front of pack had no effect on consumption of a product (in this case crackers) framed as 'healthy' but increased intake of a 'less healthy' product compared to no information (study 1).[125] The presence of both serving size and calorie information on the label had a greater effect, further increasing calorie intake. In a subsequent study, a no-FOPL control was compared to a calorie and serving size label and a 'double serving' calorie and serving size label (study 5).[125] Consumption was greater with the standard size label compared to both the 'double serving' label and control. A study comparing a 'reduced salt' FOPL with a no label control and two other label conditions related to taste[131] found no differences between the conditions, regardless of the messaging; intake varied only according to participants' interest in reducing dietary salt.

Purchasing Findings

Meta-analyses are shown in Figures 2-9 and summarised in Table 6 and Table 7. Overall, FOPLs significantly reduced the content of sugar and sodium in purchased products (Table 7 and Figure 11 and Figure 12, respectively) and showed a trend in decreasing energy and saturated fat content (Table 6 and Figure 10 and Figure 13, respectively). When examining the impact of FOPL by specific scheme, meta-analyses showed that the 'High in' schemes significantly reduced purchase content of energy, sugar and sodium (Table 7 and Figure 14, Figure 15, and Figure 17 respectively) and MTL decreased sodium content (Table 7 and Figure 17), with a trend towards reduction in the purchase content of saturated fat (Table 7 and Figure 16). The HSR scheme showed no significant findings, but trended towards a decrease in purchase content of sugar, saturated fat and sodium (Table 7 and Figure 15, Figure 16, and Figure 17, respectively).

For any FOPL vs. no label (Table 6), Egger's regression analysis found no evidence of bias for any, funnel plots showed some evidence of asymmetry, trim and fill showed no evidence of missing studies for any. For FOPL scheme vs no label (Table 7), Egger's regression analysis found no evidence of bias for any, funnel plots showed low evidence of asymmetry, trim and fill analyses showed evidence of no missing studies for the saturated fat and sugar meta-analyses, but evidence of one missing N-S study in energy meta-analysis and two missing HSR studies in the sodium meta-analysis.

Table 6: Effect sizes for nutrient content of food or beverages purchased with FOPL vs. no label, generated by a random effects model

Outcome	**Studies**	**Effect size (95% CI)**	***p*-value**	**I^2 (%)**
Energy (kcals/100g)	6	-2.030 (-4.308, 0.249)	0.081	17.20
Sugar (g/100g)	6	-0.403 (-0.690, -0.116)	0.006	0.0
Saturated fat (g/100g)	4	-0.154 (-0.331, 0.024)	0.091	0.0
Sodium (mg/100g)	4	-24.482 (-43.648, -5.316)	0.012	0.0

CI, confidence interval; I^2, statistic indicating the degree of heterogeneity across studies

Table 7: Effect sizes for nutrient content of food or beverages purchased by specific FOPL scheme vs. no label, generated by a random effects model

Outcome	**Studies**	**Effect size (95% CI)**	***p*-value**	**I^2 (%)**
Energy (kcals/100g)				
'High in'	4	-4.430 (-8.741, -0.119)	0.044	49.40
Health Star Rating	5	-1.381 (-4.388, 1.626)	0.368	27.2
Multiple Traffic Light	5	-4.217 (-11.425, 2.992)	0.252	45.6
Nutri-Score	2	-1.163 (-2.956, 0.629)	0.203	0.0
Daily Intake Guide	1	-1.912 (-30.867, 27.043)	0.897	-
Sugar (g/100g)				
'High in'	4	-0.668 (-1.059, -0.277)	0.001	0
Health Star Rating	5	-0.339 (-0.694, 0.015)	0.061	0
Multiple Traffic Light	5	-0.272 (-0.655, 0.110)	0.162	0
Nutri-Score	2	-0.273 (-0.712, 0.167)	0.224	0
Daily Intake Guide	1	-0.650 (-3.395, 2.095)	0.643	-
Saturated fat (g/100g)				
'High in'	2	-0.209 (-0.475, 0.058)	0.126	0
Health Star Rating	3	-0.211 (-0.435, 0.013)	0.065	0
Multiple Traffic Light	4	-0.207 (-0.426, 0.013)	0.065	0
Nutri-Score	1	0.114 (-0.172, 0.400)	0.436	-
Daily Intake Guide	1	-0.190 (-1.530, 1.150)	0.781	-
Sodium (mg/100g)				
'High in'	2	-33.778 (-59.395, -8.161)	0.010	0.0
Health Star Rating	3	-23.614 (-47.600, 0.373)	0.054	0.0
Multiple Traffic Light	4	-34.938 (-58.525, -11.351)	0.004	0.0
Nutri-Score	1	-8.222 (-34.579, 18.135)	0.541	-
Daily Intake Guide	1	4.500 (-101.266, 110.266)	0.934	-

CI, confidence interval; I^2, statistic indicating the degree of heterogeneity across studies.

Figure 10: Forest plot of comparison: FOPL vs. no label, and energy (kcals/100g) of food or beverages purchased. 95% CIs and study weights are indicated. Effect sizes generated by a random effects model. (CI, confidence interval; WMD, weighted mean difference).

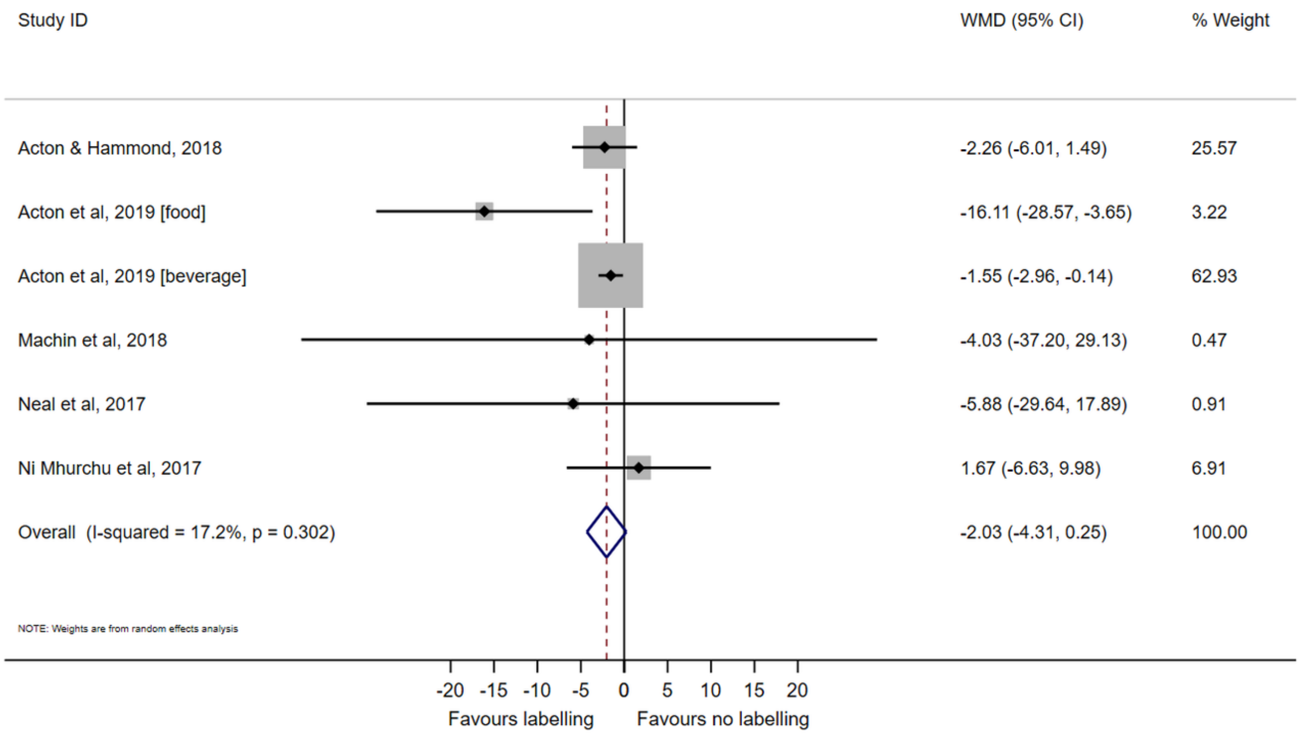


Figure 11: Forest plot of comparison: FOPL vs. no label, and sugar (g/100g), of food or beverages purchased. 95% CIs and study weights are indicated. Effect sizes generated by a random effects model. (CI, confidence interval; WMD, weighted mean difference).

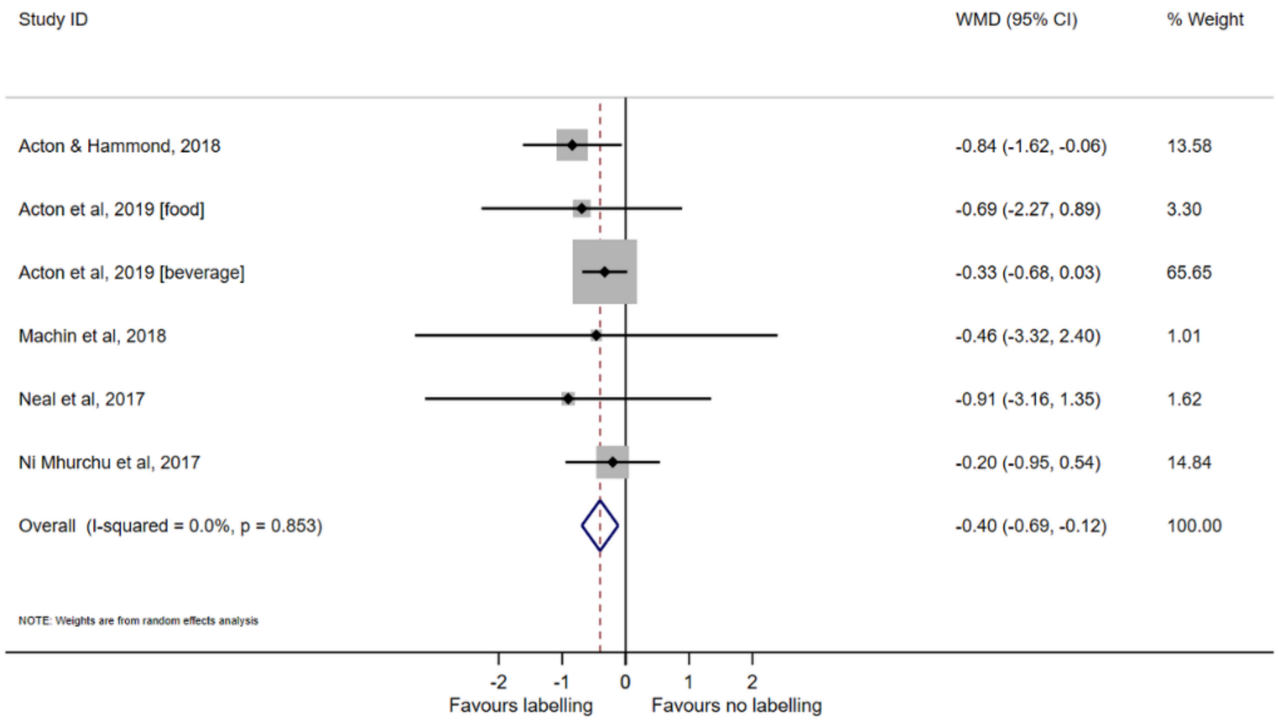


Figure 12: Forest plot of comparison: FOPL vs. no label, and saturated fat (g/100g) of food or beverages purchased. 95% CIs and study weights are indicated. Effect sizes generated by a random effects model. (CI, confidence interval; WMD, weighted mean difference).

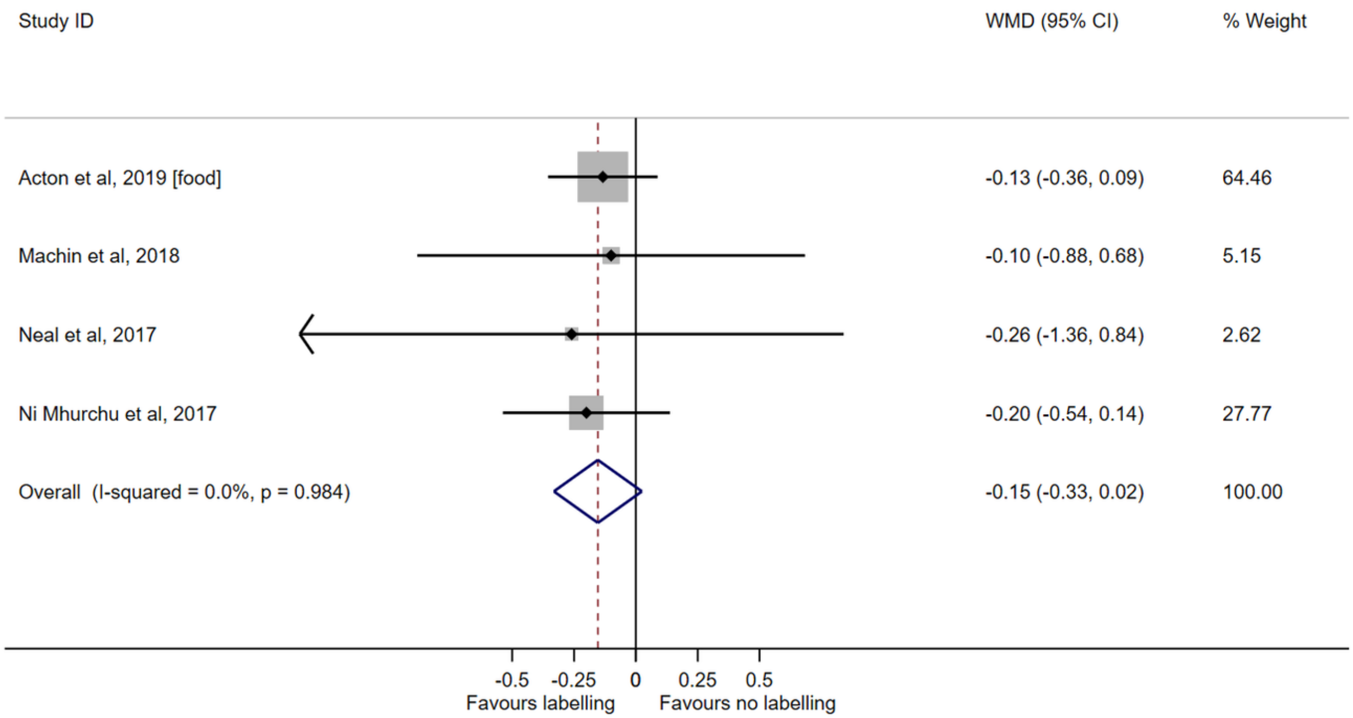


Figure 13: Forest plot of comparison: FOPL vs. no label, and sodium (mg/100g) of food purchased. 95% CIs and study weights are indicated. Effect sizes generated by a random effects model. (CI, confidence interval; WMD, weighted mean difference).

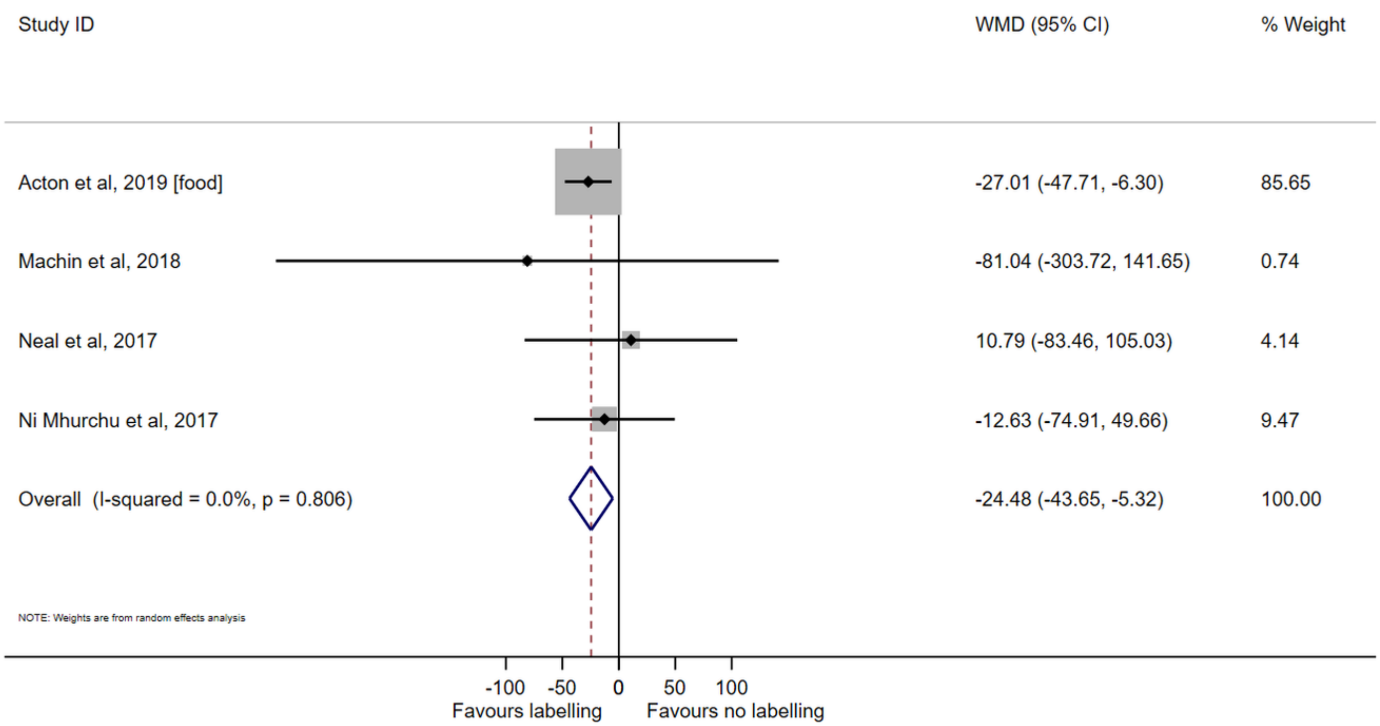


Figure 14: Forest plot of comparison: by FOPL scheme vs. no label, and energy (kcals/100g) of food or beverages purchased. 95% CIs and study weights are indicated. Effect sizes generated by a random effects model. (CI, confidence interval; WMD, weighted mean difference).

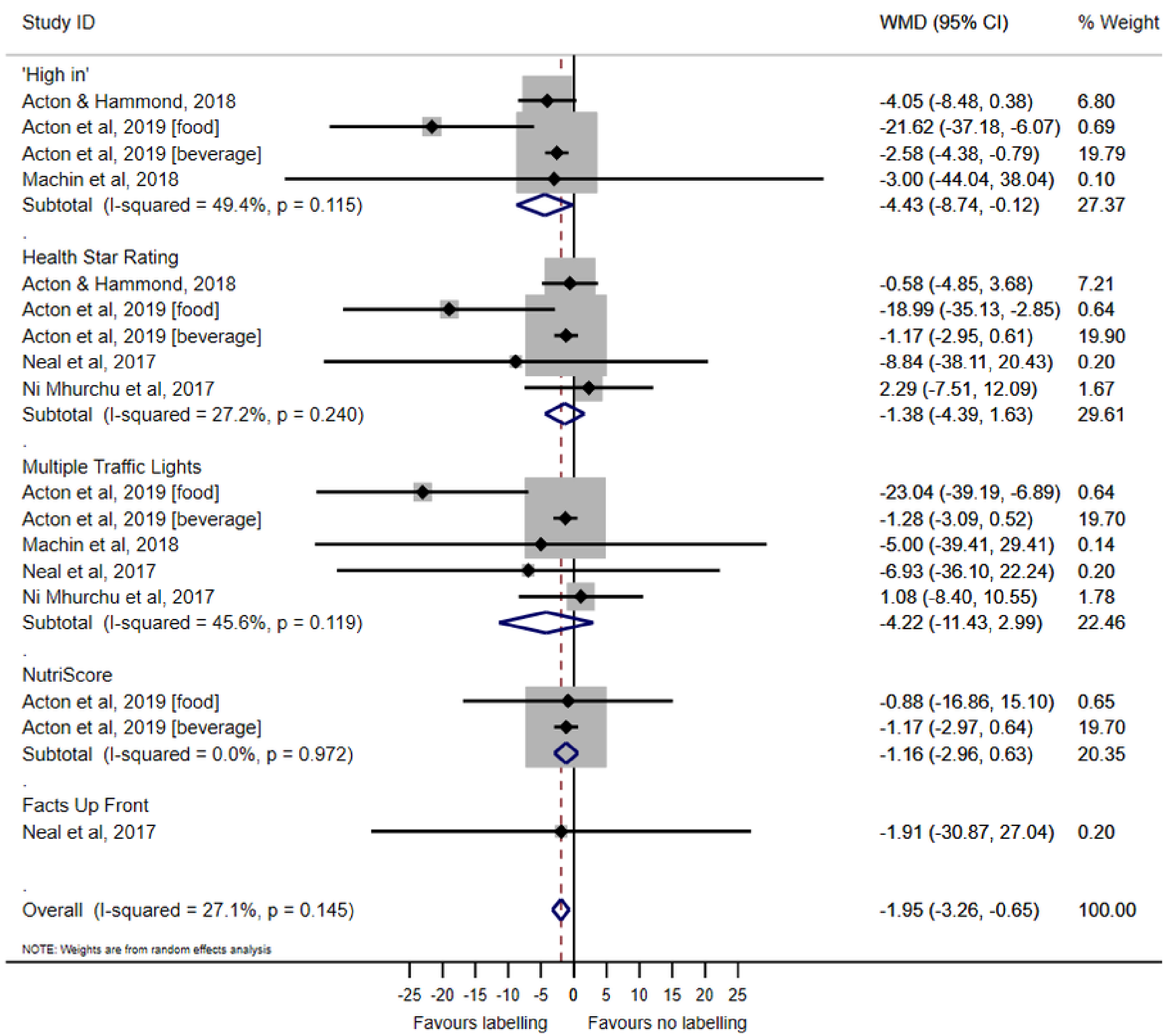


Figure 15: Forest plot of comparison: by FOPL scheme vs. no label, and sugar (g/100g) of food or beverages purchased. 95% CIs and study weights are indicated. Effect sizes generated by a random effects model. (CI, confidence interval; WMD, weighted mean difference).

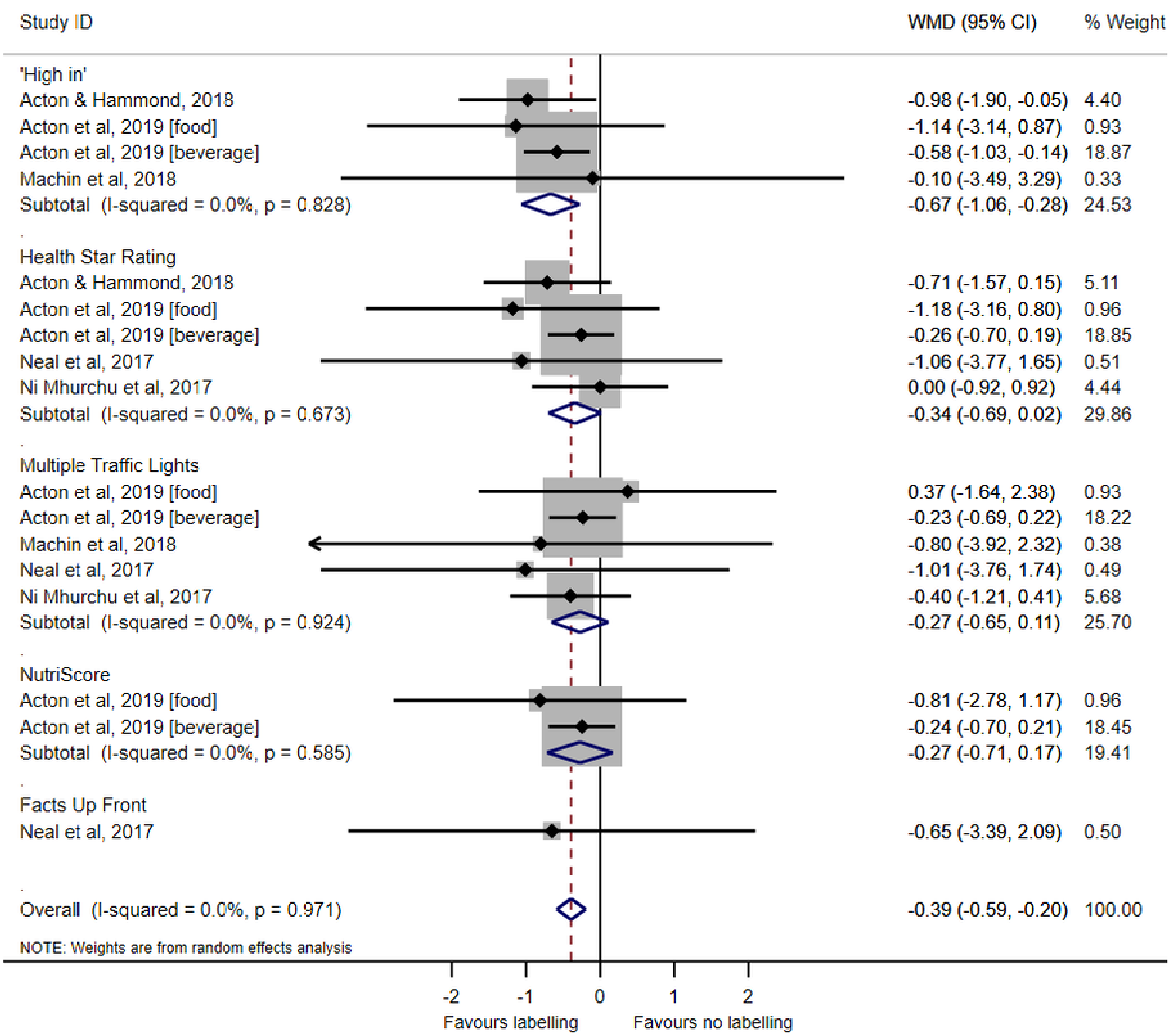


Figure 16: Forest plot of comparison: by FOPL scheme vs. no label, and saturated fat (g/100g) of food or beverages purchased. 95% CIs and study weights are indicated. Effect sizes generated by a random effects model. (CI, confidence interval; WMD, weighted mean difference).

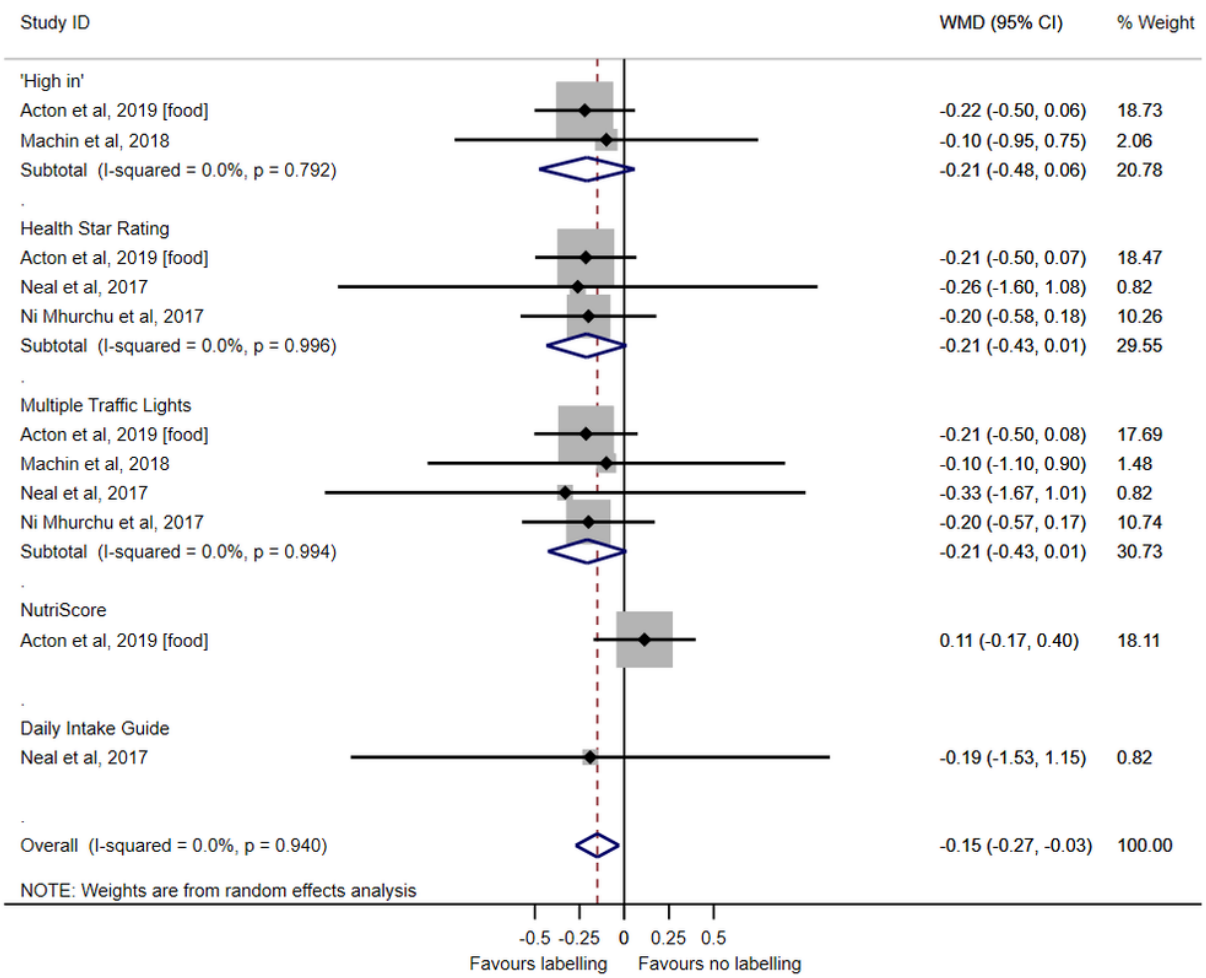
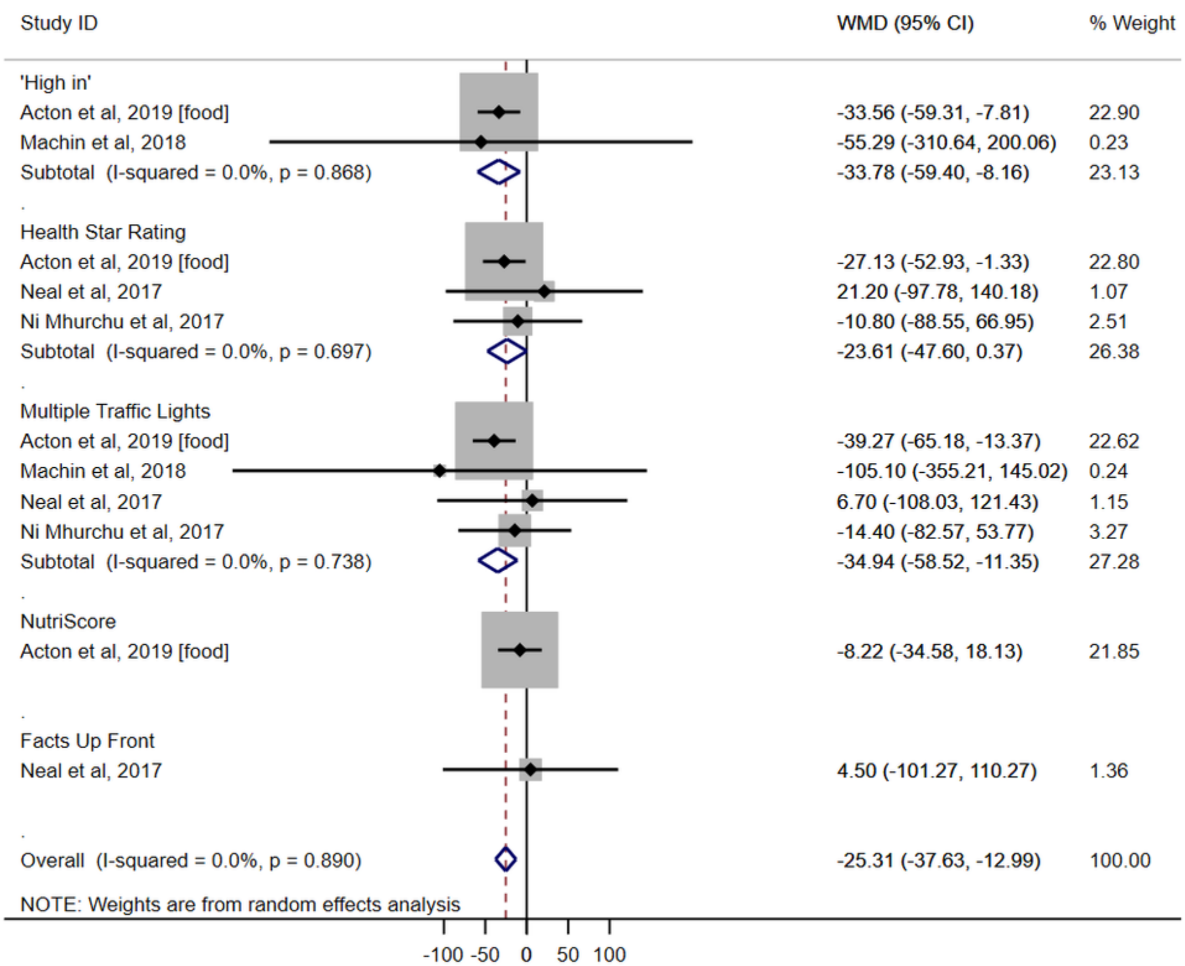


Figure 17: Forest plot of comparison: by FOPL scheme vs. no label, and sodium (mg/100g) of food purchased. 95% CIs and study weights are indicated. Effect sizes generated by a random effects model. (CI, confidence interval; WMD, weighted mean difference).



The findings from the three purchasing outcome studies not included in the meta-analyses were broadly consistent with the meta-analysis results.[126-128] Two of the studies found the quality of purchased products improved nutritionally with most FOPLs compared to control groups, with significant effects for NutriColors (equivalent to MTL), NutriMark (equivalent to HSR), NutriRepere (equivalent to DIG) and N-S [128] and a text-based health warning, but not for the WL for 'high in sugar'.[127] One study found no differences in the energy, sugar, saturated fat or sodium of selected products to purchase with either MTL or DIG FOPL, compared to no-FOPL.[126] One study also measured fibre, protein and a combined point score for fruits, vegetables, nuts, and legumes, with no significant differences between FOPL or no-FOPL.[130]

Few studies reported the effect of FOPL according to socio-demographic or other characteristics. One study found a less pronounced (but still significant) reduction in purchasing of unhealthy foods in those of lower socio-economic status (SES), with N-S performing best in this group.[128] There was no observed effect of BMI, education or household income on purchasing in another study.[126] One study found no impact of age, ethnicity, education but the back of pack nutrition information panel (control) appeared to perform better than the HSR and MTL labels in low income groups and men, but the authors commented that the numbers in these sub-groups were small.[130]

Interrupted Time Series (ITS) studies

Study Descriptions

Three ITS studies were identified, conducted in Chile,[132] the UK[133] and Ecuador.[134] All examined the impact of introducing new labelling schemes on product sales in real-world settings by measuring purchasing behaviours before and after the introduction of the schemes. The labelling schemes were the mandatory Chilean WL,[132] the voluntary UK FOP Guideline Daily Amounts[133] and the mandatory TL labelling in Ecuador.[134] Two of the studies used customer scanner data from major grocery retail stores[132, 133] and one used Kantar data from a random sample of households.[134] Outcome measures were the quantity of products purchased and/ or the energy content of the purchased products. The products varied between studies, two studies included foods and drinks (fruit juice, breakfast cereal, chocolate and cookies;[132] biscuits, breakfast cereals and soft drinks[133] and the other study focused on soft drinks.[134]

Findings

All of the studies found purchasing of unhealthy products decreased for at least some of the included products. One study found a significant reduction in purchases of juices (-23.8%) and cereals (-11.0%) following the introduction of the Chilean WL, but no impact on chocolates and candies (11.2%), or cookies (1.7%).[132] The UK study found that customers purchased products with fewer calories after the introduction of GDA, with a 9.5% decrease in calories across the three included product categories (cookies, breakfast cereals, soft drinks).[133] The study from Ecuador found that, following introduction of a mandatory TL scheme, the purchase of soft drinks reduced by 0.003L and the mean sugar content of soft drinks decreased by 0.93g/ 100ml, the latter due to reformulation.[134]

Bias Assessments

For the experimental studies, those examining consumption outcomes were rated as having 'some concerns' to 'high' risk of bias and those measuring purchasing outcomes were rated as having 'some concerns' or 'low' risk for one study. The risk of bias for the ITS studies were all assessed as low risk and rated as good quality. The bias assessments are provided in Appendix B.

Chapter 3 Discussion

This systematic review set out to identify studies published since the 2018 Cochrane review by adapting the review with more inclusive purchasing outcomes and focusing on the impact of FOPL on objectively measured purchasing and consumption of pre-packaged foods. It led to 11 experimental studies (eight of which reported purchasing outcomes and three reported consumption outcomes) and three ITS studies meeting the inclusion criteria. The meta-analyses were informative about the impact of FOPLs on the sugar, energy, saturated fat and sodium content of food purchases, thus further developing the evidence base. There was a significant overall effect of any FOPL compared to no FOPL for the sugar and sodium content of purchases, and a trend for energy and saturated fat content. The ‘high in’ FOPL significantly reduced the sugar, calorie and sodium content of purchased products compared to no FOPL and the MTL FOPL significantly reduced the sodium content of purchased products compared to no FOPL. There were no effects on purchasing from N-S, HSR or DIG, although the HSR FOPL approached significance for sugar, saturated fat and sodium. It should be noted that few studies were identified which examined N-S or DIG. Data on consumption were limited and findings inconsistent. The three ITS studies indicated that labelling schemes (WL, GDA, TL) resulted in healthier purchasing patterns.

Taken together, these findings suggest that overall, FOPLs had an impact on consumer behaviour but this was stronger for purchasing compared to consumption, and there appears to be less evidence to support ‘straight up’ nutrition information (including GDA/DIG). Non-interpretive or 'straight up' nutrition information is likely to be more difficult for people to understand and act upon, as suggested in the experimental study across 12 studies that found N-S performed best.[69] This is further supported in the literature, with evidence that FOPLs that use interpretive elements (colour or text) were easier to interpret than just numeric information.[43] This could be because the inclusion of information giving an indication of the healthiness of the food, rather than solely providing nutritional information that requires interpretation, is helpful to people with a poorer understanding of nutrition. If people do have adequate knowledge/nutritional understanding, this still requires processing time, which may be unrealistic in a shopping scenario, which is typically time limited. Studies have indicated that nutrition knowledge is associated with level of education[135] and SES[136] meaning that individuals from less educated and poorer background are likely to find interpreting these 'straight up' FOPLs more challenging. This is particularly important given the marked inequalities in obesity.[137, 138] Few studies assessed the impact of SES on the results. Of the two studies that did there was no evidence of effect, but an indication that N-S was most effective in impacting the purchases of the lower SES group. This suggests that, even if to a lesser extent, individuals of lower SES may be able to take advantage of the information in FOPL to make healthier purchasing decisions. Another study, which used real purchasing data from the UK between 2005-2008, found that the presence of FOPLs (a mix of MTL, GDA and hybrid) reduced the purchasing of 'unhealthy' products to a greater extent in households of lower SES compared to higher SES.[139] The authors hypothesised that this was due to the FOPLs presenting new information to the lower SES group, due to the assumed lower baseline nutritional knowledge.

The three ITS studies identified in the current review showed healthier purchasing following the introduction of the Chilean WL, TL labelling in Ecuador and the UK FOP GDA. These studies were rated as being high quality with no major concerns which gives confidence in the findings. An additional study, carried out in the Netherlands using household purchasing data, but published after the cut-off point for inclusion, also found favourable effects of labelling for most products.[140] In this study, products displaying the voluntary Dutch Choices endorsement FOPL (which indicates that a product is 'healthy') experienced significant increases in market share after implementation of the scheme. Another recent study examined national household purchasing data of beverages before and after the mandatory WL in Chile.[141] They found that following the policy implementation, purchases of beverages high in energy, sugar, sodium, and saturated fat content decreased by 23.7% or by 22.8mL per capita per day. This supports the meta-analyses findings, that 'high in' systems are effective at reducing consumption of energy, sugar and sodium.

I was unable to quantify the effect of FOPL according to product, but there was some evidence that FOPLs may have an impact when they provide unexpected information. One of the ITS studies found significant reductions in purchasing of juice and cereals but not chocolates, candies and cookies; the authors hypothesised that information disclosure may only be effective at reducing purchasing when the information is unexpected, i.e., juice and cereals are not typically viewed as 'unhealthy'.[132] However, this is contrary to a more recent study which found changes in purchasing for most products but no change for cereals, a possible explanation being that consumers may use FOPLs less for products perceived as healthy.[140] A study looking at changes in knowledge with FOPLs found smaller increases in knowledge for cereals compared to cakes.[69] These findings require further exploration to establish how FOPL schemes impact on purchasing patterns of different products, which will be explored in Chapters 5 and 6.

Limitations

The limitations of this review include the small number of included studies that examined DIG and N-S (due to this being a newly developed FOPL scheme). There was also limited data comparing individual FOPL schemes meaning meta-analysis was not possible. Care needs to be taken when interpreting the meta-analysis results as the studies differed greatly, as evidenced by the high heterogeneity between 'high in' and MTL for energy (kcals/100g). The outcome of purchasing is insightful and is viewed as a valid proxy measure for food consumption.[142] It is also not possible to understand consumption at an individual level as many outcomes were measured at the household level. Purchases that were experimental or hypothetical in nature (not using their own money) were also included, so care with interpretation is needed, although all studies strongly replicated actual retail environments. The risk of bias for the experimental studies was moderate to high for the consumption studies and mostly moderate for the purchasing studies. The study quality for the ITS studies was high overall. Study quality overall was reasonable albeit mixed. Owing to the approach in adopting a more inclusive eligibility criteria than the Cochrane review, and the narrow timeframe of the searches, there are possibly earlier studies that exist that were not considered. This review provides a good platform for further work, including updated meta-analyses as more studies are published or extending outcomes to examine the effects of FOPL on purchase intentions. It also highlights the need for real-world evaluation studies of any FOPL policies.

Chapter 3 Conclusions

This systematic review has established that FOPLs can be potentially effective at improving consumer behaviours, through the purchasing of healthier products, and the evidence suggests interpretive labels are the most effective. This has built on the literature review presented in Chapter 2 that showed that FOPLs can improve the ability of consumers to correctly understand product healthiness/nutritional quality, but no research has been conducted in a representative British sample. This systematic review then addressed the next stage in the pathway presented in the Crockett et al. logic model, of whether FOPLs can impact on the purchasing of healthier food and a suggested effect on consumption. This review included international evidence and was not able to assess which FOPL scheme was most effective due to the different comparisons across the studies. Therefore, it is still unclear which FOPL option is the most effective at impacting consumer behaviours. Throughout the rest of the thesis, I detail an experiment undertaken to test and better understand FOPL options that are of interest to UK policymakers. In order to fill this evidence gap I co-designed an experiment, in collaboration with the DHSC and NatCen, which was set up as the best way of generating quality evidence. A high quality experiment was needed to appropriately address the policy question. Using the synthesis of the results from Chapter 2 and 3 I co-developed an online randomised controlled experiment that I will outline in Chapter 4, including the design, rationale and methodology used to test the FOPL schemes. In Chapter 5, I address the objective understanding of FOPLs in terms of understanding product healthiness between different FOPL schemes, compared to no label. In Chapter 6, I address other factors, including subjective understanding (if participants *think* they understand the label), time taken to use the labels and perceptions of the labels and ranking ability. In Chapter 7, I present the results of a latent class analysis, examining the impact that consumer food habits and motivational characteristics may have on FOPL engagement and by extension consumer behaviours.

Chapter 4: Experiment Design and Methodology

Chapter 4 Background

In this Chapter, I will explain the design of the study and the rationale, including decisions around the food and drink products to include and the decision to use mock products. I outline information on the steering group, study participants, setting, materials, ranking task, survey development, measures, procedures, statistical analysis and ethical approval. The results of the experiment and related research are presented in Chapter 5 (primary outcomes), Chapter 6 (secondary outcomes), Chapter 7 (latent class analysis using experimental findings and sample) and Chapter 8 (a Patient and Public Involvement and Engagement session relating to the findings).

Steering Group

This study was commissioned by the Department of Health and Social Care (DHSC) to explore which FOPLs would help consumers in understanding the healthiness of products, to assist them with making healthier choices when buying food and drink products. The research project involved close collaboration with DHSC and Public Health England (PHE) to ensure the delivery of a study that would meet their specifications and be most useful to inform policy makers thinking. A steering group was created to assist this collaboration, comprising of both international experts and officials from UK public health departments: DHSC – Officials from the Science, Research and Evidence Directorate, Nutrition and Healthy Weight Branch and Population Health Analysis Branch; PHE – Officials from the Diet, Obesity and Physical Activity Branch; Food Standards Agency in Northern Ireland; Food Standards Scotland; and the Welsh Government. The group deliberated and advised upon which labels to use, label design, product selection and design, as further explained in this Chapter.

Participants

The NatCen panel, a probability-based research panel, was chosen as the participants are nationally representative of Great Britain.[143] The panel was established in 2015, when participants from the British Social Attitudes Survey were invited to join via email, texts or letters. It was essential that the experiment sample was nationally representative, ensuring the external validity of the study as it will be used to inform national policy decisions. The NatCen sample is broadly representative of the adult British population[144] and weights are calculated to assist with the adjustment of the sample based on non-response bias (see Appendix C for the NatCen Panel Technical Information and the reference for greater detail on this process). Using the NatCen panel for this experiment has many strengths. The cohort is established (allowing for tight policy timeframes to be met), the participants of NatCen are experienced in taking surveys with high participation response rates (important for appropriately powered sample sizes). However, using this panel does include some weaknesses, which will be addressed in later Chapters (limitations section of Chapter 5, 6 and 9).

Setting

The study took place entirely online, due to the requirement for visual stimulus in the survey.

Experiment Design

The experiment was a five arm randomised controlled trial using a between-subjects experimental survey design. The main trial aims were to identify if specific front of pack labels (Multiple Traffic Light, MTL; Nutri-Score, N-S; Warning Label, WL; Positive Choice Tick, PC) are effective at improving participants' ability to determine the healthiness of food items, and if there is a difference compared to the no label control group.

Materials

Label Selection and Design

Four labels were selected for testing by DHSC including MTL, N-S, WL and PC. MTL, N-S, and WL are established labels with published specifications. Copyright permission and specification documents were sought and received from the appropriate institutions: Public Health France (Santé Publique France) for N-S[145] and the Ministry of Health Chile (Ministerio de Salud) for WL.[146] MTL does not require copyright permission but the graphic designer created the MTL according to the published specifications.[83] The PC label was created for this study based on other endorsement logos (e.g., Healthy Choices Programme), see Figure 18.

The established criteria were used to assign FOPLs based on each product's nutritional profile (see Table 8 for threshold details, section 'Nutritional profiles of products' below and Appendix

C for further detail on the N-S and MTL label specifications). The MTL and N-S label thresholds were taken directly from the relevant technical guidance. The WL criteria were adapted from the technical guidance, with warnings assigned to products with a red MTL label (specific to each nutrient) and an energy warning displayed if energy content was greater than 275kcal/100g for food products or 70kcal/100ml for drinks (these were the Chilean thresholds specified at 36 months, except no WL was assigned to products that crossed energy thresholds but did not contain added sugars or saturated fat). The PC label was assigned where the product did not qualify for any red MTL labels and did not cross the energy threshold used for assigning the Chilean energy WL. The three products within each category were designed to have distinct label profiles, except this was not possible for the PC label. Due to the PC label being a binary label, it was unfeasible to create a scenario where three products could be differentiated between. It was also impossible to create products within the cake and crisps category that were a) realistic and b) would qualify for a PC, while still meeting the requirements for distinct label profiles for the other FOPLs.

Table 8: Thresholds and agreed approaches for each FOPL

FOPL scheme	Approach
MTL	Existing UK voluntary scheme in line with DHSC FoP guidance.* Fat (per 100g): low/green ≤ 3g; medium/amber >3g to ≤ 17.5g; high/red >17.5g Saturates (per 100g): low/green ≤ 1.5g; medium/amber >1.5g to ≤ 5g; high/red >5g Total sugars (per 100g): low/green ≤ 5g; medium/amber > 5g to ≤22.5g; high/red > 22.5g Salt (per 100g): low/green ≤ 0.3g; medium/amber >0.3g to ≤ 1.5g; high/red >1.5g *There is additional criteria per portion
N-S	N-S calculated for each product in line with guidance.* Favourable component (fruits, vegetables, nuts, protein, dietary fibre) and unfavourable components (sugar, salt, saturates, energy) are attributed nutritional values/calculated as points and the overall score translates to a Nutri-Score value. Points and corresponding Nutri-Score value for food products per 100g: min to -1 = A/dark green; 0 to 2 = B/light green; 3 to 10 = C/yellow; 11 to 18 = D/light orange; 19 to max = E/dark orange. Points and corresponding Nutri-Score value for drink products per 100ml: water only = A /dark green; min to - 1 = B/light green; 2 to 5 = C/yellow; 6 to 9 = D/light orange; 10 to max = E/dark orange.
WL	Use of UK MTL 'high/red' for total fat, saturated fat, total sugars and salt. Use of Chile energy cut-off at ≥275kcal/100g.
PC	Use of UK MTL 'not high/red' for total fat, saturated fat, total sugars and salt and product is also <275kcal/100g.

For further details see *Appendix C.

Product Category Selection

The selection of product categories was crucial for the success of the experiment. The steering group agreed that the product categories needed to be commonly purchased and consumed in the UK and have enough nutritional variability to allow for three distinct and varied products (a 'most healthy', an 'in between' and a 'least healthy'). To select the product categories, the 2004/5 UK Nutrient Profiling Model[147] was utilised, with references to the distribution of scores, to ensure that it was feasible to have three distinct products per category. Following collaboration with DHSC and PHE, six categories were selected: pizza, instant hot chocolate, cake, crisps, yoghurt and breakfast cereal. These categories each had enough variability to create realistic products for each healthiness level, for example pizza products, the 'most healthy' had a thin crust, the 'in between' was standard and 'least healthy' had extra cheese. Examples of products that were not chosen due to lacking nutritional variability include energy drinks, chocolate mousse and muesli bars. Within product categories, similar product types were selected. For example, for the breakfast cereal product category, only muesli-type products were selected rather than selecting a broader mix of cereal types such as flavoured or plain cereals (e.g., chocolate-flavoured cereals or porridge).

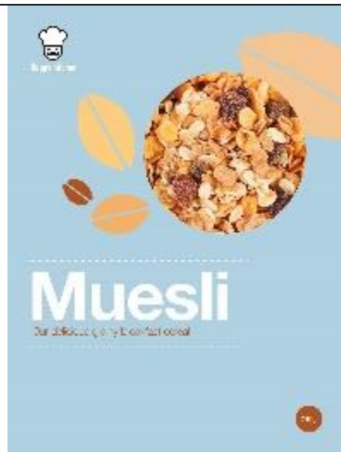
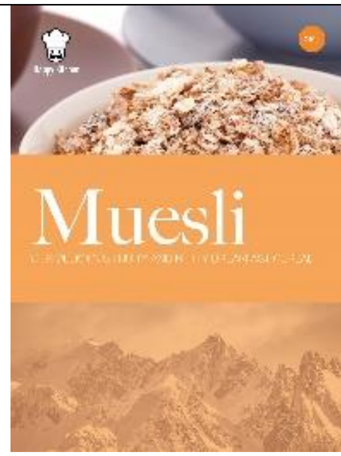
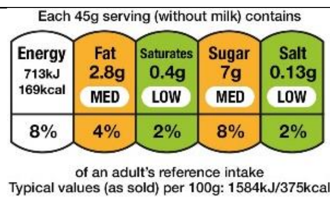
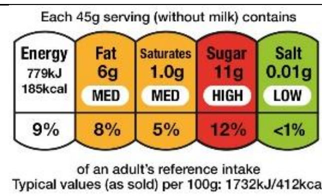
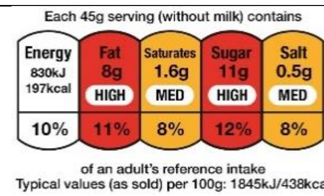
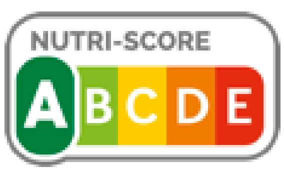
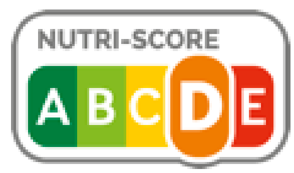
Nutritional Profiles of Products

The nutritional profiles for each of the fictional products were created by me and approved by the research team which included a dietitian, as this information was required to inform the FOPL assignments. The nutrient information required for this included energy (kilocalories/kilojoules), fat, saturated fat, sugar, salt, protein and fibre; per 100g and per portion (see Appendix C, Table C1 for full details of the products). The nutritional profiles and corresponding FOPL assignments were intentionally designed to reduce ambiguity between products and to ensure participants with some nutritional knowledge would be able to successfully differentiate between the products and theoretically rank them correctly. The aim was to have distinct label profiles for the three products within each category (e.g., for N-S, scores of 'A', 'C', 'D'; for MTL, different colours patterns, i.e., no scenario with two products having two orange and two red of different nutrients) although this was not possible for the PC label, as outlined in the section 'Label section and design'.

Product Packaging Design

Mock packaging for the 18 fictional products (six product categories x three products) were designed in collaboration with a UCL graphic designer. This decision was made in collaboration with the steering group in order to avoid copyright issues, minimise any biases from real brand or product familiarity and to allow the standardisation of product characteristics. All products had the same fictional brand name and logo (Happy Kitchen) as well as identical product size and serving sizes within each product category. Additionally, no nutrition, allergen or quality claims were included on the packaging as these could influence participants' responses (e.g., high in protein, organic, gluten-free) (see Figure 18 for images of the cereal products and Appendix C, Table C2 for all images). The packaging designs were created to be realistic and contain cues about the underlying nutritional quality. These cues would enable participants with nutritional knowledge to correctly distinguish between the ‘most healthy’, ‘in between’ and ‘least healthy’ products within each category. The cues included product descriptions (e.g., “thick strawberry layer” to indicate a less healthy yoghurt), product characteristics (e.g., granola with visible nuts, cake with thick chocolate icing) and colour (e.g., lighter colours used for healthier products and darker for least healthy products), see Table C2 for all product packaging images, applicable labels and design features. The position and size of the FOPLs on the packaging was also standardised across the products to be placed in the bottom left (see Appendix C for product packaging for the ‘least healthy’ pizza with MTL, N-S and WL FOPLs). This was a compromise between the Chilean WL specifications which suggests the upper right corner position[146] and MTL guidance which says it just needs to be provided in the ‘principal field of vision’ but the position can be determined by other factors (brand position, pack size and shaped etc.).[83]

Figure 18: Example of the breakfast cereal product packaging used for the ranking tasks, with the associated front of pack labels, by experimental group

Group	Breakfast cereal		
No label (baseline and control)			
Multiple Traffic Light			
Nutri-Score			
Warning label			
Positive Choice tick			

Ranking Task

The ranking task was adapted from a recently conducted high-quality experimental study.[70] In the experimental study by Egnell et al.,[70] the research question and aims were similar, but they did not include all FOPLs of interest to the UK government, the study was not conducted in a nationally representative UK sample and they did not include a no-label control with repeated measures to control for learning. Five FOPL conditions were tested (HSR, MTL, N-S, RI and WL) and participants were presented with three products from a single product category of varying healthiness and then asked to rank the products according to healthiness. Baseline rankings with no labels were completed before follow-up rankings with FOPLs shown. It was decided that the experiment would use a similar ranking task but a no-label control group should be added to strengthen the reliability of the findings (see section 'Primary outcome measure' for further details). This inclusion was important to confirm that any differences in ranking ability were due to the inclusion of the FOPLs in the second set of ranking tasks, rather than inherent learning from repeating the ranking task.

Survey Development

Question Development

In addition to the ranking task, a survey was developed to test factors beyond the ability to correctly identify the healthiness of products, which are important for actual label use. These factors were primarily informed by the Grunert and Wills conceptual framework (as outlined in Chapter 1) and included; subjective understanding, perception, liking and use.[37] See section 'Measures' for specific details and Appendix C for the full survey. Following expert guidance from NatCen regarding survey design, agree-disagree statements were avoided due to acquiescence bias (where some respondents are likely to acquiesce/agree to the statement) and questions were designed to be construct-specific rather than leading. For example, instead of 'did you find the labels easy to understand' or 'do you think that these labels would take too long to use when shopping', questions were instead phrased 'how easy or difficult did you find it to understand these labels' and 'which of the following best represents your views on how long it takes to use these labels'. The questions that could be included by NatCen were limited due to time constraints, so the inclusion process was very selective.

User Testing

To ensure the questions were perceived how we intended them, I conducted some informal pilot surveys with colleagues and acquaintances. This included a mock version of the survey created using PowerPoint. I conducted 14 of these pilot sessions and the results informed the wording of the questions, identified redundancies in the questions, provided insight into the user experience of the ranking task and the overall time needed to complete the survey. Once the survey questions and design were finalised, NatCen also conducted some user testing. This was conducted on their survey platform with members of NatCen and our research team. These tests were aimed at identifying any errors within the survey and testing the functionality.

Measures

Primary Outcome Measure

<u>Ranking task</u>

The ability to accurately assess product healthiness was measured using the ranking tasks. This first consisted of baseline rankings with plain packaging for all of the products with no FOPLs displayed. Followed by the follow-up rankings with the same product packaging now displaying the assigned FOPLs. The ranking was classified as correct if all three products within each category were ranked in the correct order. Conversely, the ranking was incorrect if any were ranked in the wrong place. The primary outcome measure was examined in three ways and the results are presented in Chapter 5:

Ranking of product healthiness at baseline and follow-up, by product category. Coded as all products correct = 1, any incorrect = 0.

Change in ranking from baseline to follow-up, by product category. The change in ranking was classified as an improvement (incorrect at baseline and correct at follow-up), no change (incorrect OR correct at both baseline and follow-up) or worsened (correct at baseline and incorrect at follow-up). Coded as improved = +1; no change = 0; worsened = −1.

Change in global food score. This score was a summed overall change score, calculated using the change in ranking outcome for each of the five food categories i.e., excluding the drink (see section ‘Chapter 5 Statistical Analysis’ for justification). The overall change scores ranged from -5, if ranking ability in all five food categories, to +5, if ranking ability improved in all five food categories.

Secondary Outcome Measures

All of the secondary outcome measures are presented in Chapter 6.

Ranking task perceptions

Confidence in ranking was assessed by asking participants how many products they thought they had ranked correctly after the baseline rankings and once again after the follow-up rankings. Confidence was dichotomised as ‘more confident’ (all/ most) vs ‘less confident’ (some/ none/ don’t know) for ease of interpretation.

Perceiving enough information was assessed by asking participants if they felt they had enough information to rank the product from most healthy to least healthy. For enough information to rank products there were 12 responses, recoded as ‘yes to all’, ‘no to all’ or ‘mixed response’ at baseline and follow-up, for ease of interpretation.

Perceptions of labels

To examine the perceptions of label use, liking and subjective understanding, participants randomised to a FOPL group (i.e., not the control group) were asked the following questions:

- Recall of seeing the label in the ranking tasks (dichotomised as yes vs no/not sure for ease of interpretation)
- Use of the label in the ranking tasks (all/ some/ did not use)
- Ease of label understanding (dichotomised as ‘easy’ [very easy/ quite easy] vs ‘difficult’ [quite difficult/ very difficult])
- View on implementation of label on products in the UK (yes-all/ yes-some/ no-none)
- View on helpfulness of label for food shopping (dichotomised as ‘helpful’ [very helpful / quite helpful] vs ‘not helpful’ [not very helpful/ not at all helpful])
- Views on time to use label when shopping (quick enough/ too long)

Time taken to complete ranking tasks. This was a proxy measure, i.e., the time spent on that page of the survey extracted from the software paradata (administrative data collected) in seconds. The median and interquartile range of time taken in seconds to complete each ranking task was calculated per product category, at baseline and follow-up and summed time for baseline and follow-up rankings.

Other Measures

Current Food Habits

- Food shopping responsibility (yes – I do/ no – someone else does)
- Product consumption or purchase in the last 12 months (selected all that applied across the 6 product categories – pizza, hot chocolate, cake, crisps, yoghurt, breakfast cereal or none)
- Front of pack nutritional information use when buying food (very often/ quite often/ occasionally/ rarely/ never)
- Front of pack nutritional information influence on buying food (very often/ quite often/ occasionally/ rarely/ - never)
- Knowledge of healthy eating (a lot of knowledge/ some knowledge/ a little knowledge/ no knowledge)
- Interest in healthy eating (very interested/ quite interested/ not very interested/ not at all interested)
- Currently trying to lose weight (yes/ no/ prefer not to say)

Medical/ Additional demographics

- Height and weight – in metres and centimetres/ feet and inches or a prefer not to say option
    - Body Mass Index (BMI) was derived from these self-reported measures
- Physical or mental health conditions or conditions or illnesses lasting or expected to last for 12 months or more which affect vision/ learning, understanding or concentration/ eating

- Medical diagnosis of: Coronary heart disease/Angina/Heart attack or myocardial infarction/Type 1 Diabetes /Type 2 Diabetes/High blood pressure/hypertension /None of these
- Current pregnancy status

Fed-forward demographic data provided by NatCen

Including but not limited to categories listed here, for full details, see Appendix C for NatCen Panel Technical Information. Key information is updated every six months.

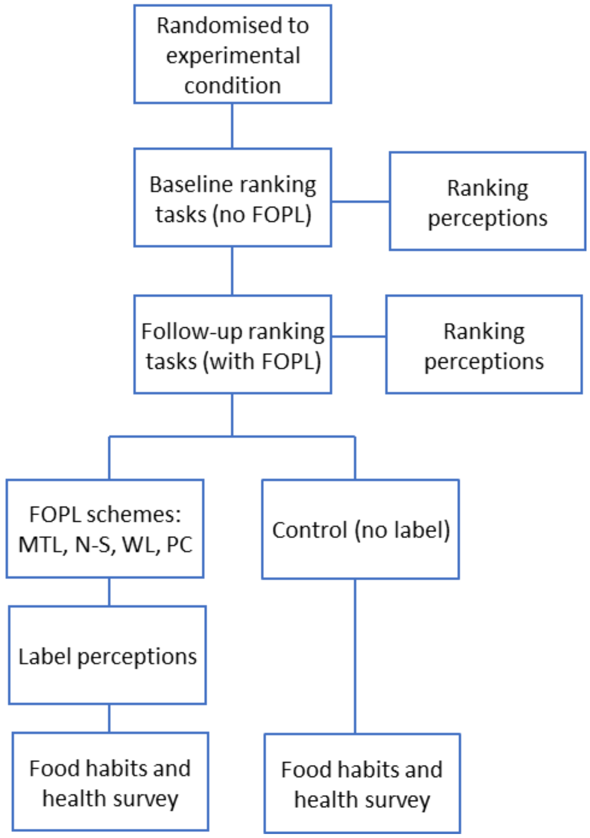
- Sex (female/ male)
- Age category (18–29/ 30–39/ 40–49/ 50–59/ 60–69/ 70+)
- Ethnicity (white British/ white other, mixed or multiple ethnic groups, Asian or Asian British, black or black British, other);
- Highest educational qualification achieved (degree or equivalent/ A levels or vocational level 3 or equivalent/ qualifications below A levels or equivalent/ other qualification/ no qualification)
- Equivalised household income (£800 or less/ £801-1250/ £1251-2000/ More than £2000)
- Children in the household under 16 years of age (yes vs no)

Survey Procedure

The survey procedure and data collection was conducted by NatCen. All active NatCen panel members were contacted by letter, email and SMS, with an invitation and link to participate in the online experiment. As per NatCen standard procedures, the invitation contained information about the study, outlined that participation was voluntary, how the data would be securely held and details of the incentive offered as a thank you on completion (a £5 Love2Shop voucher). Opting in to the study was accepted as informed consent, as per standard NatCen procedures and approved by the ethics committee. The survey was created with NatCen’s established survey template using UNICOM Intelligence software. Participants could complete the survey on any computer device, including smartphones. For accessibility, the survey was optimised and tested by NatCen across a range of browsers, operating systems and devices (including laptops, desktops, tablets and smartphones). However, participants were encouraged in the invitation email to use a device with a larger screen and if phone use was detected by the software during the survey, participants were prompted to switch to a device with a larger screen. If the participant did not change the device, the images were shown vertically instead of horizontally, to improve the readability. Participants were able to leave and re-join the survey and email reminders to complete the survey were sent. Participants were recruited and data collected between the 28th of October to the 15th of November 2020. See Figure 19 for the experimental procedure.

NatCen used SPSS to randomly allocate equal numbers of participants to one of five experimental groups: MTL, N-S, WL, PC or control (no-label). Randomisation was based on key stratification variables known before the study: recruitment to panel, sex, age, government office region and household income. Participants were told that they would be asked their views on food labelling on behalf of the UK DHSC. To provide an indicator of the time taken to use labels, participants were instructed to click 'next' as soon as they had completed each task. Participants were asked to complete 12 ranking tasks where each involved viewing three images of food/drink products within a category and ranking them in order of healthiness: 'most healthy', 'in between' and 'least healthy'. The order of categories and of products within categories were randomised to avoid presentation bias. A repeated-measures design was used; all participants completed the same 12 ranking tasks, six with no FOPLs (baseline) and six with packaging displaying the FOPL according to their assigned experimental group (follow-up). Participants could enlarge the images, but no other information (such as the mandatory back of pack nutritional information) was provided. To minimise missing data, participants were unable to rank two products as the same level of healthiness and an error message would appear if they attempted to. They were also only given a "don't know" option if they tried to move onto the next page without completing the ranking task.

Figure 19: Experimental procedure. MTL = Multiple Traffic Light; N-S = Nutri-Score; WL= Warning Label; PC = Positive Choice Tick



Following completion of each ranking task, participants were asked if they had enough information to complete the ranking (12 times) and their confidence in their ability to correctly rank products (twice: at baseline and follow-up). Participants then answered questions specific to the FOPL group they were randomised to (i.e., control group did not answer these questions), including questions on perceptions of labels with a picture on their allocated labels shown. Following that, all participants answered current food habits; medical/ additional demographics (as outlined in the section 'Measures'). Full details of the protocol are provided on Open Science Framework (https://osf.io/k9v2p/).[148]

Chapter 4 Statistical Analysis

Sample Size Calculation

The sample size calculations were completed by the independent trial statistician (DR) and conducted using assumptions based off of the Egnell et al. study, using the smallest relative change of the change in correct ranking of product healthiness (44% ranked pizza correctly at baseline, 57% at follow-up).[69] Assuming a correlation of 0.3 between paired observations, a sample size of 2,400 was calculated (480 per group) to detect differences of five percent at 90% power, accounting for design effects, non-response and ineligibility.

General Statistical Methods

The analysis methods will be presented in each respective Chapter. The primary outcome results of ability to rank all products correctly are presented in Chapter 5; the secondary outcome results of ranking perceptions, label perceptions, speed of ranking and the ability to correctly rank the healthiest product correctly are presented in Chapter 6. Generally, all analyses were adjusted for stratification variables and the pre-specified covariates: ethnicity, highest education level, household composition, food shopping responsibility and current FOPL use; and models were weighted to account for non-response and to ensure findings were representative of the British adult population.

Ethics, Data Protection and Data Management

This study involved primary research collection using the NatCen panel. The main ethical and management issues were around the collection of data that may be perceived as sensitive and data protection. Participants were able to withdraw their data up to one week after fieldwork finished. The study was reviewed by the NatCen Ethics Review committee on the 28th of August 2020 (application reference: P15640). Ethical approval was granted on 4th of September 2020. The project was registered with the UCL Data Protection office, confirmed on 9th of September 2020 (reference number Z6364106/2020/09/31) and granted UCL Research and Development (number 20PP07). See Appendix C for the NatCen ethics committee decision letter, the UCL Data Protection project registration email and the data flowchart.

Chapter 4 Conclusions

In this Chapter, I have presented the design and methodology of the online randomised controlled experiment and now can present the results of the experiment that address the main thesis aims and objectives. In Chapter 5, I will present the key results of the primary outcome, the ability of participants to understand product healthiness assessed through a ranking task; and the impact of participant characteristics on this ability. In Chapter 6, I will present the secondary outcomes (correct ranking of the healthiest product, speed of ranking, ranking perceptions and label perceptions), followed by Chapter 7 where I will present the exploratory latent class analysis examining the impact of participant engagement and motivation characteristics on the experiment outcomes.

Chapter 5: Ability to Correctly Rank the Products According to Healthiness, by FOPL Scheme and Demographics

Chapter 5 Background

In this Chapter, I will present the findings addressing the primary thesis aim and outcome of the experiment, to see if FOPLs can improve objective understanding of product healthiness. This was tested through the ability of participants to correctly rank all three products according to the product healthiness. I will also present the primary outcome results by demographic characteristics, which is important for policy decisions. The secondary outcomes of the experiment are presented in Chapter 6, which address other important factors of label engagement, including speed of ranking, perceptions of ranking and of the labels.

As outlined in Chapter 1, the UK Government has committed to reviewing the current FOPL policy, to ensure that the country's labelling scheme remains based on the latest evidence.[20] Although evidence to date supports FOPLs as a tool to improve consumers' ability to identify healthier foods (as addressed in Chapter 2) and encourage healthier food purchasing (as addressed in Chapter 3), research in representative UK samples has been limited. This research sought to inform policy making around future FOPL options in the UK. I aimed to examine whether FOPLs were effective at improving participant understanding of the healthiness of foods and drinks in a large population-based British sample (including England, Scotland and Wales), and to explore whether this is influenced by level of education. The primary objectives were to identify if FOPLs (MTL, N-S, WL, PC) were more effective at improving participants' understanding of the healthiness of food items compared to no label control (research objective 2), and to identify which FOPL scheme was most effective (research objective 3). The main outcomes were the ability to correctly rank the products at baseline and follow-up (FOPLs vs. no label control), and the change in correct ranking from baseline to follow-up for products categories individually and combined. Given the evidence supporting N-S, a secondary objective was to compare the effectiveness for enabling accurate product ranking for healthiness between N-S and the current UK label (MTL). I also conducted descriptive analyses of participant characteristics (including demographics, engagement/motivation to use labels and medical conditions) for correct ranking at follow-up and mean global food score (research objective 4).

Publication status: These results have been published in Nutrients and delivered as a briefing paper to policy makers. Packer, J., Russell, S. J., Ridout, D., Hope, S., Conolly, A., Jessop, C., Robinson, O. J., Stoffel, S. T., Viner, R. M., & Croker, H. (2021). Assessing the Effectiveness of Front of Pack Labels: Findings from an Online Randomised-Controlled Experiment in a Representative British Sample. Nutrients, 13(3), 900. https://doi.org/10.3390/nu13030900

Chapter 5 Methods

The overall methods for the online randomised controlled experiment have been described in Chapter 4. The primary outcome examined in this Chapter was participants' ability to accurately rank products according to their healthiness. This outcome was examined in three ways:

- Ranking of healthiness of products (correct for all products = 1, any incorrect = 0) at baseline and follow-up;
- Change in ranking from baseline to follow-up (improved = +1; no change = 0; worsened = −1);
- Change in global food score, with change for each of the five food categories, i.e., excluding the instant hot chocolate (see section 'Chapter 5 Statistical Analysis') summed to give an overall score (range +5, all categories improved, to −5, all categories worsened)

The primary outcome was also descriptively examined by participant characteristics, including:

- Sex – female, male
- Age – categorised by NatCen as 18–29, 30–39, 40–49, 50–59, 60–69, 70+ years
- Ethnicity – dichotomised as: white (white British/white other) vs non-white (mixed or multiple ethnic groups/ Asian or Asian British/ black or black British/ other)
- Educational level – presented categorically and dichotomised as: higher education (A-levels or vocational level 3 or equivalent and above, equivalent to end of high school education in the UK) vs. lower education (below A-level or equivalent).

- Equivalised household income per month – presented categorically and dichotomised as: More than £2000 vs £2000 and below (£800 or less, £801-1250, £1251-2000)
- Household composition (any children under 16 years)
- Whether respondents were responsible for food shopping (yes/no)
- Current FOPL use – dichotomised as: often (very often/ quite often) vs not often (occasionally/ rarely/ never)
- Interest in healthy eating – dichotomised as: very (very interested) vs not very (quite interested/ not very interested/ not at all interested)
- Whether they had consumed or purchased each viewed food in the last 12 months (yes/no)
- Body Mass Index (BM) – categorised as underweight (<18.5), normal (18.5-<25), overweight (25-<30), obese (30+)

Chapter 5 Statistical Analysis

Participants' demographics (age, sex, SES, educational level), shopping habits, nutritional knowledge and label use were summarised for the full sample and by experimental group. The ability of participants to rank products according to healthiness was analysed using three statistical tests, 1) the proportion of correct rankings at follow-up, adjusted for baseline ranking was analysed using a multilevel log-binomial model; 2) the improvement in correct rankings from baseline to follow-up was tested using a log-binomial model; 3) the change in global food score from baseline to follow-up was tested using a multiple linear regression analysis. For these outcomes, the difference between each FOPL group compared to the no label control and a policy-specific comparison between MTL and N-S was tested. All models were adjusted for the five stratification factors and the following pre-specified covariates: ethnicity, highest education level, household composition, food shopping responsibility, and current FOPL use. Models were weighted to account for non-response and to ensure findings were representative of the British adult population. To prevent that lack of familiarity with products did not impact results, participants who reported they had not bought or consumed products from a particular product category in the last 12 months were excluded from analyses (see Table 1 for these numbers). For assessment of the change in global food score, only participants who reported buying or consuming all five food products were included. Given this requirement, the change in ranking of the instant hot chocolate was excluded from this analysis, as only 34% reported consuming/purchasing this product, compared to 74% for pizza, 70% for cake, 82% for crisps, 83% for yoghurt and 84% for cereal, see Figure D1 in the Appendix D. Analyses were not blinded, but the primary outcomes were carried out by a statistician (DR) who had not been involved with setting up and running the experiment, as per best practice for randomised controlled trials; I then led the interpretation and writing up of the findings and all other analyses. Stata software (Release 16, StataCorp LLC., College Station, TX, USA) was used for all analyses and a significance level of 5% was used.[149] Results are presented as relative risk (RR) of linear regression coefficient, with 95% confidence intervals.

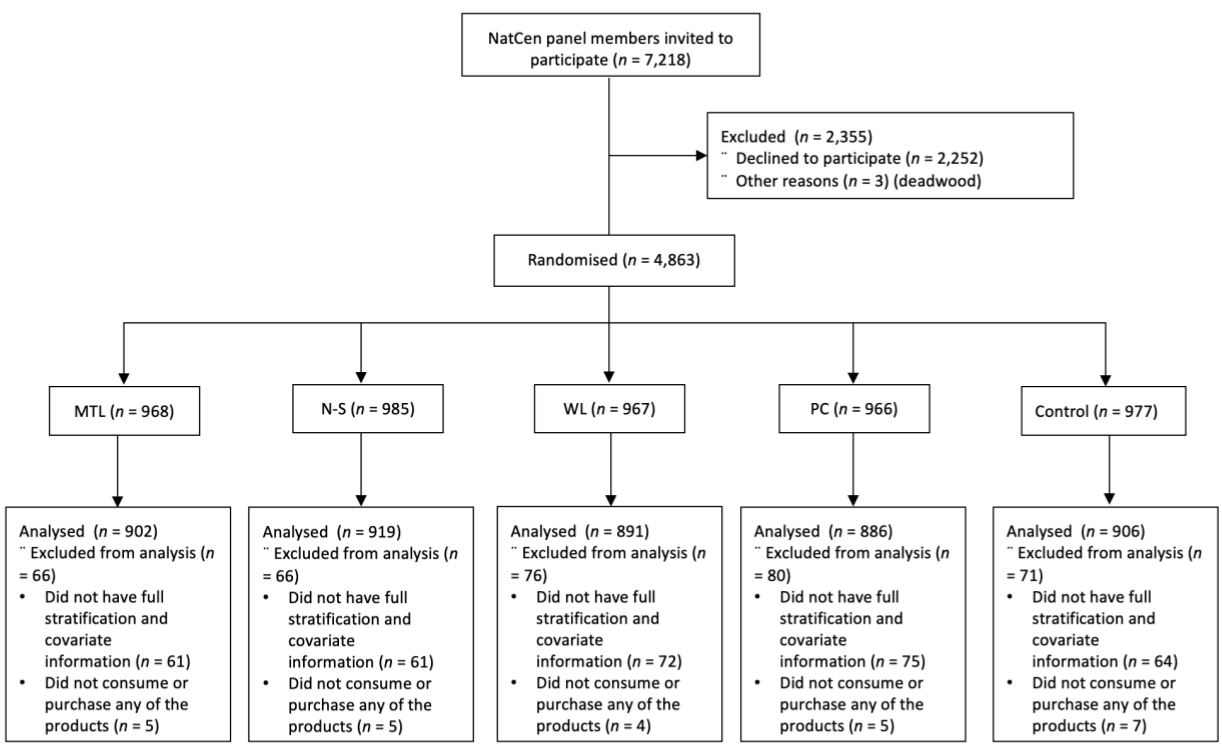
Statistical analyses were limited to a pre-defined set of analyses agreed with DHSC and stated in the statistical analysis plan. Additional analyses were confined to descriptive analyses, to limit errors from multiple testing. Descriptive findings for two variations of the primary outcome (proportion of participants correct at follow-up ranking and mean global food score) are presented by participant characteristics, as outlined above. Both correct ranking at follow-up and mean global food score results are presented, as they assess different concepts i.e., overall number of participants who could rank the products correctly or the improvement in ranking ability from baseline to follow-up. I ran sensitivity analyses, using the full sample of participants in the models (n = 4,530), irrespective of whether they consumed the particular food, see Table D2. To avoid multiple statistical testing, additional statistical analyses are presented in Appendix D and were limited to assessment of whether education, income and self-reported interest in healthy eating were significantly associated with change in global food score, comparing only between MTL and N-S groups. This was undertaken using multiple linear regression analysis with interaction terms for education, income, interest in healthy eating and change in the global food score as the outcome. This additional exploratory analysis was conducted following a request from DHSC.

Chapter 5 Results

Sample Characteristics

Of the 7,218 panel members contacted, 4,863 agreed to participate (67% response rate) and were randomly allocated to: MTL (*n* = 968), N-S (*n* = 985), WL (*n* = 967), PC (*n* = 966) and control (*n* = 977). See CONSORT flowchart for more detail (Figure 20). The average time taken to complete the online survey was 13 min (SD 8 min), not including the demographic data routinely collected by NatCen. Complete stratification and covariate information was available for 4,530 participants, of which 4,504 were included in the primary outcome analysis. The number with missing data for the primary outcomes was very low (≤2% missing per variable, see Table D1 in Appendix D). The sample for each product category analysis differed slightly as inclusion was dependent on consuming or purchasing that specific product (see Table D2 in Appendix D for the sensitivity analyses using the full 4,530 sample, results were consistent with the main analyses). The global food score sample was 1,976, as inclusion was dependent on participants consuming or purchasing all five food products.

Figure 20: CONSORT flowchart. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.



The individual participant characteristics are provided in Table 9 by the primary outcome analysis sample (*n* = 4,504) and the sample with full covariate information (*n* = 4,530). Baseline characteristics and sample sizes were similar across experimental groups (see Table D3 in Appendix D for characteristics by experimental group). Overall, 57% of participants were female, 93-94% were white British or white other, 48% had a degree or higher, 30% had children under 16 years old in the household, 96% had food shopping responsibilities, and 55% reported currently using food labels very or quite often. The proportion of participants consuming each product category was roughly equivalent across the five food groups.

Table 9: Sample Characteristics - individual characteristics of the primary outcome analysis sample (*n* = 4,504; unweighted) and sample with full covariate information (*n* = 4,530; unweighted)

	***n* = 4,504**	***n* = 4,530**
	***n* (%)**	***n* (%)**
Sex		
Female	2,570 (57)	2,586 (57)
Male	1,934 (43)	1,944 (43)
Age[i]		
18–29	278 (6)	279 (6)
30–39	639 (14)	642 (14)
40–49	878 (20)	881 (19)
50–59	962 (21)	969 (21)
60–69	924 (21)	933 (21)
70+	820 (18)	823 (18)
Ethnicity		
White British	3,954 (88)	3,974 (87)
White other	257 (6)	259 (6)
Mixed or multiple ethnic groups	54 (1)	54 (1)
Asian or Asian British	152 (3)	155 (3)
Black or Black British	71 (2)	72 (2)
Other	16 (0)	16 (0)
	***n* (%)**	***n* (%)**
Education		
Degree or equivalent	2,182 (48)	2,196 (48)
A-levels or vocational level 3 or equivalent	874 (19)	878 (19)
Other qualifications below A-levels or equivalent	781 (17)	784 (17)
Other qualification	269 (6)	271 (6)
No qualifications	398 (9)	401 (9)
Latest household income equivalised (grouped)[ii]		
800 or less	628 (14)	631 (14)
801 to 1250	880 (20)	887 (20)
1251 to 2000	1,142 (25)	1,148 (25)
More than 2000	1,692 (38)	1,702 (38)

Children in household		
Yes	1,373 (30)	1,377 (30)
No	3,131 (70)	3,153 (70)
Shopping responsibility		
Yes—some or all	4,316 (96)	4,340 (96)
No—someone else does	188 (4)	190 (4)
Current label use		
Very often	934 (21)	945 (21)
Quite often	1,536 (34)	1,546 (34)
Occasionally	1,316 (29)	1,318 (29)
Rarely	558 (12)	559 (12)
Never	160 (4)	162 (4)
Reported consuming or buying product in past 12 months		
Pizza	3,361 (75)	3,361 (74)
Drink	1,630 (36)	1,630 (36)
Cake	3,163 (70)	3,163 (70)
Crisps	3,723 (83)	3,723 (82)
Yoghurt	3,779 (84)	3,779 (83)
Breakfast cereal	3,802 (84)	3,802 (84)
Currently trying to lose weight		
Yes	2,115 (47)	2,125 (47)
No	2,226 (49)	2,240 (49)
Prefer not to say	163 (4)	165 (4)
Interested in healthy eating		
Very interested	1,881 (42)	1,894 (42)
Quite interested	2,320 (52)	2,332 (52)
Not very interested	280 (6)	280 (6)
Not at all interested	23 (1)	24 (1)
Knowledge in healthy eating[iii]		
A lot of knowledge	1,262 (28)	1,275 (28)
Some knowledge	2,617 (58)	2,629 (58)
A little knowledge	602 (13)	602 (13)
No knowledge	22 (1)	23 (1)
Medical conditions that affect eating[iii]		
Yes	734 (16)	740 (16)
No	3,769 (84)	3,789 (84)
Type 1 Diabetes		
Yes	39 (1)	39 (1)
Type 2 Diabetes		
Yes	310 (7)	310 (7)
	Mean (SD)	**Mean (SD)**
BMI[iv]	26.7 (5.5)	26.7 (5.5)

[i]Adds up to 4,501, [ii]Criteria not met for calculating (n = 162, 3.6%), [iii] 1 refused to answer (0.02%). BMI, body mass index, [iv] participants could choose not to say, the sample reporting this information = 3396/4504 and 4018/ 4530

Primary Outcome – Correct Ranking of All Three Products

Figure 21 shows the number of participants who correctly ranked products at baseline and follow-up and the percentage change in correct ranking for each FOPL group, by product category. The proportion of participants who correctly ranked the products at baseline and follow-up are shown in Table D4 in Appendix D. The associations between the experimental group and correct ranking of the products at follow-up, adjusted for covariates, are shown in Table 10. The probability of participants correctly ranking the products at follow-up was significantly greater across all products for N-S, MTL and WL groups, compared to control (all $p < 0.001$; except WL drink, $p = 0.01$). N-S was associated with the greatest probability of correctly ranking the products followed by MTL, then WL and PC, compared to the no label control. The PC group showed significant differences only for the drink and yoghurt categories. The comparison between N-S and MTL groups showed that participants in the N-S group were significantly more likely than the MTL group to correctly rank the drinks only.

Figure 21: Number of participants who correctly ranked products at baseline and follow-up, by FOPL group and product category: (a) Pizza; (b) Instant hot chocolate; (c) Cake; (d) Crisps; (e) Yoghurt; (f) Breakfast cereal.† Cake and crisps categories had no products qualify for Positive Choice tick. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

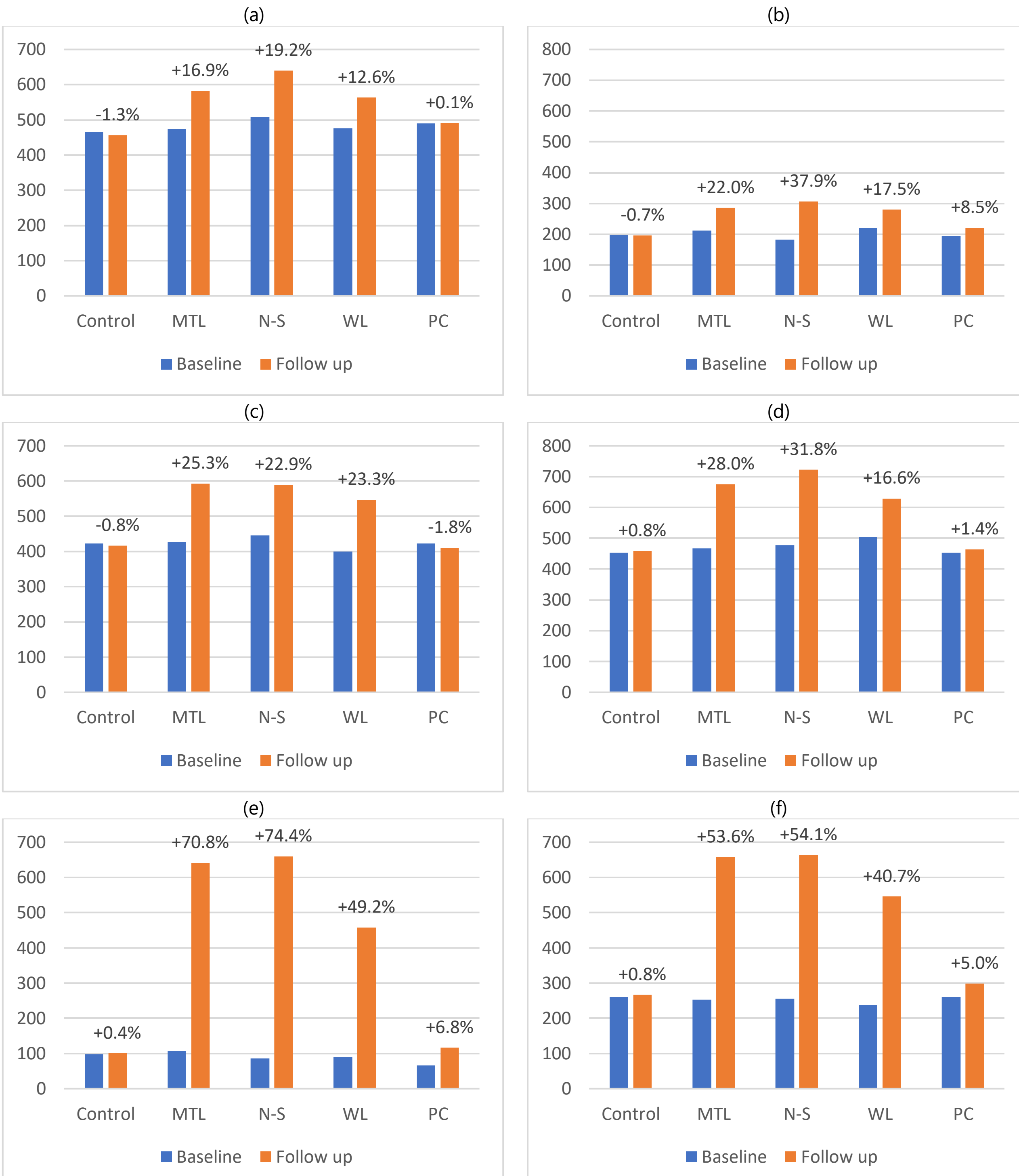


Table 10: Multilevel log-binomial regression results—follow-up (FOPL group) correct (yes/no) and adjusted for baseline rank compared to control (also adjusted for design effects and covariates as planned).

	MTL vs. Control RR (95%CI)	N-S vs. Control RR (95%CI)	WL vs. Control RR (95%CI)	PC vs. Control RR (95%CI)	N-S v MTL RR (95%CI)
Pizza	1.23 (1.11, 1.36) p < 0.001	1.29 (1.18, 1.41) p < 0.001	1.20 (1.09, 1.32) p < 0.001	0.98 (0.90, 1.07) 0.69	1.05 (0.95, 1.16) p = 0.32
Drink	1.29 (1.13, 1.46) p < 0.001	1.61 (1.41, 1.84) p < 0.001	1.16 (1.03, 1.30) p = 0.01	1.18 (1.04, 1.33) p < 0.01	1.25 (1.08, 1.46) p < 0.01
Cake	1.47 (1.32, 1.63) p < 0.001	1.42 (1.29, 1.57) p < 0.001	1.38 (1.25, 1.52) p < 0.001	1.04† (0.96, 1.12) 0.38	0.97 (0.87, 1.08) p = 0.55
Crisps	1.45 (1.32, 1.59) p < 0.001	1.48 (1.35, 1.62) p < 0.001	1.26 (1.16, 1.37) p < 0.001	1.03† (0.95, 1.11) 0.50	1.02 (0.92, 1.14) p = 0.70
Yoghurt	5.72 (4.30, 7.59) p < 0.001	6.86 (4.90, 9.60) p < 0.001	4.22 (3.02, 5.90) v < 0.001	1.73 (1.25, 2.38) p <0.01	1.20 (0.83, 1.74) p = 0.34
Breakfast cereal	2.48 (2.09, 2.96) p < 0.001	2.61 (2.18, 3.12) p < 0.001	2.35 (1.93, 2.88) p < 0.001	1.14 (0.95, 1.36) p = 0.15	1.05 (0.87, 1.27) p = 0.60

† Cake and crisps categories had no products qualify for Positive Choice tick; all analyses were adjusted for baseline ranking (correct/incorrect), stratification factors (year of recruitment to panel, sex, age, government office region, household income), and the following pre-specified covariates: ethnicity, highest education level, household composition, food shopping responsibility, and current FOPL use. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick; RR: Relative Risk; CI: Confidence Interval.

The associations between the experimental groups and an improved ranking score, adjusted for covariates, are shown in Table 11. The results showed that N-S, MTL and WL significantly increased the probability of participants improving their score from baseline to follow-up, compared to control (all $p < 0.001$). The probability of improved ranking varied across N-S, MTL and WL groups, with the magnitude of RRs in that order. PC was associated with a greater probability of improved ranking score for only the drink, yoghurt, and breakfast cereal categories. The comparison between N-S and MTL showed that N-S significantly increased the probability of improving in ranking score compared to MTL for drinks only.

Table 11: Log-binomial regression results–relative risk that ranking improved (follow-up vs. baseline) between FOPL group and control (adjusted for design factors and covariates as planned).

	MTL vs. Control RR (95%CI)	N-S vs. Control RR (95%CI)	WL vs. Control RR (95%CI)	PC vs. Control RR (95%CI)	N-S vs. MTL RR (95%CI)
Pizza	2.66 (1.72, 4.11) *p* < 0.001	2.70 (1.76, 4.15) *p* < 0.001	2.36 (1.51, 3.67) *p* < 0.001	0.90 (0.56, 1.46) 0.67	1.02 (0.78, 1.33) *p* = 0.90
Drink	4.27 (2.36, 7.74) *p* < 0.001	6.10 (3.43, 10.84) *p* < 0.001	2.95 (1.61, 5.41) *p* < 0.001	2.37 (1.17, 4.77) *p* < 0.001	1.43 (1.04, 1.96) *p* = 0.03
Cake	6.89 (3.80, 12.50) *p* < 0.001	6.21 (3.46, 11.14) *p* < 0.001	5.60 (3.11, 10.10) *p* < 0.001	0.91† (0.41, 2.02) 0.82	0.90 (0.70, 1.15) *p* = 0.40
Crisps	5.28 (3.59, 7.76) *p* < 0.001	5.83 (3.97, 8.56) *p* < 0.001	3.69 (2.46, 5.53) *p* < 0.001	1.40† (0.82, 2.40) 0.22	1.10 (0.90, 1.35) *p* = 0.34
Yoghurt	22.36 (13.77, 36.31) *p* < 0.001	22.80 (14.04, 37.04) *p* < 0.001	15.81 (9.69, 25.81) *p* < 0.001	2.92 (1.65, 5.16) *p* < 0.001	1.02 (0.93, 1.12) *p* = 0.68
Breakfast cereal	7.03 (4.96, 9.98) *p* < 0.001	7.75 (5.47, 10.97) *p* < 0.001	6.43 (4.51, 9.17) *p* < 0.001	1.95 (1.29, 2.95) *p* < 0.01	1.10 (0.97, 1.25) *p* = 0.13

† Cake and crisps categories had no products qualify for Positive Choice tick; all analyses were adjusted for the five stratification factors (year of recruitment to panel, sex, age, government office region, household income) and the following pre-specified covariates: ethnicity, highest education level, household composition, food shopping. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick; RR: Relative Risk; CI: Confidence Interval.

The associations between global food score and FOPL groups, adjusted for covariates, are shown in Table 12. The results showed that N-S, MTL and WL were significantly associated with an increase in global food score compared to control (all $p < 0.001$), with the magnitude of effects in that order. No significant change was found for PC. The comparison between N-S and MTL showed that the difference in global food score was small but significant.

Table 12: Multiple regression analysis results—association between global food score and FOPL group (adjusted for design factors and covariates as planned).

	MTL vs. Control RR (95%CI)	N-S vs. Control RR (95%CI)	WL vs. Control RR (95%CI)	PC vs. Control RR (95%CI)	N-S vs. MTL RR (95%CI)
Score (−5, +5) Regression (coefficients)	1.7 (1.6, 1.9) *p* < 0.001	2.1 (1.9, 2.2) *p* < 0.001	1.4 (1.3, 1.6) *p* < 0.001	0.1 (−0.02, 0.3) 0.09	0.3 (0.2, 0.5) *p* < 0.001

Global food score was an aggregated score of correct ranking in the five food products, range −5 to +5 (− indicates worsening and + indicates improvement); all analyses adjusted for stratification factors (year of recruitment to panel, sex, age, government office region, household income) and covariates: ethnicity, highest education level, household composition, food shopping responsibility, current FOPL use. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick; RR: Relative Risk; CI: Confidence Interval.

Descriptive Analysis– Global Food Score by Participant Characteristics

Descriptive analyses of the mean global food score by experimental group and participant characteristics (sex, age, ethnicity, education, equivalised household income, children in household, current label use, trying to lose weight, interest in healthy eating, medical conditions that affect diet, Type 1 and Type 2 diabetes, and BMI category) are shown in Table 13, with a higher global food score indicating a greater improvement in ranking. Education and income were also included as dichotomised variables, consistent with other analyses presented in this thesis.

Overall, there were little to no differences for participant characteristics. Greatest improvement was seen with males compared to females (1.3 vs 1.0), higher education compared to lower education (1.2 vs 1.0), and participants with Type 1 diabetes compared to without (1.8 vs 1.1); there was a small proportion of participants with a diagnosis for Type 1 diabetes (n = 13). When comparing N-S to MTL, N-S appeared to be more stable and showed fewer differences in the mean global food scores, across education and income groups. Participants with lower education (below A-levels) had a lower mean global food score compared to higher education (A-levels and above) in the MTL group (1.7 vs 2.0), but there was no difference in the N-S group (2.1 vs 2.1). Participants with lower income (below £2000) had marginally lower mean global food score, compared to higher income (£2000 and above) in the MTL group (1.9 vs 2.1), but little evidence of this difference in N-S group (2.1 vs 2.0). The experiment was not powered to formally test these differences. See Table D2 in Appendix D for sensitivity analysis conducted in the full analysis sample (n = 4,530), without the requirement for to consume all food products (as required for the global food score analyses). The sensitivity analysis results were consistent with these findings, showing a similar stability for N-S and a larger gradient across MTL.

Table 13: Mean global food score by experimental group and participant characteristics

	Experimental group					
	Control	**MTL**	**N-S**	**WL**	**PC**	**Overall**
Sex						
Female (*n*=1,179)	0.0 (0.8)	1.8 (1.3)	2.0 (1.2)	1.3 (1.3)	0.1 (1.0)	1.0 (1.4)
Male (*n*=797)	0.1 (0.7)	2.1 (1.4)	2.1 (1.2)	1.7 (1.5)	0.1 (0.8)	1.3 (1.5)
Age						
18-29 (*n*=155)	-0.1 (0.7)	1.8 (1.2)	2.1 (0.9)	1.7 (1.3)	0.4 (0.7)	1.2 (1.3)
30-39 (*n*=418)	0.1 (0.6)	1.8 (1.3)	2.2 (1.2)	1.5 (1.2)	0.2 (1.0)	1.1 (1.4)
40-49 (*n*=523)	0.0 (0.7)	1.8 (1.2)	2.0 (1.3)	1.5 (1.4)	0.2 (1.0)	1.1 (1.4)
50-59 (*n*=428)	-0.0 (0.7)	2.0 (1.5)	2.0 (1.1)	1.3 (1.6)	0.1 (0.7)	1.0 (1.5)
60-69 (*n*=284)	0.1 (0.9)	2.3 (1.5)	2.1 (1.3)	1.4 (1.5)	0.1 (0.9)	1.3 (1.6)
70+ (*n*=167)	-0.1 (1.0)	2.3 (1.5)	2.0 (1.2)	1.6 (1.6)	0.1 (0.9)	1.2 (1.6)
Ethnicity dichotomised						
White (*n*=1,859)	0.0 (0.8)	2.0 (1.3)	2.1 (1.2)	1.5 (1.4)	0.1 (0.9)	1.1 (1.5)
Non-white (*n*=117)	0.0 (0.6)	1.5 (1.9)	2.1 (1.5)	1.2 (1.5)	0.1 (0.8)	1.0 (1.6)
Education level						
Degree or equivalent + (*n*=1,040)	0.0 (0.7)	2.0 (1.3)	2.0 (1.2)	1.5 (1.4)	0.2 (0.9)	1.2 (1.4)
A levels or vocational level 3 or equivalent (*n*=401)	0.0 (0.7)	2.0 (1.2)	2.2 (1.1)	1.4 (1.4)	0.7 (0.8)	1.2 (1.4)
Qualifications below A levels or equivalent (*n*=319)	0.0 (0.6)	2.1 (1.3)	2.0 (1.3)	1.6 (1.5)	0.0 (1.0)	1.1 (1.5)
Other qualification (*n*=84)	0.2 (1.3)	1.4 (1.6)	2.7 (1.1)	1.1 (1.2)	-0.1 (1.3)	1.0 (1.6)
No qualification (*n*=132)	-0.0 (1.1)	1.3 (1.9)	2.0 (1.4)	1.1 (1.4)	0.2 (0.9)	0.9 (1.5)
Education dichotomised*						
A Levels + (*n*=1,441)	0.0 (0.7)	2.0 (1.3)	2.1 (1.2)	1.5 (1.4)	0.2 (0.9)	1.2 (1.4)
Below A level (*n*=535)	0.0 (0.9)	1.7 (1.5)	2.1 (1.3)	1.4 (1.5)	0.1 (1.0)	1.0 (1.5)
Equivalised household income (£)						
More than 2000 (*n*=772)	0.0 (0.6)	2.1 (1.2)	2.0 (1.1)	1.6 (1.4)	0.2 (0.9)	1.2 (1.4)
1251- 2000 (*n*=522)	0.0 (0.8)	2.2 (1.3)	2.1 (1.3)	1.4 (1.4)	0.0 (0.8)	1.2 (1.5)
801-1250 (*n*=358)	0.0 (0.8)	1.9 (1.2)	2.0 (1.2)	1.5 (1.4)	0.2 (0.9)	1.1 (1.4)
800 or less (*n*=253)	0.1 (0.7)	1.3 (1.7)	2.2 (1.1)	1.1 (1.6)	0.3 (1.1)	1.0 (1.5)
Income dichotomised*						
More than £2000 (*n*=772)	0.0 (0.6)	2.1 (1.2)	2.0 (1.1)	1.6 (1.4)	0.2 (0.9)	1.2 (1.4)
£2000 and below (*n*=1,133)	0.0 (0.8)	1.9 (1.4)	2.1 (1.3)	1.4 (1.5)	0.1 (0.9)	1.1 (1.5)
Children in household						
Yes (*n*=886)	0.0 (0.7)	1.9 (1.4)	2.1 (1.2)	1.4 (1.4)	0.2 (1.0)	1.1 (1.4)

No (*n*=1,090)	0.0 (0.8)	2.0 (1.3)	2.0 (1.2)	1.5 (1.4)	0.1 (0.8)	1.2 (1.5)
Current label use dichotomised*						
Often (*n*=1,052)	0.0 (0.7)	1.9 (1.3)	2.1 (1.2)	1.5 (1.4)	0.2 (0.9)	1.2 (1.4)
Not often (*n*=924)	0.0 (0.8)	2.0 (1.5)	2.0 (1.3)	1.4 (1.5)	0.1 (0.9)	1.1 (1.5)
Currently trying to lose weight						
Yes (*n*=935)	0.0 (0.7)	1.9 (1.3)	2.1 (1.2)	1.3 (1.5)	0.2 (0.9)	1.2 (1.5)
No (*n*=965)	0.0 (0.7)	2.0 (1.4)	2.0 (1.2)	1.6 (1.4)	0.1 (1.9)	1.1 (1.5)
Interest in healthy eating dichotomised*						
Very (*n*=727)	0.0 (0.7)	2.0 (1.2)	2.1 (1.1)	1.4 (1.5)	0.2 (0.9)	1.2 (1.4)
Not very (*n*=1,249)	-0.0 (0.8)	1.9 (1.4)	2.0 (1.3)	1.5 (1.4)	0.1 (0.9)	1.0 (1.5)
Medical conditions that affect diet						
Yes (*n*=267)	0.0 (1.1)	2.2 (1.5)	1.9 (1.2)	1.4 (1.4)	0.2 (1.0)	1.2 (1.5)
No (*n*=1,708)	0.0 (0.7)	1.9 (1.3)	2.1 (1.2)	1.5 (1.4)	0.1 (0.9)	1.1 (1.5)
Type 1 diabetes						
Yes (*n*=13)	0.0 (0.0)	2.2 (1.3)	2.7 (1.5)	2.0 (0.0)	0.5 (0.7)	1.8 (1.4)
No (*n*=1,963)	0.0 (0.8)	2.0 (1.4)	2.1 (1.2)	1.5 (1.4)	0.1 (0.9)	1.1 (1.5)
Type 2 diabetes						
Yes (*n*=92)	0.1 (1.3)	2.4 (1.6)	1.9 (1.4)	1.1 (1.5)	0.2 (0.9)	1.2 (1.6)
No (*n*=1,884)	0.0 (0.7)	1.9 (1.3)	2.1 (1.2)	1.5 (1.4)	0.1 (0.9)	1.1 (1.5)
BMI category						
Underweight (*n*=21)	0.0 (0.0)	0.5 (1.6)	1.0 (0.0)	1.0 (1.4)	0.7 (1.2)	0.5 (1.2)
Normal (*n*=757)	0.0 (0.7)	2.0 (1.3)	2.1 (1.1)	1.5 (1.3)	0.1 (1.0)	1.2 (1.4)
Overweight (*n*=572)	0.0 (0.7)	1.9 (1.3)	2.2 (1.3)	1.6 (1.6)	0.1 (0.8)	1.2 (1.5)
Obese (*n*=386)	0.1 (0.7)	2.0 (1.5)	2.1 (1.3)	1.2 (1.4)	0.2 (1.1)	1.1 (1.5)

Global food score was an aggregated score of correct ranking in the five food products, range −5 to +5 (− indicates worsening and + indicates improvement). Participants required to buy/eat each food product in the last 12 months and have full co-variate information to be included. *Variables were dichotomised as: ethnicity, white (white British/white other) vs non-white (mixed or multiple ethnic groups/ Asian or Asian British/ black or black British/ other); household equivalised income, more than 2000 vs 200 and below (800 or less, 801-1250, 1251-2000); current label use, often (very often/ quite often) vs not often (occasionally/ rarely/ never); interest in healthy eating, very (very interested) vs not very (quite interested/ not very interested/ not at all interested). MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick; SD: standard deviation.

Descriptive Tables– Proportion Of Correct Ranking of All Three Products

Descriptive analyses of the proportion of participants correctly ranking products at follow-up, by experimental group and dichotomised participant sociodemographic, food and label habits/motivations and medical characteristics are shown in: Table 14 for sex (female/male); Table 15 for ethnicity (white/non-white); Table 16 for education (higher/lower); Table 17 for income (higher/lower);Table 18 for children in household (yes/no); Table 19 for current label use (often/not often); Table 20 for currently trying to lose weight (yes/no); Table 21 for interest in healthy eating (very/not very); Table 22 for medical conditions that affect diet (yes/no); Table 23 for Type 1 diabetes (yes/no); Table 24 for Type 2 diabetes (yes/no). Participants needed to have complete covariate information and consume the product specific to the analysis (but not all products).

By sociodemographic characteristics, female participants (57%; Table 14) had a higher proportion who correctly ranked products, across FOPL groups and product categories, compared to males. However, the difference between the sexes varied, with the N-S group appearing to have a smaller difference. White participants (94%; Table 15) had a higher proportion who correctly ranked products, across FOPL groups and product categories, compared to non-white. However, the difference between the two ethnicity categories varied considerably, with the N-S group appearing to have a smaller difference compared to MTL, and also the highest proportion who correctly ranked products. Participants with a higher education (68%; Table 16) had a higher proportion who correctly ranked the products, across FOPL groups and product categories, compared to those with a lower education. However, the difference between the two education groups varied considerably, with the N-S group appearing to have a smaller difference. No significant interactions between N-S and MTL (see Table D5). Participants with a higher income (39%; Table 17) had a higher proportion who correctly ranked products, across FOPL groups and product categories, compared to participants with a lower income. No significant interactions between N-S and MTL and income (see Table D5). Participants with children under 16 years old in the household (30%; Table 18) showed no consistent differences, compared to participants without children in household, across FOPL groups and product categories.

By food and label habits/motivations, participants who reported currently using labels often (very/quite) (55%; Table 19) had a higher proportion who correctly ranked products, across FOPL groups and product categories, compared to participants who did not use labels often (occasionally/ rarely/ never), (apart from drinks in N-S group). The proportion of participants who ranked products correctly was greatest in N-S across participants who did not use labels often, compared to the other FOPL groups. Participants who reported currently trying to lose weight (47%; Table 20) had a higher proportion who correctly ranked products, compared to those who were not currently trying to lose weight, across all product categories in MTL and N-S groups. Findings were mixed for the WL group. The proportion who correctly ranked products was higher overall in the N-S group, compared to other FOPL groups. Participants who were very interested in healthy eating (42%; Table 21) had higher proportion who correctly ranked products, across all FOPL groups and product categories, compared to those who were not as interested (quite/not very/ not at all). The proportion who correctly ranked products was higher overall in the N-S group, compared to other FOPL groups. No significant interactions between N-S and MTL (see Table D5) but global food score was significantly higher for very interested compared to quite/not very/ not at all interested (0.2 points, p=0.007).

By medical conditions, participants with medical conditions that affect their diet (16%; Table 22), had no difference in the proportion who correctly ranked products, compared to participants without medical conditions affecting diet, across FOPL groups and product categories. Participants with Type 1 diabetes (1%; Table 23) had a higher proportion who correctly ranked products, across all FOPL groups and product categories, compared to participants without. The N-S group appeared to have smaller differences between participants, compared to MTL. Participants with Type 2 diabetes (7%; Table 24) tended to have a lower proportion who correctly ranked products, compared to participants without type 2 diabetes, across FOPL groups and product categories. By BMI category, participants categorised as being in the normal range generally had a higher proportion correctly rank the products at follow-up, compared to those categorised as overweight/obese (Table 25). There were differences between product categories and experimental groups, with N-S and MTL the most stable within each product category.

Table 14: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and sex (female/male)

Sex	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Female	319/349 (91%)	366/392 (93%)	326/387 (84%)	287/372 (77%)	269/386 (70%)
Male	263/298 (88%)	274/294 (93%)	237/303 (78%)	204/296 (69%)	187/284 (66%)
Drink					
Female	164/192 (85%)	191/201 (95%)	185/219 (84%)	141/198 (71%)	131/198 (66%)
Male	122/144 (85%)	115/123 (93%)	95/117 (81%)	80/123 (65%)	65/115 (57%)
Cake					
Female	326/364 (90%)	348/368 (95%)	316/368 (86%)	228/346 (66%)	238/360 (66%)
Male	267/287 (93%)	242/265 (91%)	231/267 (87%)	183/271 (68%)	179/267 (67%)
Crisps					
Female	382/421 (91%)	428/456 (94%)	366/436 (84%)	268/406 (66%)	267/429 (62%)
Male	293/318 (92%)	295/316 (93%)	263/317 (83%)	196/301 (65%)	192/323 (59%)
Yoghurt					
Female	364/437 (83%)	406/469 (87%)	261/433 (60%)	80/443 (18%)	60/447 (13%)
Male	278/318 (87%)	254/303 (84%)	197/314 (63%)	37/301 (12%)	41/314 (13%)
Breakfast cereal					
Female	357/416 (86%)	388/437 (89%)	309/438 (71%)	182/436 (42%)	165/447 (37%)
Male	301/341 (88%)	276/318 (87%)	237/318 (75%)	117/319 (37%)	102/332 (31%)

MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 15: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and ethnicity (white/non-white)

Ethnicity	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
White	543/597 (91%)	604/644 (94%)	531/649 (82%)	465/626 (74%)	430/627 (69%)
Non-white	39/50 (78%)	36/42 (86%)	32/41 (78%)	26/42 (62%)	26/43 (60%)
Drink					
White	264/306 (86%)	285/301 (95%)	268/317 (85%)	212/305 (70%)	182/285 (64%)
Non-white	22/30 (73%)	21/23 (91%)	12/19 (63%)	9/16 (56%)	14/28 (50%)
Cake					
White	556/602 (92%)	555/592 (94%)	517/598 (86%)	389/583 (67%)	393/589 (67%)
Non-white	37/49 (76%)	35/41 (85%)	30/37 (81%)	22/34 (65%)	24/38 (63%)
Crisps					
White	638/689 (93%)	675/721 (94%)	595/710 (84%)	438/668 (66%)	432/707 (61%)
Non-white	37/50 (74%)	48/51 (94%)	34/43 (79%)	26/39 (67%)	27/45 (60%)
Yoghurt					
White	608/706 (86%)	619/722 (86%)	437/705 (62%)	109/693 (16%)	97/712 (14%)
Non-white	34/49 (69%)	41/50 (82%)	21/42 (50%)	8/51 (16%)	4/49 (8%)
Breakfast cereal					
White	623/705 (88%)	627/710 (88%)	522/714 (73%)	287/711 (40%)	257/741 (35%)
Non-white	35/52 (67%)	37/45 (82%)	24/42 (57%)	12/44 (27%)	10/38 (26%)

Ethnicity dichotomised as white (White British; White other) and non-white (Mixed or multiple ethnic groups; Asian or Asian British; Black or Black British; Other). MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 16: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and education level (higher/lower)

Education	FOPL group				
	MTL	N-S	WL	PC	Control
Pizza					
Higher	427/460 (93%)	462/493 (94%)	411/479 (86%)	358/478 (75%)	333/474 (70%)
Lower	155/187 (83%)	178/193 (92%)	152/211 (72%)	133/190 (70%)	123/196 (63%)
Drink					
Higher	209/240 (87%)	232/245 (95%)	205/242 (85%)	160/230 (70%)	143/226 (63%)
Lower	77/96 (80%)	74/79 (94%)	75/94 (80%)	61/91 (67%)	53/87 (61%)
Cake					
Higher	406/439 (92%)	426/451 (94%)	383/437 (88%)	302/434 (70%)	303/429 (71%)
Lower	187/212 (88%)	164/182 (90%)	164/198 (83%)	109/183 (60%)	114/198 (58%)
Crisps					
Higher	468/500 (94%)	503/533 (94%)	442/520 (85%)	312/477 (65%)	325/526 (62%)
Lower	207/239 (87%)	220/239 (92%)	187/233 (80%)	152/230 (66%)	134/226 (59%)
Yoghurt					
Higher	452/521 (87%)	481/545 (88%)	327/515 (63%)	86/520 (17%)	75/532 (14%)
Lower	190/234 (81%)	179/227 (79%)	131/232 (56%)	31/224 (14%)	26/229 (11%)
Breakfast cereal					
Higher	460/503 (91%)	477/512 (93%)	380/510 (75%)	224/512 (44%)	191/532 (36%)
Lower	198/254 (78%)	187/243 (77%)	166/246 (67%)	75/243 (31%)	76/247 (31%)

Education level dichotomised as A levels and above vs. below A levels, higher or lower education respectively.
MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 17: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and household income (higher vs lower)

Income	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Higher	215/233 (92%)	261/271 (96%)	235/280 (84%)	191/257 (74%)	198/272 (73%)
Lower	347/393 (88%)	359/395 (91%)	307/385 (80%)	278/386 (72%)	243/374 (65%)
Drink					
Higher	98/110 (89%)	122/130 (94%)	106/124 (85%)	79/110 (72%)	65/114 (57%)
Lower	179/215 (83%)	176/186 (95%)	167/204 (82%)	133/198 (67%)	120/186 (65%)
Cake					
Higher	204/221 (92%)	234/244 (96%)	220/247 (89%)	163/241 (68%)	172/239 (72%)
Lower	364/401 (91%)	340/373 (91%)	309/368 (84%)	225/349 (65%)	232/365 (64%)
Crisps					
Higher	245/263 (93%)	277/289 (96%)	251/298 (84%)	171/264 (65%)	182/297 (61%)
Lower	409/450 (91%)	427/463 (92%)	353/427 (83%)	281/417 (67%)	255/422 (60%)
Yoghurt					
Higher	236/265 (89%)	273/306 (89%)	190/294 (65%)	55/294 (19%)	38/291 (13%)
Lower	385/463 (83%)	368/446 (83%)	251/425 (59%)	58/421 (14%)	59/440 (13%)
Breakfast cereal					
Higher	238/256 (93%)	276/295 (94%)	218/283 (77%)	125/294 (43%)	106/292 (36%)
Lower	394/470 (84%)	367/439 (84%)	305/443 (69%)	164/435 (38%)	151/452 (33%)

Equivalised income dichotomised as 2000 and above vs. below 2000, higher or lower income, respectively. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 18: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and children in household (yes/no; under 18 years old)

Children in household	**FOPL group**				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Yes	208/229 (91%)	220/234 (94%)	227/269 (84%)	186/243 (77%)	155/245 (63%)
No	374/418 (89%)	420/452 (93%)	336/421 (80%)	305/425 (72%)	301/425 (71%)
Drink					
Yes	125/145 (86%)	119/122 (98%)	134/154 (87%)	108/145 (74%)	91/135 (67%)
No	161/191 (84%)	187/202 (93%)	146/182 (80%)	113/176 (64%)	105/178 (59%)
Cake					
Yes	203/223 (91%)	196/204 (96%)	202/234 (86%)	147/212 (69%)	152/216 (70%)
No	390/428 (91%)	394/429 (92%)	345/401 (86%)	264/405 (65%)	265/411 (64%)
Crisps					
Yes	216/237 (91%)	243/251 (97%)	227/269 (84%)	160/244 (66%)	174/264 (66%)
No	459/502 (91%)	480/521 (92%)	402/484 (83%)	304/463 (66%)	285/488 (58%)
Yoghurt					
Yes	205/239 (86%)	222/240 (93%)	164/255 (64%)	45/252 (18%)	34/249 (14%)
No	437/516 (85%)	438/532 (82%)	294/492 (60%)	72/492 (15%)	67/512 (13%)
Breakfast cereal					
Yes	198/229 (86%)	214/233 (92%)	191/262 (73%)	105/248 (42%)	94/252 (37%)
No	460/528 (87%)	450/522 (86%)	355/494 (72%)	194/507 (38%)	173/527 (33%)

MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 19: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and current label use (often/ not often)

Current label use	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Often	364/387 (94%)	325/338 (96%)	332/386 (86%)	249/326 (76%)	240/338 (71%)
Not often	218/260 (84%)	315/348 (91%)	231/304 (76%)	242/342 (71%)	216/332 (65%)
Drink					
Often	176/204 (86%)	160/170 (94%)	159/187 (85%)	112/157 (71%)	61/103 (63%)
Not often	110/132 (83%)	146/154 (95%)	121/149 (81%)	109/164 (66%)	93/149 (62%)
Cake					
Often	355/382 (93%)	291/304 (96%)	311/359 (87%)	191/296 (65%)	203/309 (66%)
Not often	238/269 (88%)	299/329 (91%)	236/276 (86%)	220/321 (69%)	214/318 (67%)
Crisps					
Often	413/448 (92%)	373/391 (95%)	367/429 (86%)	236/343 (69%)	234/378 (62%)
Not often	262/291 (90%)	350/381 (92%)	262/324 (81%)	228/364 (63%)	225/374 (60%)
Yoghurt					
Often	413/479 (86%)	365/421 (87%)	284/454 (63%)	64/392 (16%)	60/412 (15%)
Not often	229/276 (83%)	295/351 (84%)	174/293 (59%)	53/352 (15%)	41/349 (12%)
Breakfast cereal					
Often	418/460 (91%)	352/387 (91%)	340/450 (76%)	172/387 (44%)	147/407 (36%)
Not often	240/297 (81%)	312/368 (85%)	206/306 (67%)	127/368 (35%)	120/372 (32%)

Current label use dichotomised as often (very often / quite often) and not often (occasionally/ rarely/ never). Survey question: “How often do you read the nutritional information on the front of packaging when buying food or drink?” MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 20: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and currently trying to lose weight (yes/no)

Currently trying to lose weight	**FOPL group**				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Often	281/306 (92%)	329/346 (95%)	242/309 (78%)	220/301 (73%)	223/308 (72%)
Not often	280/317 (88%)	291/320 (91%)	299/353 (85%)	252/341 (74%)	215/333 (65%)
Drink					
Often	129/150 (86%)	164/172 (95%)	135/158 (85%)	101/140 (72%)	95/145 (66%)
Not often	147/174 (84%)	135/144 (94%)	133/163 (82%)	109/164 (66%)	91/153 (59%)
Cake					
Often	276/298 (93%)	307/322 (95%)	238/288 (83%)	173/264 (66%)	202/287 (70%)
Not often	297/330 (90%)	264/291 (91%)	291/326 (89%)	224/328 (68%)	197/313 (62%)
Crisps					
Often	328/354 (93%)	378/397 (95%)	291/343 (85%)	224/323 (69%)	211/340 (62%)
Not often	328/361 (91%)	325/354 (92%)	316/382 (83%)	222/353 (63%)	225/380 (59%)
Yoghurt					
Often	309/362 (85%)	354/402 (88%)	210/345 (61%)	59/338 (17%)	57/360 (16%)
Not often	313/369 (85%)	287/346 (83%)	232/375 (62%)	50/374 (13%)	42/368 (11%)
Breakfast cereal					
Often	302/349 (87%)	337/380 (89%)	246/342 (72%)	151/329 (46%)	120/357 (34%)
Not often	336/382 (89%)	309/354 (87%)	278/385 (72%)	137/397 (35%)	139/394 (35%)

Survey question: “Are you currently trying to lose weight?” (responses = yes, no, prefer not to say). Prefer not to say excluded from this analysis (*n*=163). MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 21: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and interest in healthy eating (very/not very)

Interest in healthy eating	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Very	250/270 (93%)	264/278 (95%)	212/256 (83%)	189/247 (77%)	169/234 (72%)
Not very	332/377 (88%)	376/408 (92%)	351/434 (81%)	302/421 (72%)	287/436 (66%)
Drink					
Very	125/146 (86%)	127/133 (95%)	107/122 (88%)	80/123 (65%)	60/106 (57%)
Not very	161/190 (85%)	179/191 (94%)	173/214 (81%)	141/198 (71%)	136/207 (66%)
Cake					
Very	251/268 (94%)	243/257 (95%)	214/238 (90%)	141/225 (63%)	139/211 (66%)
Not very	342/383 (89%)	347/376 (92%)	333/397 (84%)	270/392 (69%)	278/416 (67%)
Crisps					
Very	291/316 (92%)	303/320 (95%)	239/289 (83%)	183/270 (68%)	158/263 (60%)
Not very	384/423 (91%)	420/452 (93%)	390/464 (84%)	281/437 (64%)	301/489 (62%)
Yoghurt					
Very	298/347 (86%)	315/360 (88%)	183/317 (58%)	46/317 (14.5%)	42/301 (14%)
Not very	344/408 (84%)	345/412 (84%)	275/430 (64%)	71/427 (17%)	59/460 (13%)
Breakfast cereal					
Very	300/338 (89%)	303/333 (91%)	225/315 (71%)	147/310 (47%)	107/301 (36%)
Not very	358/419 (85%)	361/422 (86%)	321/441 (73%)	152/445 (34%)	160/478 (33%)

Interest in healthy eating dichotomised as very interested (very interested) and not very interested (quite interested/ not very interested/ not at all interested). Survey question: “How interested would you say you are in healthy eating?” . MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 22: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and medical conditions that affect diet (yes/no)

Medical conditions that affect diet	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Yes	78/86 (91%)	87/95 (92%)	93/114 (82%)	78/108 (72%)	64/89 (72%)
No	504/561 (90%)	552/590 (94%)	470/576 (82%)	413/560 (74%)	392/581 (67%)
Drink					
Yes	39/46 (85%)	48/48 (100%)	53/64 (83%)	25/47 (53%)	21/38 (55%)
No	247/290 (85%)	258/276 (93%)	227/272 (83%)	196/274 (72%)	175/275 (64%)
Cake					
Yes	83/91 (91%)	72/80 (90%)	95/109 (87%)	56/100 (56%)	52/85 (61%)
No	510/560 (91%)	517/552 (94%)	452/526 (86%)	355/517 (69%)	365/542 (67%)
Crisps					
Yes	103/113 (91%)	107/117 (91%)	118/133 (89%)	93/128 (73%)	61/106 (58%)
No	572/626 (91%)	616/654 (94%)	511/620 (82%)	371/579 (64%)	398/646 (62%)
Yoghurt					
Yes	95/112 (85%)	99/118 (84%)	84/141 (60%)	14/126 (11%)	12/110 (11%)
No	547/643 (85%)	561/653 (86%)	374/606 (62%)	103/618 (17%)	89/651 (14%)
Breakfast cereal					
Yes	301/341 (88%)	276/318 (87%)	237/318 (75%)	117/319 (37%)	102/332 (31%)
No	357/416 (86%)	388/437 (89%)	309/438 (71%)	182/436 (42%)	165/447 (37%)

Survey question: “Do you have any physical or mental health conditions or illnesses lasting or expected to last for 12 months or more which affect what you eat?” (responses = Yes/ No). . MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 23: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and Type 1 diabetes (yes/no)

Type 1 diabetes	**FOPL group**				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Yes	7/8 (88%)	6/6 (100%)	4/5 (80%)	1/2 (50%)	1/4 (25%)
No	575/639 (89%)	634/680 (93%)	559/685 (82%)	490/666 (74%)	455/666 (68 %)
Drink					
Yes	2/2 (100%)	4/4 (100%)	1/1 (100%)	N/A	1/1 (100%)
No	284/334 (85%)	302/320 (94%)	279/335 (83%)	221/321 (69%)	195/312 (63%)
Cake					
Yes	5/6 (83%)	4/4 (100%)	5/6 (83%)	2/3 (67%)	1/3 (33%)
No	588/645 (91%)	586/629 (86%)	542/629 (86%)	409/614 (67%)	416/624 (67%)
Crisps					
Yes	8/9 (89%)	8/8 (100%)	7/7 (100%)	2/3 (67%)	3/6 (50%)
No	667/730 (91%)	715/764 (94%)	622/746 (83%)	462/704 (66%)	456/746 (61%)
Yoghurt					
Yes	9/9 (100%)	9/9 (100%)	5/7 (71%)	1/3 (33%)	0/5 (0%)
No	633/746 (85%)	651/763 (85%)	453/740 (61%)	116/741 (16%)	101/756 (13%)
Breakfast cereal					
Yes	9/9 (100%)	10/10 (100%)	4/6 (67%)	1/3 (33%)	2/6 (33%)
No	649/748 (87%)	654/754 (89%)	542/750 (72%)	298/752 (40%)	265/773 (34%)

Survey question: "Has a doctor or other health professional ever told you that you have any of these conditions?" (response = could select 'Type 1 Diabetes')

Table 24: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and Type 2 diabetes (yes/no)

Type 2 diabetes	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Yes	30/35 (86%)	28/32 (88%)	27/40 (68%)	30/44 (68%)	19/34 (56%)
No	552/612 (90%)	612/654 (94%)	536/650 (82%)	461/624 (74%)	437/636 (69%)
Drink					
Yes	15/18 (83%)	13/14 (93%)	10/14 (71%)	8/18 (44%)	7/18 (38.9%)
No	271/318 (85%)	293/310 (95%)	270/322 (84%)	213/303 (70%)	189/295 (64%)
Cake					
Yes	36/37 (97%)	28/35 (80%)	28/37 (76%)	22/44 (50%)	19/34 (56%)
No	557/614 (91%)	562/598 (94%)	519/598 (87%)	389/573 (68%)	398/593 (67%)
Crisps					
Yes	43/49 (88%)	43/48 (90%)	38/48 (79%)	46/61 (75%)	27/41 (66%)
No	632/690 (92%)	680/724 (94%)	591/705 (84%)	418/646 (65%)	432/711 (61%)
Yoghurt					
Yes	41/47 (87%)	33/43 (77%)	24/48 (50%)	8/60 (13%)	5/44 (11%)
No	601/708 (85%)	627/729 (86%)	434/699 (62%)	109/684 (16%)	96/717 (13%)
Breakfast cereal					
Yes	41/49 (84%)	39/47 (83%)	29/49 (59%)	17/61 (28%)	12/47 (26%)
No	617/708 (87%)	625/708 (88%)	517/707 (73%)	282/694 (41%)	255/732 (35%)

Survey question: "Has a doctor or other health professional ever told you that you have any of these conditions?" (response = could select 'Type 2 Diabetes'). MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table 25: Proportion of participants correct (yes /no) at follow-up, by product category, FOPL group and BMI category classified as normal weight vs overweight or obese

BMI category	FOPL group				
	MTL	N-S	WL	PC	Control
Pizza					
Normal	223/241 (93%)	242/259 (93%)	233/262 (89%)	190/255 (75%)	175/250 (70%)
Overweight/Obese	283/316 (90%)	339/360 (94%)	275/354 (78%)	235/329 (71%)	215/324 (67%)
Drink					
Normal	112/130 (86%)	103/111 (93%)	114/132 (86%)	86/125 (69%)	80/116 (69%)
Overweight/Obese	131/150 (87%)	162/172 (94%)	125/155 (81%)	103/154 (67%)	89/156 (57%)
Cake					
Normal	233/251 (93%)	213/230 (93%)	210/238 (88%)	160/241 (66%)	150/236 (64%)
Overweight/Obese	283/305 (93%)	313/334 (94%)	270/311 (87%)	196/299 (66%)	207/307 (67%)
Crisps					
Normal	255/272 (94%)	263/276 (95%)	239/285 (84%)	168/264 (64%)	161/282 (57%)
Overweight/Obese	327/360 (91%)	379/407 (93%)	315/378 (83%)	237/354 (67%)	234/371 (63%)
Yoghurt					
Normal	256/291 (88%)	252/294 (86%)	188/283 (66%)	40/286 (14%)	41/288 (14%)
Overweight/Obese	300/355 (85%)	343/399 (86%)	224/378 (59%)	55/357 (15%)	48/372 (13%)
Breakfast cereal					
Normal	269/294 (92%)	253/280 (90%)	225/294 (77%)	123/297 (41%)	119/301 (40%)
Overweight/Obese	304/357 (85%)	351/403 (87%)	271/377 (72%)	138/363 (38%)	123/381 (32%)

Survey question: We told participants that we would like to know how tall they were and how much they weighted (without shoes and clothes), with the ability to opt out and to respond with imperial or metric measurements. . MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Chapter 5 Discussion

In this randomised controlled experiment, I found that, compared to a no label control, all FOPLs significantly improved the ability of participants to correctly rank food and drink products according to healthiness. This was a consistent finding across all products and the three main outcomes, but the magnitude of the effects differed between FOPL groups. The largest effects were seen for N-S, then MTL, WL and lastly, PC. The global food score analysis indicated that N-S led to improved rankings in an average of two product categories (2.1), compared to less than two categories for MTL (1.7) and WL (1.4) (see Table 4). I also found that there was limited learning between baseline and follow-up ranking in the control group, indicating that the observed effects were likely due to the inclusion of the FOPL. The descriptive analyses examining correct ranking of products at follow-up and mean global food score showed that overall, the results were similar across participant groups. The main comparison of interest is N-S vs MTL. Across participant characteristics, N-S led to stable mean global food scores across education and income groups, whilst MTL led to a larger gradient with scores declining among lower education and income groups.

These findings are consistent with previous studies, where N-S out-performed other FOPLs, but the literature is limited given that N-S was introduced in 2017.[69, 70] Results were similar in a Canadian study that included the HSR label (N-S and HSR are both summary indicator labels), suggesting that overall, these labels are more effective than the MTL label.[71] However, the differences between MTL and either N-S or HSR were greater than the current study. As evidenced in these studies and the current study, FOPLs consistently improve the ability to rank products.[33, 59, 60] All of the FOPLs in this experiment were interpretive, in that they provide consumers with some judgement of the healthiness of the food. These labels typically perform better than non-interpretive labels and those that are simple and use aids (such as colour) have been found to be the easiest and most likely to be used.[33, 59, 150] The use of colour could have contributed to the larger effects seen for N-S and MTL compared with WL and PC. N-S is a summary indicator label, which provides an overall nutritional assessment, whereas MTL provides nutrient-specific assessments, but no overall judgement. Other studies have found that interpretive labels providing an overall summary of healthiness are more helpful for consumers than those not providing a summary, but can lead to health halo effects (improving health perceptions of both healthy and unhealthy products, although only influenced purchase intentions of healthy products).[33, 71]

The difference seen in this experiment between N-S and MTL appeared to be smaller than in previous studies. This could be due to the study design, as larger and more rigorously conducted trials can lead to smaller effects compared to smaller trials.[151] This could also be explained by participants being more familiar with the MTL, since the MTL has been in use in the UK since 2013 and understanding and knowledge have been shown to improve with familiarity with the label.[33, 81] Familiarity can impact on subjective understanding directly (whether consumers *think* they have understood correctly) and through subconscious perception (exposure to familiar concepts leads to subconscious activation).[37] Such factors could also have contributed to the larger effects seen for the MTL label than the WL and PC labels. Concerns over the accuracy of FOPL have been found to have a negative impact on use, and this could have reduced effects for less well-known labels such as N-S, WL, and PC. PC was least effective across the three outcomes; this is unsurprising, since endorsement labels are binary, meaning they provide limited information and two of the products did not qualify for any PC.

Differences between product categories showed that the yoghurt and breakfast cereal product categories had the greatest improvements in correct ranking from baseline to follow-up, compared to the other foods. This could be due to the design of the mock packaging, which is addressed in the limitations section of this Chapter. Alternatively, the FOPL effectiveness might differ depending on product type or consumer motivation for buying, such as hedonic pleasure.[50] Both of these products may also be perceived as "healthy" products and often feature health claims (e.g., "light", "low-fat"). Health claims (which were not included in the mock images) can bias opinions of food, which FOPLs have shown potential in reducing.[52] These factors may explain the difference in responses and highlight the potential effectiveness of mandatory labels in these seemingly confusing products.

To make use of FOPLs, shoppers first need to notice and engage with the label. This is easier if they like the label and its format.[37] Participants had high recall for all the labels, except the PC (no cake or crisp products qualified for PC, so this could explain the lower recall). Label recall and use is more likely if they like the label and the format; additionally, factors such as context (for example, time, shopping location, price, marketing) and previous purchase and product preference also influence label engagement which will be addressed in Chapter 6.[50]

FOPL effects were independent of SES, ethnicity, age, sex, household composition, food shopping responsibility, and current FOPL use. However, the descriptive analysis showed there was a tendency for participants with higher education and income to be able to rank products with greater accuracy than those with less education or lower income. The results from this experiment suggest some differences in responses to FOPL according to the level of education and income. N-S seemed to show the smallest difference between dichotomised education and income groups compared to other FOPL groups (especially marked for pizza, instant hot chocolate, cake, and crisps, less so for yoghurts and breakfast cereals). This finding is consistent with previous literature, which found the N-S label had the strongest positive impact on individuals with lower nutritional knowledge.[152] This could be due to N-S being a summary indicator label that provides composite information with less detail, compared to the information presented in MTL.

The findings presented in this Chapter support the use of FOPLs as a policy intervention to improve British consumers' ability to understand the healthiness of food products. These findings provide evidence that both the MTL and N-S FOPLs are likely to be most effective at improving knowledge. Whilst there was some evidence that N-S had the largest effects, the difference between MTL and N-S was relatively small and not significant for individual product categories. It is important that policies to improve food choices do not disproportionately affect individuals from more deprived or less educated backgrounds. These findings suggest that there are benefits for all individuals regardless of education level, but there were some indications that N-S was most effective for those with a lower education level. Caution is needed when interpreting these findings, since these were descriptive analyses. Knowledge can support people in making healthier choices, and evidence from Chapter 3 has shown that FOPLs have the potential to encourage healthier purchases.

The strengths of this analysis were that it was appropriately powered, had a pre-specified protocol and statistical analysis plan, and the addition of a control group. The sample of British adults was large and was comparable to the full NatCen panel and to the UK population.[143, 144] The primary outcome analysis was conducted blinded by an independent statistician. There are some limitations of this analysis. The analysis sample for global food score was limited to 44% (n = 1,976), it was pre-specified that only participants who had consumed or purchased the product within the last 12 months would be included. Due to the low level of instant hot chocolate consumption or purchase, this category was excluded from the global score. Sensitivity analyses showed that results were stable when all participants were considered. The PC label does not provide sufficient information to rank three products, as it is a binary label, and two food categories had no products that qualified for it, as would be expected for this particular label. This is indicative of real-world issues related to PC. To deal with this limitation, analysis of the correct ranking of the healthiest product is presented in Chapter 6. Broader limitations including study design, analysis and packaging issues will be discussed in Chapter 9.

Chapter 5 Conclusions

The results presented in this Chapter show the primary outcome analysis of the experiment outlined in Chapter 4, addressing research objectives 2-4. I found that MTL and NS were the most effective labels at improving understanding of products healthiness and the NS label was more stable across SES groups. What I would really like to know now is: do composite labels enable faster use; does speed of ranking differ by socioeconomic characteristics; how the perceptions of ranking ability and perceptions of the labels differ by FOPL scheme. The descriptive results of improvement in ranking ability by participant characteristics showed participants with higher education, who were white, females, higher income tended to have a greater proportion ranked correctly compared to those with lower education, non-white and lower income respectively. When examined by FOPL scheme there is evidence that N-S may be the most effective FOPL for participants with lower education and income. In the next Chapter, I will present the secondary outcomes of the experiment including the correct ranking of the healthiest product, speed of ranking and the perceptions of the ranking task and the labels. In Chapter 7, I present the latent class analysis, examining the experimental outcomes based on label motivation and engagement characteristics. In Chapter 8, I present the findings of a Patient and Public Involvement and Engagement session conducted with children and young people, where I presented these the results from this Chapter and discussed their views on the findings, FOPLs generally and UK policy options.

Chapter 6: Ability to Correctly Rank the Healthiest Product, Speed of Ranking, and Perception Outcomes

Chapter 6 Background

In Chapter 5, I presented the primary outcome findings, examining the ability of participants to rank all three products correctly according to their healthiness. I found that front of pack labels (FOPLs) were effective at helping participants to identify the healthiness of the products. Nutri-Score (N-S) performed best, followed closely by Multiple Traffic Light (MTL) and then Warning Label (WL). Positive Choice tick (PC) labels performed the worst, but all labels performed better than the no label control group. But it is still unknown which FOPLs consumers could use correctly the quickest, the perceptions of their ability to use the labels in the ranking task, perceptions of the FOPLs and if the finding for ranking only the healthiest product differed compared to ranking all three products, so in line with the study protocol I present these pre-specified outcomes in this Chapter. The correct ranking of the healthiest product is especially relevant for PC, as it is a binary label and therefore it is not possible to correctly rank three products (the primary outcome for the experiment). The ability to correctly understand the comparative healthiness of products is important, however other factors are also important, especially those that impact on actual use and therefore the likely effects of any FOPL policy.

The effectiveness of FOPLs at equipping consumers with information to guide healthier choices relies on the ability of consumers to accurately understand the information and then to use that information when making food and drink purchase and consumption decisions.[38] The theoretical framework on how FOPLs impact on consumer responses presented by Grunert and Wills, introduced in Chapter 1, outlines that the perceptions of FOPLs impact on use/label effectiveness, including seeing and perceiving the label, and if they think they understand the label (subjective understanding), if they like the label.[37] Other perception factors shown to impact actual use include familiarity, credibility and trust in the accuracy of the label (including who developed the label e.g., industry vs Government), the context (including time pressures, price, marketing, product claims), product preferences and previous purchase.[50, 53] Familiarity with MTL is assumed to be higher in this British sample, compared to the other FOPL schemes, and therefore perceptions will likely be positively impacted.

The impact of time on the likelihood of using labels when food shopping is evidenced in the literature, with label use being more likely if more time is available or there is less time pressure.[37, 42, 50, 53, 129] Observational studies from Uruguay and the UK found that on average shoppers take 22-28 seconds choosing a product, and only seven seconds from picking up a product to putting it in their basket.[38, 46] Time to pick a product also differed between product categories, with UK shoppers spending more time viewing ready meals (41 seconds) compared to carbonated soft drinks (23 seconds), showing the impact of context.[38] Findings from a qualitative study that discussed the use of nutrition information when shopping found consumers felt like they did not have time to look at the labels and that the use of colours were helpful so they did not need to look at the numbers and figures.[50] Evidence suggests that summary labels are easier for consumers to interpret compared to nutrient-specific labels, especially for at-risk populations.[49, 62] It is unclear if this interpretation is also quicker, in addition to being easier.

In this Chapter, I will present the secondary analyses for the randomised controlled experiment. I examined whether FOPLs were effective at improving participant understanding of the healthiest product; whether the time taken to complete the ranking tasks differed between FOPLs; participants' perceptions of the FOPLs. The main objectives were 1) to identify if FOPLs (MTL, N-S, WL, PC) improved ability to rank the healthiest product to a greater extent than no label control; 2) identify if FOPLs differ in the time taken for participants to complete ranking tasks; 3) descriptively examine participant perceptions of the FOPLs. Consistent with previous analyses, I directly compared N-S and MTL since evidence suggests these are the best performing labels. The healthiest product analysis aimed to deal with the limitations of the PC label, as it does not provide sufficient information to rank products due to it being a binary label.

Publication status: These results have been published in Nutrients and delivered as a briefing paper to policy makers: Packer, J., et al., Assessing the Effectiveness of Front of Pack Labels:

Findings from an Online Randomised-Controlled Experiment in a Representative British Sample. Nutrients, 2021. 13(3).

Chapter 6 Methods

The materials and methodology of this experiment have been detailed in Chapter 4. Full details of the study protocol are available on Open Science Framework (see footnote[2]).

Procedure Overview

Participants viewed images of three products and were asked to rank them in order of healthiness, from most to least healthy. Participants completed the ranking tasks once with no FOPLs (baseline ranking tasks) and then again with their assigned FOPL shown on the products (follow-up ranking tasks). A total of 12 ranking tasks were completed by each participant (six product categories by two ranking conditions: baseline and follow-up). Participants were able to zoom in on the image, but no other information was presented (including no back of pack information).

After the follow-up ranking tasks, to check labels had been noticed, participants were shown an image of their label condition and asked if they had seen the FOPL. They were then asked if they had used the FOPL to complete the rankings, how easy/difficult they found the labels to understand, if they wanted the labels on packaging in the UK, how helpful in choosing what to buy they would be, and views on how long it takes to use the labels. These questions were not applicable to the control group as they did not see any FOPLs.

The time taken to complete rankings were taken from the survey software, which recorded the length of time (seconds) spent on each web page (each ranking task was on a separate page). Participants were advised to click *next* as soon as they had completed each task. The healthiest product outcomes and speed of ranking outcomes were based on the ranking tasks. None of the FOPLs were introduced or explained to participants.

[2] https://osf.io/k9v2p/

Measures

Secondary Outcome Measures

The primary outcome of the study was the ability to correctly rank products according to healthiness, the results of which were presented in Chapter 5 and have been published.[153] The secondary outcomes of the experiment presented in this Chapter include:

Healthiest product ranking outcomes were examined in three ways:

- Ranking of the healthiest product (correct = 1, incorrect = 0) at baseline and follow-up
- Change in ranking from baseline to follow-up (improved = +1, no change = 0, worsened = -1)
- Change in healthiest product global food score, aggregated change in rankings for the five food products (score range was –5 to +5)

Speed of ranking:

The time taken on each page of the experiment (in seconds) was provided by NatCen from the UNICOM Intelligence software paradata (data that is collected about the survey process).

- Time to complete each ranking (seconds) at baseline and follow-up (by product category and summed)

Ranking perceptions:

- Confidence in ranking was assessed by asking participants how many products they thought they had ranked correctly after the baseline rankings and once again after follow-up rankings -dichotomised as ‘more confident’ (all/ most) vs ‘less confident’ (some/ none/ don’t know) for ease of interpretation
- Enough information was assessed by asking participants if they felt they had enough information to rank the product from most healthy to least healthy after each ranking task - recoded as ‘yes to all’, ‘no to all’ or ‘mixed response’ at baseline and follow-up

Label perceptions, specific to the participants label condition (control group were excluded):

- Saw the label in the ranking tasks – dichotomised as yes vs no/not sure

- Used the label in the ranking tasks – all/some/did not use
- Ease of label understanding – dichotomised as ‘easy’ (very easy/quite easy) vs ‘difficult’ (quite difficult/very difficult)
- View on implementation of label in UK – support for mandatory labelling (yes-all); support for voluntary labelling (yes-some); no support for labels (no-none)
- View on helpfulness of label for food shopping – dichotomised as ‘helpful’ (very helpful/quite helpful) vs ‘not helpful’ (not very helpful/not at all helpful)
- View on time to use label when shopping – quick enough/too long

Chapter 6 Statistical Analysis

The experiment was powered on the primary outcomes (presented in Chapter 5) so formal testing of secondary outcomes was limited. I limited formal statistical analyses to healthiest product outcomes and speed of ranking outcomes, focusing on the N-S and MTL comparison to avoid over-testing. For the healthiest food outcome analyses, I used the same three tests as for the primary analyses (see Chapter 5 for the primary outcome analysis for correct ranking of all three products or publication)[153]: multilevel log-binomial regression analysis to compare the proportion of correct rankings between baseline and follow-up; log-binomial regression to compare the change in rankings from baseline to follow-up; and multiple linear regression analysis to compare the change in the global food score from baseline to follow-up. For time taken to complete each ranking task, I cleaned the timing paradata provided by NatCen and merged multiple page visits by participants, if applicable. I also created an indicator for the device used by each participant (mobile or desktop or both – as participants were able to leave and re-join the survey). For the speed of ranking analyses, I log-transformed the data to normalise the distribution, then used multiple regression analysis to test the association between time taken to rank the products at follow-up, before back transforming for interpretation. For these four outcomes, I compared each FOPL group with no label control, and additionally compared MTL and N-S.

All models were adjusted for the five stratification factors used for randomisation (year of recruitment to panel, sex, age, government office region, and household income) and the following pre-specified covariates: ethnicity, highest education level, household composition, food shopping responsibility and current FOPL use. For product level analyses, I excluded participants if they self-reported not buying or consuming that product in the last 12 months, to ensure lack of familiarity of a product did not impact results (see Appendix E, Table E3 for descriptive analysis with no requirement to consume products). For global food score analysis, I only included participants who reported buying or consuming all five food products, so lack of familiarity with a product type did not impact the results, as was pre-specified in the protocol. For speed of ranking, participants were required to correctly rank the product under the label condition, so the results were comparable and people who were incorrect did not impact the median (see Table E4 in Appendix E for sensitivity analysis with no requirement to be correct at follow-up). Additionally for the speed of ranking outcome in the multiple regression analysis I adjusted for device used, as mobile users were found to be faster; see Table E1 in Appendix E for the proportion of participants and devices used).

Models were weighted using a weight provided by NatCen to account for non-response and to ensure findings were representative of the British adult population. Stata software (Release 16, StataCorp LLC., College Station, TX, USA) was used for all analyses and a significance level of 5% was used.[149] Output for healthiest product outcomes was presented as relative risk (RR) of linear regression coefficient, with 95% confidence intervals and output for speed of ranking was presented as relative mean (RM), i.e., the ratio of the mean outcome in one group relative to the mean outcome in another group. The speed of ranking results are also presented descriptively, by dichotomised education and income level, FOPL group and product category (with the requirement of being correct). Perceptions of the ranking task are also presented descriptively, including the proportion of participants who recalled seeing the FOPL in the ranking tasks, reporting they had enough information to rank the products (following all ranking tasks) and confidence that they ranked the products correctly (at baseline and follow-up). The associations between label perceptions and FOPL groups are presented descriptively. I did not formally test differences between labels. These questions were specific to the label conditions that participants were randomised to, therefore not applicable to the control group as they saw no labels.

Chapter 6 Results

Participant Characteristics

The participant numbers were the same as for the primary outcome analyses shown in Chapter 5 (see section titled 'Sample Characteristics') and will not be repeated here. There was a high proportion of participants who reported being very or quite interested in healthy eating (94%) and having some or a lot of knowledge about healthy eating (86%). Nearly half of the sample reported they were currently trying to lose weight (47%).

Healthiest Product Ranking Outcome

Figure 22 shows the number of participants who correctly ranked the healthiest product at baseline and follow-up, and the percentage change in correct ranking for each FOPL group, by product category (see Table E2 in Appendix E for the number and proportion of participants correctly ranking at baseline and follow-up). The associations between experimental group and correct ranking of healthiest product at follow-up, adjusted for covariates, are shown below in Table 26. The probability of participants correctly ranking the healthiest product at follow-up was significantly greater across all product categories in MTL, N-S and WL, compared to control. N-S had the highest RR of correctly ranking the healthiest product followed by MTL and WL. The PC group had mixed results, with significant differences for drink, yoghurt, and cereal, but not pizza, crisps, or cake (no products in the cake and crisps category were eligible for a PC). There were no significant differences between MTL and N-S in the direct comparison. I found that the proportion of participants correctly ranking the healthiest product at baseline were similar across the FOPL conditions.

Figure 22: Number of participants who correctly ranked the healthiest product at baseline and follow-up, by FOPL group and product category: (a) Pizza; (b) Drink; (c) Cake; (d) Crisps; (e) Yoghurt; (f) Cereal. †Cake and crisps categories had no products qualify for Positive Choice tick

Table 26: Multilevel log-binomial regression results - healthiest product ranked correctly (yes /no) at follow-up (adjusted for baseline rank), compared to control by FOPL group (adjusted for design effects and covariates)

	MTL v Control RR (95%CI)	**N-S v Control RR (95%CI)**	**WL v Control RR (95%CI)**	**PC v Control RR (95%CI)**	**N-S v MTL RR (95%CI)**
Pizza	1.12 (1.04, 1.20) *p* = 0.002	1.20 (1.12, 1.28) *p* < 0.001	1.09 (1.03, 1.16) *p* = 0.005	1.03 (0.97, 1.09) *p* = 0.29	1.07 (0.99, 1.15) *p* = 0.10
Drink	1.42 (1.26, 1.59) *p* < 0.001	1.57 (1.39, 1.78) *p* < 0.001	1.20 (1.08, 1.34) *p* = 0.001	1.22 (1.10, 1.34) *p* < 0.001	1.11 (0.97, 1.26) *p* = 0.12
Cake	1.20 (1.12, 1.29) *p* < 0.001	1.22 (1.14, 1.31) *p* < 0.001	1.19 (1.12, 1.26) *p* < 0.001	1.02† (0.97, 1.07) *p* = 0.45	1.02 (0.94, 1.10) *p* = 0.63
Crisps	1.26 (1.17, 1.35) *p* < 0.001	1.31 (1.22, 1.40) *p* < 0.001	1.20 (1.12, 1.28) *p* < 0.001	1.04† (0.97, 1.10) *p* = 0.26	1.04 (0.96, 1.13) *p* = 0.37
Yoghurt	4.89 (3.78, 6.33) *p* < 0.001	5.64 (4.22, 7.54) *p* < 0.001	3.81 (2.83, 5.12) *p* < .001	2.50 (1.89, 3.30) *p* < 0.001	1.15 (0.85, 1.57) *p* = 0.37
Cereal	1.95 (1.69, 2.24) *p* < 0.001	1.98 (1.71, 2.30) *p* < 0.001	1.76 (1.49, 2.06) *p* < 0.001	1.18 (1.03, 1.35) *p* = 0.018	1.02 (0.88, 1.17) *p* = 0.84

†Cake and crisps categories had no products qualify for Positive Choice tick; all analyses were adjusted for baseline ranking (correct/incorrect), stratification factors (year of recruitment to panel, sex, age, government office region, household income), and the following pre-specified covariates: ethnicity, highest education level, household composition, food shopping responsibility, and current FOPL use. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick; RR: Relative Risk; CI: Confidence Interval.

The associations between the FOPLs and an improved change score, adjusted for covariates, are shown in Table 27. The probability of participants improving their score from baseline to follow-up was significantly greater across all product categories in MTL, N-S and WL, compared to control. N-S led to the greatest probability of improved scores, followed by MTL and WL. PC had mixed results, with a greater probability of improved score for drinks, yoghurt, and cereal, but not for pizza, cake or crisps. The comparison between N-S and MTL showed no significant differences.

Table 27: Log-binomial regression results – relative risk that ranking of healthiest product improved (follow-up vs baseline) between FOPL group and control (adjusted for design factors and covariates)

	MTL v Control RR (95%CI)	N-S v Control RR (95%CI)	WL v Control RR (95%CI)	PC v Control RR (95%CI)	N-S v MTL RR (95%CI)
Pizza	2.91 (1.81, 4.68) *p* < .001	3.29 (2.05, 5.27) *p* < .001	2.57 (1.60, 4.15) *p* < .001	1.35 (0.81, 2.26) *p* = 0.24	1.13 (0.79, 1.61) *p* = 0.50
Drink	8.04 (3.83, 16.86) *p* < .001	10.68 (5.16, 22.09) *p* < .001	5.06 (2.37, 10.78) *p* < 0.001	3.98 (1.74, 9.09) *p* = 0.001	1.33 (0.95, 1.85) *p* = 0.09
Cake	8.59 (4.45, 16.60) *p* < .001	8.98 (4.70, 17.16) *p* < .001	7.57 (3.94, 14.57) *p* < 0.001	1.07† (0.41, 2.76) *p* = 0.89	1.05 (0.71, 1.54) *p* = 0.82
Crisps	4.86 (3.05, 7.72) *p* < .001	5.65 (3.60, 8.97) *p* < .001	3.73 (2.30, 6.03) *p* < .001	1.62† (0.87, 3.00) *p* = 0.13	1.16 (0.91, 1.49) *p* = 0.23
Yoghurt	18.13 (10.78, 30.49) *p* < .001	18.50 (10.99, 31.15) *p* < .001	13.07 (7.72, 22.11) *p* < .001	4.57 (2.62, 7.99) *p* < .001	1.02 (0.93, 1.12) *p* = 0.68
Cereal	5.20 (3.50, 7.73) *p* < .001	5.78 (3.90, 8.56) *p* < .001	4.73 (3.17, 7.08) *p* < .001	1.95 (1.26, 3.01) *p* = 0.003	1.11 (0.96, 1.29) *p* = 0.17

†Cake and crisps categories had no products qualify for Positive Choice tick; all analyses were adjusted for the five stratification factors (year of recruitment to panel, sex, age, government office region, household income) and the following pre-specified covariates: ethnicity, highest education level, household composition, food shopping. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick; RR: Relative Risk; CI: Confidence Interval.

The associations between an improved healthiest product global food score and FOPLs groups, adjusted for covariates, are shown in Table 28. I found that all FOPL groups were associated with a significant increase in the healthiest product global food score, compared to control. Comparison between N-S and MTL showed that N-S led to a greater improvement in the healthiest product global food score.

Table 28: Multiple regression analysis results- association of global food score between FOPL group and control (adjusted for design factors and covariates)

	MTL v Control RR (95%CI)	N-S v Control RR (95%CI)	WL v Control RR (95%CI)	PC v Control RR (95%CI)	N-S v MTL RR (95%CI)
Score (-5, +5) Regression (coefficients)	1.5 (1.3, 1.6) $p < 0.001$	1.8 (1.6, 1.9) $p < 0.001$	1.2 (1.0, 1.3) $p < 0.001$	0.3 (0.2, 0.5) $p < 0.001$	0.3 (0.2, 0.5) $p < 0.001$

Global food score was an aggregated score of correct ranking of the healthiest product in the five food products, range -5 to +5 (- indicates worsening and + indicates improvement); all analyses adjusted for stratification factors (year of recruitment to panel, sex, age, government office region, household income) and covariates: ethnicity, highest education level, household composition, food shopping responsibility, current FOPL use. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick; RR: Relative Risk; CI: Confidence Interval.

Speed of Ranking

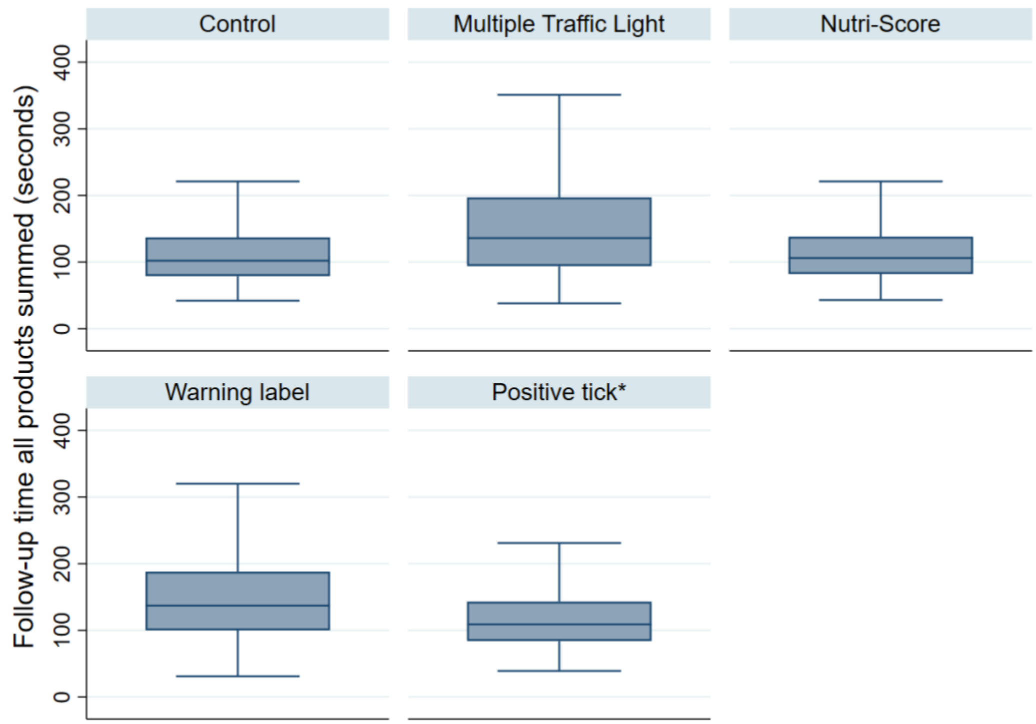
The median and interquartile ranges of the time to complete the baseline and follow-up rankings in seconds, by experimental group and product category, are presented in Table 29. The baseline median ranking times were similar within product categories and between groups. Follow-up rankings were faster compared to baseline rankings, for all categories and experimental groups. Overall, there was variation between the product categories at baseline (cake quickest, yoghurt and cereal longest), but little variation between product categories at follow-up. Comparing N-S and MTL, participants using N-S were fastest across all categories and participants using MTL had slower ranking times (please note that the MTL label contains more information than the other labels). See Table E3 in Appendix E for results without the need to consume the product. The median ranking times and interquartile ranges in seconds of the summed follow-up ranking times (seconds) by experimental condition are shown in Figure 23. This figure shows the distribution of the ranking data excluding outliers, see Figure E1 in Appendix E for the distribution including outliers.

Table 29: Summary of baseline and follow-up ranking times (seconds) by product category, overall and experimental group

	Control (*n*=913)		**MTL (*n*=907)**		**N-S (*n*=924)**		**WL (*n*=895)**		**PC (*n*=891)**		**Overall (*n*=4530)**	
	B	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**
	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)
Pizza	34 (24-50)	17 (12-25)	35 (24-51)	20 (14-33)	34 (26-52)	16 (12-24)	34 (24-50)	22 (15-33)	35 (24-49)	17 (12-26)	35 (25-50)	18 (13-28)
Drink	33 (23-47)	15 (11-22)	34 (23-49)	20 (14-31)	33 (24-47)	16 (12-23)	32 (22-48)	19 (13-27)	32 (23-47)	16 (11-23)	33 (23-48)	17 (12-25)
Cake	25 (18-38)	14 (9-20)	27 (19-38)	19 (13-32)	27 (19-40)	16 (12-23)	27 (18-38)	22 (14-35)	25 (18-37)	15 (10-23)	26 (18-38)	17 (12-26)
Crisps	30 (22-42)	15 (11-22)	29 (22-41)	20 (14-32)	30 (22-42)	17 (12-24)	30 (22-43)	21 (15-29)	29 (21-41)	18 (12-25)	30 (22-42)	18 (13-26)
Yoghurt	36 (25-53)	16 (11-24)	36 (25-52)	21 (14-32)	36 (25-53)	16 (12-23)	36 (25-52)	21 (15-30)	35 (23-51)	17 (12-25)	36 (25-52)	18 (12-27)
Cereal	35 (24-52)	17 (12-25)	36 (25-54)	21 (14-33)	36 (25-52)	17 (16-32)	36 (25-52)	23 (16-32)	34 (24-50)	18 (12-27)	36 (24-52)	19 (13-28)

B = baseline; FU = follow-up; IQR = interquartile range *Participants needed to have complete covariate information and buy/eat each product to be included in each product analysis.

Figure 23: Median and interquartile ranges of ranking times (seconds) of summed follow-up rankings by experimental condition. Excluding participants who did not have full co-variate information and outliers.



*Cake and crisps categories had no products qualify for Positive Choice tick; outliers not presented in the graph.

The associations between time to complete rankings of each product and the FOPL conditions, adjusted for covariates, are shown in Table 30, with higher RM indicating time to complete ranking was slower i.e., worse (see Table E4 in Appendix E for sensitivity analysis without the requirement for being correct at follow-up). By FOPL group, for pizza, I found that the time to complete ranking significantly increased by 25% for MTL and 28% for WL compared to no label control; no significant differences for N-S or PC compared to no-label control. The pattern was broadly consistent across the product categories. Comparison between N-S and MTL showed that the time to complete ranking decreased by 15-22% for participants in N-S group compared to MTL, across all product categories (all $p < 0.001$).

Table 30: Multiple regression analysis results- association between time taken to rank products at follow-up and FOPL group compared to control (adjusted for baseline ranking time, device used, design factors and covariates)

	MTL v Control RM (95%CI)	**N-S v Control RM (95%CI)**	**WL v Control RM (95%CI)**	**PC v Control RM (95%CI)**	**N-S v MTL RM (95%CI)**
Pizza	1.25 (1.17, 1.33) $p < 0.001$	1.00 (0.94, 1.06) $p = 0.96$	1.28 (1.20, 1.37) $p < 0.001$	1.05 (0.98, 1.12) $p = 0.16$	0.80 (0.75, 0.85) P < 0.001
Drink	1.28 (1.18, 1.38) $p < 0.001$	1.08 (1.00, 1.17) $p = 0.056$	1.28 (1.18, 1.39) $p < 0.001$	1.10 (1.01, 1.19) $p = 0.035$	0.85 (0.79, 0.91) $p < 0.001$
Cake	1.43 (1.35, 1.53) $p < 0.001$	1.18 (1.10, 1.26) $p < 0.001$	1.61 (1.51, 1.72) $p < 0.001$	1.09 (1.02, 1.17) $p = 0.01$	0.82 (0.78, 0.86) $p < 0.001$
Crisps	1.39 (1.31, 1.46) $p < 0.001$	1.13 (1.07, 1.19) $p < 0.001$	1.30 (1.23, 1.38) $p < 0.001$	1.12 (1.05, 1.19) $p < 0.001$	0.82 (0.78, 0.86) $p < 0.001$
Yoghurt	1.40 (1.26, 1.56) $p < 0.001$	1.10 (0.99, 1.22) $p = 0.08$	1.44 (1.29, 1.60) $p < 0.001$	1.16 (1.01, 1.33) $p = 0.10$	0.78 (0.74, 0.83) $p < 0.001$
Cereal	1.31 (1.22, 1.41) $p < 0.001$	1.02 (0.95, 1.10) $p = 0.08$	1.41 (1.31, 1.52) $p < 0.001$	1.11 (1.02, 1.21) $p = 0.03$	0.78 (0.74, 0.82) $p < 0.001$

All analyses were adjusted for the five stratification factors (year of recruitment to panel, sex, age, government office region, household income) and the following pre-specified covariates: ethnicity, highest education level, household composition, food shopping responsibility, current FOPL use, baseline ranking time and device used. Participants needed to have complete covariate information and buy/eat the specific product to be included and needed to correctly rank the products. To deal with outliers the data was log-transformed for analyses and then back transformed for interpretation. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning Label; PC: Positive Choice tick; RM: Relative Mean; CI: Confidence Interval.

Descriptive Analysis– Speed of Ranking by Participant SES Characteristics

Descriptive analyses of the speed of ranking by experimental group and participant SES characteristics (dichotomised education level and equivalised household income, for consistency with other analyses presented in this thesis) are shown in Table 31 and Table 32 (see Table E5 and Table E6 in Appendix E for sensitivity analyses without the need to be correct). Overall, the higher education and income groups were faster at completing the ranking tasks across all of the product categories, compared to the lower education and income groups. When comparing N-S to MTL, participants in N-S regardless of education or income level, were faster than all participants in MTL (ranking time with N-S faster in lower SES groups, compared to MTL ranking time in higher SES groups). The interquartile ranges for the ranking times in the N-S condition were the smallest, indicating greater stability.

Table 31: Median and interquartile ranges of ranking times (seconds) at follow-up, by product category, FOPL group and education level (higher/lower)

Education	**FOPL group**				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Higher	18 (13-30)	16 (12-22)	22 (15-34)	17 (12-23)	16 (11-23)
Lower	22 (15-39)	17 (13-26)	25 (17-35)	19 (13-30)	19 (13-26)
Drink					
Higher	18 (13-26)	15 (11-21)	18 (13-25)	15 (11-21)	14 (10-20)
Lower	22 (15-32)	18 (13-26)	21 (15-29)	17 (11-27)	17 (11-24)
Cake					
Higher	18 (13-31)	15 (12-21)	23 (15-36)	14 (10-22)	13 (9-18)
Lower	21 (14-34)	16 (12-25)	23 (15-36)	17 (11-24)	15 (10-23)
Crisps					
Higher	19 (14-29)	16 (12-23)	21 (15-28)	18 (12-24)	15 (10-21)
Lower	23 (16-36)	19 (14-27)	23 (16-31)	19 (13-27)	17 (12-23)
Yoghurt					
Higher	19 (14-30)	15 (12-21)	21 (15-29)	16 (12-26)	16 (11-21)
Lower	25 (17-36)	18 (13-25)	25 (17-36)	20 (15-29)	15 (12-19)
Breakfast cereal					
Higher	19 (14-31)	16 (12-22)	22 (15-31)	17 (12-26)	16 (11-22)
Lower	26 (17-40)	18 (13-25)	25 (18-37)	21 (13-32)	18 (12-24)

Education level dichotomised as A levels and above vs. below A levels, higher or lower education, respectively. Participants needed to have complete covariate information correctly rank the products, specific to the product category.

Table 32: Median and interquartile ranges of ranking times (seconds) at follow-up, by product category, FOPL group and income category (higher/lower)

Income	FOPL group				
	MTL	N-S	WL	PC	Control
Pizza					
Higher	18 (13-30)	15 (11-20)	21 (15-30)	17 (13-24)	16 (10-23)
Lower	20 (14-33)	17 (13-24)	24 (16-36)	17 (12-26)	17 (12-25)
Drink					
Higher	17 (13-28)	14 (11-21)	18 (13-24)	15 (11-21)	14 (10-20)
Lower	20 (14-27)	17 (12-23)	20 (14-28)	15 (11-22)	15 (11-23)
Cake					
Higher	18 (13-31)	14 (12-20)	23 (15-36)	15 (10-23)	14 (9-21)
Lower	20 (13-33)	16 (12-24)	24 (15-37)	16 (10-23)	13 (9-19)
Crisps					
Higher	20 (14-29)	16 (12-22)	20 (14-28)	18 (13-25)	15 (11-21)
Lower	21 (15-32)	17 (13-25)	22 (16-30)	18 (12-25)	16 (11-22)
Yoghurt					
Higher	19 (14-29)	14 (12-19)	20 (15-29)	16 (12-23)	16 (11-23)
Lower	23 (15-33)	17 (12-23)	22 (17-33)	18 (13-29)	16 (11-20)
Breakfast cereal					
Higher	19 (14-30)	15 (11-21)	21 (15-31)	16 (12-26)	15 (11-22)
Lower	22 (15-34)	17 (13-24)	25 (17-34)	19 (13-31)	17 (12-22)

Equivalised income dichotomised as 2000 and above vs. below 2000, higher or lower income, respectively. Participants needed to have complete covariate information correctly rank the products, specific to the product category.

Ranking perceptions

Participants' perceptions of the ranking tasks are presented in Table 33, including the number who: reported seeing the FOPLs; reported having enough information to rank the products; and reported having confidence in their ranking being correct. The experiment was not powered to formally test statistical differences in ranking perceptions. The number of participants who reported needing more information to rank each of the products at follow-up, overall, across each of the products, the PC and control groups had the lowest proportion of participants reporting having enough information, followed by WL, N-S and MTL the highest. When categorised as responding 'yes to all', 'no to all' or 'mixed response' at baseline and follow-up, the results indicated that responses at baseline were similar across all experimental groups. At follow-up, the responses were varied between experimental groups, with MTL having the highest proportion of participants responding 'yes' to all product categories, followed by N-S and WL, with similarly low proportions for PC and control. Responses to confidence in rankings at baseline appeared similar across all experimental groups but differed at follow-up. Compared to the control group, the FOPLs combined led to greater confidence at follow-up (52% vs 16%). Confidence in ranking varied between the FOPL groups; MTL led to the highest proportion of participants who were more confident in the follow-up rankings (81%), followed by N-S (68%), WL (41%) and then PC. (19%). PC had similarly low levels of confidence to the no label control (16%).

Table 33: Results of post-ranking questions across the study sample by FOPL group (*n* = 4,530), including reporting seeing FOPLs, having enough information to rank products at baseline and follow-up, and confidence in ranking at baseline and follow-up

	FOPL group					
	Control (*n*=913)	**MTL** (*n*=907)	**N-S** (*n*=924)	**WL** (*n*=895)	**PC** (*n*=891)	**Overall** (*n*=4,530)
Enough information to rank categorised						
Baseline						
Yes (across all)	56 (6)	50 (6)	59 (6)	52 (6)	51 (6)	268 (6)
No (across all)	648 (71)	631 (70)	648 (71)	637 (71)	657 (74)	3,221 (71)
Mixed response[v]	209 (23)	226 (25)	217 (23)	206 (23)	183 (21)	1,041 (23)
Follow-up						
Yes (across all)	70 (8)	695 (77)	475 (51)	201 (22.5)	64 (7)	1,505 (33)
No (across all)	694 (76)	95 (10.5)	313 (34)	414 (46)	704 (79)	2,220 (49)
Mixed response[v]	149 (16)	117 (13)	136 (15)	280 (31)	123 (14)	805 (18)
Confidence in ranking						
Baseline						
More confident	131 (14)	133 (15)	136 (15)	123 (14)	122 (14)	514 (14)
Less confident	782 (86)	774 (85)	788 (85)	772 (86)	769 (86)	3,103 (86)
Follow-up						
More confident	149 (16)	732 (81)	631 (68)	363 (41)	168 (19)	1,894 (52)
Less confident	764 (84)	175 (19)	293 (32)	532 (59)	723 (81)	1,723 (48)

[i]Control group not applicable, *n* = 3,598*; [ii] Does not equal 100% due to: 1 Don't know response (n = 1, 0.1%); [iii] or 2 Don't know responses (n = 2, 0.6%). MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick. [iv] For having enough information to rank, needed to consume each specific product to be included. [v] Mixed response includes yes/ no/ don't know (14 don't know responses total). Needed to have full co-variate information to be included. Confidence dichotomised as more confident (all/most) vs less confident (some/none/don't know). All FOPL = MTL, N-S, WL, PC.

Label Perceptions – Descriptive Analyses of Perceptions

Descriptive analyses of the label perceptions, by FOPL group, are shown in Table 34. Overall, 74% reported seeing the FOPLs that they were randomised to, this differed by FOPL condition. N-S (87%) had the highest recall rate followed by WL (78%), MTL (77%) and PC (55%). For using the label for all ranking tasks, the overall proportion was roughly half (52%), N-S had the highest proportion (75%) followed by MTL (69%), WL (54%) and PC (10%). Overall, the perceptions of the labels differed greatly between FOPL groups. The majority of participants favourably viewed labels and supported mandatory labelling in the UK (73%). The N-S and MTL groups had the highest proportion of participants who perceived their respective labels favourably. N-S had the highest proportion of participants who reported: the label was easy to understand; and quick enough to use. MTL had the highest proportion of participants who reported perceiving the label to be helpful in making purchasing decisions; and supported having labels on all food and drink products in the UK; and otherwise, N-S was most favourably perceived, followed by MTL for the other perceptions.

Table 34: Proportion of participant FOPL perception outcomes, by FOPL group

	MTL (*n*=907) *n* (%)	N-S (*n*=924) *n* (%)	WL (*n*=895) *n* (%)	PC (*n*=891) *n* (%)	Overall (n=3,617) *n* (%)
Saw label					
Yes	697 (77)	807 (87)	698 (78)	490 (55)	2,692 (74)
No/ not sure	210 (23)	117 (13)	197 (22)	401 (45)	925 (26)
Used label					
All	627 (69)	697 (75)	479 (54)	88 (10)	1,891 (52)
Some	51 (6)	85 (9)	190 (21)	214 (24)	540 (15)
Did not use	19 (2)	25 (3)	29 (3)	186 (21)	259 (7)
Not applicable	210 (23)	117 (13)	197 (22)	401 (45)	925 (26)
Understanding labels					
Easy	669 (74)	717 (78)	601 (67)	273 (31)	2,260 (62)
Difficult	28 (3)	89 (10)	97 (11)	217 (24)	431 (12)
Not applicable	210 (23)	117 (13)^	197 (22)	401 (45)	925 (26) ^
Label helpfulness					
Helpful	885 (98)	845 (91)	822 (92)	660 (74)	3,212 (89)
Not helpful	22 (2)	79 (9)	73 (8)	230 (26)^	404 (11)
Time to use label					
Quick enough	846 (93)	902 (98)	816 (91)	821 (92)	3,385 (94)
Too long	60 (7)	21 (2)	77 (9)	67 (8)	225 (6)
Labels in UK					
Yes- all	813 (90)	718 (78)	667 (75)	451 (51)	2,649 (73)
Yes- some	81 (9)	150 (16)	174 (19)	289 (32)	694 (19)
No- none	13 (1)	56 (6)	54 (6)	148 (17)*	271 (7)

*3 don't know responses ^1 don't know response. Control group not included as they were not asked these questions. Needed to have full co-variate information to be included. Helpful dichotomised as helpful (very helpful, quite helpful) vs not helpful (not very helpful, not at all helpful). MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Chapter 6 Discussion

In this analysis of the secondary outcomes of the experiment, I found that N-S and MTL performed best across a range of outcomes, including the correct ranking of the healthiest product and speed of ranking. The correct ranking of the healthiest product analysis showed N-S, MTL and WL performed the best, and PC performed worst in the ranking tasks; in terms of timing, analysis showed that participants using N-S ranked products significantly faster than those using MTL. The descriptive analysis of the ranking tasks perceptions showed that a higher proportion of participants used N-S and MTL for all of the ranking tasks; and descriptive analysis of label perceptions showed that FOPLs were perceived favourably overall, especially by participants who used N-S and MTL. This was consistent with the primary outcomes presented in Chapter 5, that N-S then MTL, led to the largest improvements in the ability to correctly rank products.

These findings are broadly consistent with previous research. Results from experimental ranking studies have shown that N-S performs the best in ranking tasks[69, 70] and a meta-analysis found summary indicator labels are the most effective at helping consumers identify the healthiest products.[59] In this Chapter, I conducted the healthiest product analysis, to supplement the primary outcome analysis of ranking all three correctly (presented in Chapter 5). This was to allow a fairer comparison for the binary PC label (except cake and crisps, which did not qualify) as it is not possible to differentiate between three products. I found that the results across the three outcomes were very similar to the primary analyses and that although PC performed better, it was still the worst performing label. For time needed to use labels, no studies have directly compared N-S and MTL to my knowledge, but experimental studies conducted in Uruguay and Mexico found that participants using MTL and WL were faster at completing tasks compared to GDA.[101, 154] I found that N-S was significantly faster to use than MTL, for the follow-up ranking task, and crucially also led to the highest proportion of correct rankings in primary analyses, compared to the other experimental groups. This is interesting as MTL is the voluntary label currently used in the UK and is therefore likely to be familiar to most participants. N-S provides an overall summary of product healthiness, compared to MTL which contains nutrient-specific information. This additional information may be useful for people with nutrient specific dietary needs, who need to assess different nutrients rather than overall healthiness; but may be unnecessary for most consumers. Time has been identified as a factor that impacts upon actual label use, so the finding that N-S is fastest to use while also leading to the highest proportion of correctly ranked products (as seen in Chapter 5), suggests N-S may be a more effective label than the other labels tested.[42, 50, 51, 53] The exploratory descriptive analyses examining time taken to correctly rank products by dichotomised income and education level, indicated that overall, the higher education and income groups were faster at correctly completing the rankings. But, when comparing N-S to MTL, all participants in the N-S condition (both lower and higher SES groups) were faster than the higher SES participants in the MTL condition; and the interquartile ranges were narrower, indicating greater stability. Overall, these results suggest that N-S may be more suitable for use across the population.

The ranking task in this study is not representative of how long the tested labels would take to use in real decision-making scenarios with additional factors such as cost, preference, and product claims, but I believe these factors are likely to broadly consistent regardless of FOPL used, therefore supporting N-S as the most appropriate FOPL based on these factors. Indeed, research indicates that people spend less time making purchasing decisions than the median time shown to complete the ranking tasks and identify the healthiness of the products. Observational studies show consumers spend less than seven seconds from picking up a product to putting it in their basket and less than 30 seconds in-aisle choosing a product and, therefore the additional time burden of using a FOPL, especially MTL and WL is less likely than N-S.[38, 46]

Overall, 75% of participants recalled seeing the FOPL that they were randomised to, N-S had the highest recall and PC the lowest. Recall of labels is likely to be important, as attention drawn by labels and the perception of labels is a key factor in engagement, as highlighted in the Grunert and Wills conceptual framework.[37] In a large-scale randomised controlled trial that tested four FOPLs on products in real supermarkets (N-S, MTL, Reference Intakes and a warning label variant), they found that the N-S drew the most attention and was also the most effective label in improving the nutritional quality of purchases.[155] Participants in the MTL group scored highest for reporting they had enough information to rank, and confidence in their ranking, despite N-S group generally performing better. This finding highlights the difference between objective understanding (outcomes of the ranking task) and subjective understanding (confidence that they ranked the products correctly), where N-S resulted in highest objective but not highest subjective understanding. If a summary indicator label is chosen to replace MTL, improving the subjective understanding and confidence in using the label would be important, as this would be a new FOPL and consumers may be more hesitant to use it if they do not feel they understand it correctly or are not confident using it. The latent class analysis presented in Chapter 7 further explores these factors, by looking at baseline food habits, motivation and engagement with labels and how this relates to sociodemographic characteristics and the experiment outcomes. FOPLs overall were perceived as easy to understand, helpful in choosing what to buy, and quick enough to use. N-S and MTL label had similarly favourable responses, followed by WL and then PC. This mirrored the primary outcome findings in Chapter 5. N-S had the highest proportion of participants who said the label would be 'quick enough' to use, thus supporting the speed of ranking findings. The majority of participants supported mandatory labelling in the UK, responding that they would like to see labels on all food and drink products in the UK, especially participants in the MTL and N-S groups. This is discussed further in Chapter 8, during the PPIE session. As previously stated, none of the labels were explained to participants and I assume that prior exposure to labels other than MTL (in voluntary use in UK) was low; therefore, I expected low subjective understanding of these labels and lower trust (of the underlying logic and who developed the FOPL e.g., industry vs government). Both N-S and PC are abstract labels that may need some explanation, contrary to MTL and WL which are self-explanatory. Due to this, the positive perceptions of N-S and MTL indicate both summary indicator and nutrient specific labels may be liked by British consumers. As discussed, MTL is the FOPL currently used in the UK on a voluntary basis, so the understanding and use of this label is assumed to be higher compared to the other labels. Prior understanding and familiarity combined may have negatively impacted on the confidence in using the other labels, compared to MTL.[37, 57] Other experimental studies testing the perceptions of FOPLs (including MTL and N-S) following use in ranking tasks found respondents generally viewed labels favourably and strongly supported mandatory labelling across all label conditions.[156-158] They found that participants who viewed MTL responded most favourably across 12 countries,[156] but found no difference in replicate studies in Belgium and Netherlands.[157, 158] A large cross-sectional study in France found N-S was the most favourably perceived compared to MTL, RI and a simplified nutrition labelling system, including the perception that it was the quickest label to process.[73] This study was conducted in June 2016, before N-S was implemented in 2017 and crucially appealed to participants with lower nutritional adherence, indicating it appeals to the population groups that are the most important to reach in order to reduce health inequalities.

The secondary outcomes of timing and perception presented in this Chapter relate more closely to actual use of FOPL, although still within a highly controlled experiment, as opposed to the ability of FOPLs to improve understanding of product healthiness. For FOPLs to be effective, they need to easily communicate nutritional information about the product, but also to be useable, especially for people with lower education or nutritional knowledge. This is vital, as a main consideration for these policies is to reduce and not widen health inequalities in obesity and for FOPLs to be effective for all population groups. The descriptive results presented in Chapter 5, indicate the N-S label may work better across all income groups and education groups, as the global food score results were more stable, compared to MTL and WL. These results, together with the timing and perception outcomes, may imply that summary indicators with composite scores may be easier to understand and use. This is intuitive and supported in the literature, with simpler and directive labels requiring less processing, which indicates they could be more useful for consumers, regardless of SES, age, literacy, etc.[49, 159] The main aim for FOPLs, as outlined in the WHO guidance, is to provide convenient and easy to understand information to assist the selection of healthier products.[160] For nutrient-specific or warning labels, it is hard to make judgements on which nutrient is more important to avoid, this requires additional processing which is time consuming, and also potentially alienating or overwhelming.

Despite being a high-quality experimental study, there were limitations of the secondary analyses. The study was powered on the primary aims, so I limited statistical testing of the secondary aims. A further limitation of the PC label is that it did not qualify for use in two of the food categories (crisps, cakes), which is reflected in the responses to participants in the PC group who reported seeing the label (lower compared to the other FOPL groups). This is reflective of the real-world application and limitations of the PC label, as many products would not qualify and therefore would not have a label or provide any information to consumers. Data from the software was used for the timing outcomes and participants were given the following instructions: “we would like to know how long it takes for you to rank the different foods, so please click ‘next’ as soon as you have completed each task.” These directions were provided to make participants aware that this data would be collected but with no intention to rush them, as timing was not the primary outcome; therefore, care needs to be taken with the interpretation of the findings. Participants only answered perception questions about the FOPL they used, therefore it is not possible to compare preferences between FOPLs. This was partly to reduce the time burden on participants but also meant that the responses to these questions were based on actual use of the label. The perception questions followed the use of the FOPLs in a highly controlled experiment i.e., not actual use when shopping.

Chapter 6 Conclusions

The findings were consistent with the primary results presented in Chapter 5, showing that across the outcomes N-S performed best, closely followed by MTL and then WL and PC. All FOPLs were better than the no-label control. The findings presented in this Chapter provide support for FOPLs, with evidence that they improved participants’ ability to identify the healthiest product, improved their confidence in understanding the healthiness of products and were perceived favourably. This builds on the evidence presented in the Chapter 2 and 3 reviews that assessed the impact on objective understanding and objectively measured consumer behaviours. In the next Chapter I will present the latent class analysis I conducted, which further explores relationship between participant food habit and label engagement and motivation characteristics, sociodemographic variables and the experiment outcomes.

Chapter 7: Latent Class Analysis

Chapter 7 Background

In this Chapter, I will present the methods and findings of the latent class analysis. In this analysis I was able to group participants according to their baseline food shopping habits and motivations and then examine the difference between the groups by sociodemographic characteristics and the experiment outcomes. This analysis builds on the results of the experiment presented in Chapters 5 and 6, and aims to understand more about the indicators/factors that may impact on real-world FOPL engagement.

The aim of FOPLs is to help consumers make informed and healthier choices, but this is dependent on their engagement with labels when buying or eating food.[32, 33, 38] Correctly understanding FOPLs to determine product healthiness is a crucial factor, and my experiment results presented in Chapter 5 are directly relevant to this. But evidence suggests that label engagement is impacted by many factors, as discussed in Chapter 1, including individual, social and environmental contextual factors (such as motivations for eating/buying food, sociodemographic characteristics, household factors).[33] In Chapter 6, I presented the secondary outcomes of the experiment that assessed other factors known to be associated with engagement, including time taken to use labels, ranking perceptions (including confidence in using the labels) and label perceptions (including liking of the label).

For FOPLs to be an equitable public health intervention, it is important that the chosen label is suitable for use across the entire UK population and does not widen inequalities. To assist this, it is important to identify which factors may impact (positively or negatively) on engagement and thereby understand which groups may benefit from focused public education or media campaigns to improve the real-world effectiveness. In Chapter 4, I examined the ability of participants to correctly rank all products according to product healthiness and found that all FOPLs improved the ability. I then conducted descriptive analyses examining ranking ability by participant groups, which indicated that there were some differences in ability. This was conducted descriptively as the experiment was not powered to examine this and to avoid errors from multiple testing. Therefore, to further understand the experiment results and identify underlying subgroups in the population I used latent class analysis (LCA). LCA is described as a ‘person-centred’ approach, which identifies groups of similar individuals that are distinctly different to the other groups (called classes). [161-163] A strength of LCA is that it is a finite mixture model, which reduces the total number of patterns by creating sub-groups that are representative of the main patterns in the dataset, thereby reducing multiple statistical testing. It also enables the possible identification of clusters of predictors and the examination of higher-order interactions. This is method is favourable to standard regression methods that are commonly used to test a range of predictive factors, which can be problematic due to the increased risk of chance findings due to multiple statistical testing (Type 1 error) and low statistical power.[163] LCA typically uses cross-sectional data (measured at one time point), therefore causality cannot be determined.

The aims of this study were to 1) identify if participants can be grouped into classes based on food habit and motivation characteristics, 2) identify if the classes differ according to sociodemographic characteristics and, 3) assess if outcomes from the online experiment that tested effectiveness of FOPLs (i.e., ability to understand product healthiness, ranking perceptions and label perceptions) differ by predicted class.

I hypothesised that participants who are more engaged by and engaged with labels and food shopping as assessed with the baseline measures will be better at using FOPLs in ranking tasks (correctly rank more products according to healthiness). I further hypothesized that for ranking ability, the difference between engaged and less engaged participants using interpretative labels (N-S and MTL) will be less than non-interpretive (PC). This is based on research that indicates interpretive labels are easier to engage with.[33, 44, 60, 61]

Publication status: These results have been presented at the European Congress on Obesity in Dublin. Packer, J.; Russell, S.; Viner, R (2023) PO2.084 Front of pack labels, who engages with them and why: a latent class analysis from an online experiment testing label effectiveness Obes Facts (2023) 16 (Suppl. 1): 1–351. https://doi.org/10.1159/000530456

Chapter 7 Methods

Overview

The study sample were participants from the online randomised controlled experiment, as specified in Chapter 4 (*n* = 4,863) (see Appendix F for analyses conducted in the *n* = 4,530 sample, as used in previous analyses). The measures used in this study are described fully in Chapter 4, including the baseline measures food habits (e.g., use/ influence of FOPLs), motivations to use FOPLs (e.g., to eat healthily/ trying to lose weight); household factors (e.g., children in household); sociodemographic characteristics; and experimental outcome measures (ranking ability, label perceptions, ranking perceptions).

Chapter 7 Statistical Analysis

Latent class analysis was used to identify possible subgroups within the study population related to food habits and motivations.[161, 162] A three-step approach, as illustrated in Figure 24, was used to 1) identify the latent classes reflecting motivation with food choices and engagement with labelling, using baseline measures of food habits, motivations to use FOPLs and household factors, 2) assign participants into latent classes based on the highest posterior probability (the probability of the model accurately predicting class membership for each individual), 3) explore associations between class membership and sociodemographic variables and the experiment outcomes (ability to rank and ranking/label perception outcomes).[162, 164, 165] For step one, seven categorical indicator variables measuring baseline food shopping habits (shopping responsibility, use and influence of labels), motivations to use labels (nutritional knowledge, interest in healthy eating and trying to lose weight), and household factors (children in the household) were used to build the latent class model, shown in Table 35. These seven variables were chosen as they were collected across all participants, regardless of their allocated experimental group, and were all baseline behavioural measures, so data was available for all participants and the measures were assessing the same time point. The indicator variables included both binary and ordinal categorical variables, using the 'gsem' command in Stata.[166] Goodness of fit statistics were calculated to determine the optimal number of classes, including: Bayesian Information Criterion (BIC) coefficient, Akaike Information Criteria (AIC) coefficient, entropy values and Lo-Mendell likelihood ratio test (LMT-

LRT) coefficient. The LMT-LRT test examines the fit of a model compared to a model with one less class, (e.g., performing a significance test for the three-class model compared to a two-class model). The best-fitting model was chosen using a combination of low BIC, significant LMT-LRT, high entropy value, and interpretable and appropriate class size (i.e., no group less than 5% of the sample).[167, 168] For step two, participants were assigned to a class (i.e., membership), using the highest posterior probabilities. Chi-square analysis was used to test differences in the frequencies of the indicator variables between predicted classes and a trend test was conducted to see the change across the classes (see Table F1 and Table F2 in Appendix F). For step three, the associations between assigned class membership and external variables were assessed, including proximal sociodemographic variables (e.g., education, income, BMI) and distal experiment outcome variables (e.g., ability to correctly rank products at baseline and follow-up, global food score, speed of ranking and perception outcomes).[169] For the proximal sociodemographic analysis, chi-square analysis and multinomial logistic regression were used, with predicted class membership as the outcome variable (a similar approach to James et al.)[165] and relative risk ratio (RRR) as the output. A RRR greater than 1.0 indicates an increased risk for the group compared to the reference group, and less than 1.0 indicates a decreased risk. For the distal outcome analyses, bivariate analyses for each dependent variable were conducted to describe the association between predicted class membership and ranking outcomes (chi-square for categorical, analysis of variance (ANOVA) for means, or nonparametric equality-of-medians test for median)[162, 170, 171] and multinomial logistic regression for distal outcomes.[164] Analyses examining mean global food score and label perceptions (e.g., quick enough to use, view on mandatory labelling) by predicted class and experimental group were also conducted, using either one-way ANOVA or chi-square analysis, depending on the variable. No formal corrections of p-values were applied to adjust for testing of multiple outcomes. However, the results of all statistical tests conducted are reported so that an informal adjustment of p-values can be performed while reviewing the data.[172] Results at follow-up for ranking perceptions (having enough information to rank and confidence in ranking) are presented descriptively by predicted class and experimental group. Class membership uncertainty was not taken into account, as entropy was greater than 0.8 and this has been found not to inflate or deflate standard error.[164] I conducted sensitivity analysis in Appendix F (Table F6-F14) using the analysis sample from Chapter 5 and 6 who had complete covariates variables (*n* = 4530).

Figure 24: The three-step latent class analysis method

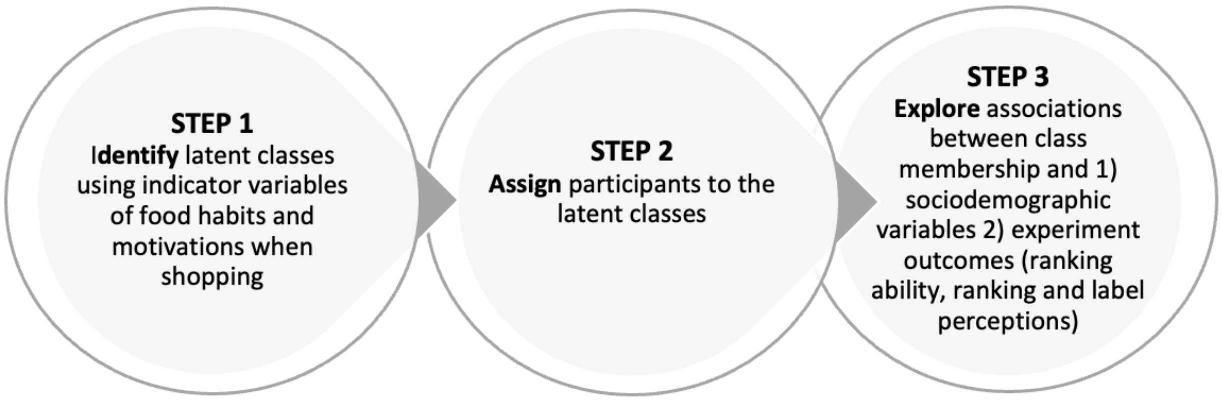


Table 35: Indicator variables

Type	Variable	Categorical coding
Food shopping habits	Shopping responsibility	Binary: yes or no
	Currently use labels	Ordinal: very/quite often, occasionally, rarely
	Influenced by labels	Ordinal: very/quite often, occasionally, rarely
Motivations to use labels	Nutritional knowledge	Ordinal: a lot, some, a little/none
	Interest in healthy eating	Ordinal: very, quite, not very/not at all
	Trying to lose weight	Ordinal: yes, no, prefer not to say
Household factors	Children in the household	Binary: yes or no

Chapter 7 Results

Fit Statistics and Model Selection – Step One

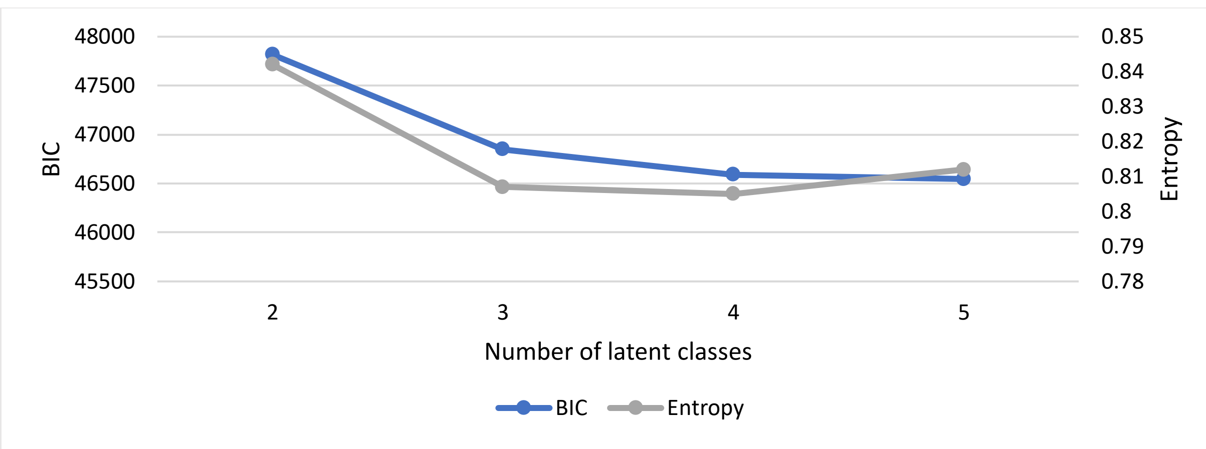
Table 36 displays the summary of goodness of fit results for models one through five classes, and Figure 25 displays a graph with the BIC and entropy values for classes two through five. The 'elbow' in Figure 2 occurs when there are three classes, indicating that after this point the decrease in BIC becomes notably smaller and there is little difference in entropy values. Therefore, the three-class model was chosen due to the combination of a significant LMT-LRT ($p<0.0001$); relatively low BIC (46847.85 vs 46589.65); greater entropy value compared to four classes (0.807 vs 0.805); and latent class proportions with appropriate and easily interpretable sample sizes.

Table 36: Goodness of fit statistics for the one- through five-class models, with delta percent change

Number of classes	p-value (LMT - LRT)	AIC	BIC	Entropy
1	0.000	52435.84	52513.71	-
2	0.000	47654.24 ∇ (-9.1%)	47816.48 ∇ (-8.9%)	0.842
3	0.000	46601.25 ∇ (-2.2%)	46847.85 ∇ (-2.0%)	0.807 ∇ (-4.2%)
4	0.820	46258.69 ∇ (-0.7%)	46589.65 ∇ (-0.6%)	0.805 ∇ (-0.2%)
5	1.000	46130.65 ∇ (-0.3%)	46545.97 ∇ (-0.1%)	0.812 Δ (+0.7%)

LMT-LRT = Lo-Mendell Rubin - Likelihood ratio test. AIC = Akaike Information Criteria. BIC = Bayesian Information Criteria. Note: LMT-LRT test p-values relate to the comparison between *k* number of classes vs. *k*-1 classes for *k*=2,3,4,5.

Figure 25: Graph of BIC and entropy values of two- through five-class models



BIC = Bayesian Information Criteria

Indicator Variables Distribution by Predicted Class

The distribution and proportions of the categorical indicator variables by latent class are presented in Table 37. Due to their respective ratings for food shopping habits, label engagement and motivations when shopping (their engagement with labels, their level of motivation to use labels, and knowledge of healthy eating), the three classes were titled 'highly engaged, knowledgeable and motivated', 'somewhat engaged, knowledgeable and motivated', and 'less engaged, knowledgeable and motivated'). All indicator variables differed significantly between the three predicted classes (P<0.001), except for having any children in the household (P=0.060).

Class 1 was characterised by individuals who were more likely to be very interested in healthy eating, report having a lot of knowledge of healthy eating, more likely to use labels often, often be influenced by labels and more likely to report trying to lose weight, compared to the other classes. Approximately half of the overall sample (53%) were classified as class 1 'highly engaged'.

Class 2 was characterised by individuals who were more likely use labels occasionally and have some knowledge in healthy eating compared to the other classes, and a higher probability of having children in the household. Approximately a third of the sample were classified as class 2 'somewhat engaged' (31%).

Class 3 was characterised by individuals who were more likely to report not being very interested and having no knowledge of healthy eating, more likely to rarely or never use labels or be influenced by them, compared to the other classes. Less than a fifth of the sample were classified as class 3 'less engaged' (17%).

Table 37: Distribution and proportion of **i**ndicator variables by predicted class membership and overall

Indicator variables	**Total sample *n* = 4,863 (%)**	**'Highly engaged' Class 1 *n* = 2,564 (%)**	**'Somewhat engaged' Class 2 *n* = 1,489 (%)**	**'Less engaged' Class 3 *n* = 810 (%)**	
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	**χ^2 (*p*-value)**
Current label use					
Often	2,669 (55)	2,388 (93)	278 (19)	3 (0)	
Occasionally	1,419 (29)	163 (6)	1,098 (74)	158 (20)	
Rarely or never	775 (16)	13 (1)	113 (8)	649 (80)	5500 (P<0.001)
Influenced by labels					
Often	2,710 (56)	2,469 (96)	241 (16)	0 (0)	
Occasionally	1,432 (29)	93 (4)	1,241 (83)	98 (12)	
Rarely or never	721 (15)	2 (0)	7 (0)	712 (88)	7300 (P<0.001)
Interested in healthy eating					
Very interested	2,053 (42)	1,708 (67)	210 (14)	135 (17)	

Quite interested	2,490 (51)	856 (33)	1,179 (79)	455 (56)	
Not very or at all interested	320 (7)	0 (0)	100 (7)	220 (27)	1800 (P<0.001)
Knowledge about healthy eating					
A lot of knowledge	1,380 (28)	1,077 (42)	179 (12)	124 (15)	
Some knowledge	2,809 (58)	1,415 (55)	1,033 (69)	361 (45)	
A little or no knowledge	673 (14)	72 (3)	276 (19)	325 (40)	1100 (P<0.001)
Currently trying to lose weight					
Yes	2,272 (47)	1,325 (52)	694 (47)	253 (31)	
No	2,412 (50)	1,161 (45)	728 (49)	523 (65)	
Prefer not to say	179 (4)	78 (3)	67 (5)	34 (4)	107.6 (P<0.001)
Responsible for shopping					
Yes	4,656 (96)	2,494 (97)	1,423 (96)	739 (91)	
No	206 (4)	70 (3)	66 (4)	70 (9)	53.4 (P<0.001)
Children in the household					
Yes	1,432 (30)	730 (29)	474 (32)	228 (28)	
No	3,410 (70)	1,818 (71)	1,012 (68)	580 (72)	5.6 (0.060)

Posterior Analysis/ Proximal Variables

Sociodemographic Variables

The sociodemographic composition analyses of the three predicted latent classes are presented in Table 38 (proportion by predicted class with chi-square analysis) and Table 39 (multinomial logistic regression). The 'highly engaged' class were significantly more likely to comprise participants who were older and have an A-levels education or above, compared to the other classes; they were also more likely to be from a high household income group (£3701 or more per month) compared to the 'less engaged' class. In contrast, the 'less engaged' class were more likely to include participants who were younger, from a low household income group (less than £1200 per month) and have a below A-levels education, compared to the other classes. Participants classified as belonging to the 'somewhat engaged' class were more likely to be middle-aged, belong to a middle household income group (£2201-3700 per month) and less likely to be white, compared to the other classes. The mean BMI was classified as overweight in all classes ('highly' = 26.4; 'somewhat' = 27.0; 'less' = 26.9), the likelihood of having a lower BMI was significant for 'highly engaged' compared to 'somewhat' and 'less'.

Table 38: Sociodemographic characteristics (frequency or mean) by predicted class membership

	Total sample	**'Highly engaged' Class 1**	**'Somewhat engaged' Class 2**	**'Less engaged' Class 3**	
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	**χ^2 (*p*-value)**
Sex					
Female	2,768 (57)	1,619 (63)	803 (54)	346 (43)	
Male	2,095 (43)	945 (37)	686 (46)	464 (57)	112.6 (P<0.001)
Age					
18-29	289 (6)	116 (5)	97 (7)	76 (9)	
30-39	677 (14)	321 (13)	223 (15)	133 (16)	
40-49	919 (19)	493 (19)	303 (20)	123 (15)	
50-59	1017 (21)	541 (21)	303 (20)	173 (21)	
60-69	1023 (21)	591 (23)	275 (18)	157 (19)	
70+	932 (19)	498 (19)	287 (19)	147 (18)	53.3 (P<0.001)
Household income per month (£)					
Less than 1,200	758 (16)	370 (15)	214 (15)	174 (22)	
1,201-2,200	1,156 (25)	591 (24)	360 (25)	205 (26)	
2,201-3,700	1,319 (28)	691 (28)	417 (29)	211 (27)	
3,701 or more	1,431 (31)	804 (33)	437 (31)	190 (24)	36.9 (P<0.001)
Ethnicity					
White	4,470 (93)	2,370 (94)	1,358 (92)	742 (93)	
Non-white	317 (7)	150 (6)	114 (8)	53 (7)	4.8 (0.089)
Education					
A levels +	3,304 (68)	1,866 (73)	980 (66)	458 (57)	
Below A levels	1,543 (32)	686 (27)	507 (34)	350 (43)	81.5 (P<0.001)
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**F (*p*-value)**
BMI	26.7 (5.5)	26.4 (5.5)	27.0 (5.5)	26.9 (5.7)	6.7 (0.0013)

SD: standard deviation; BMI = body mass index.

Table 39: Multinomial logistic regression of sociodemographic variables on predicted class membership.

	'Somewhat engaged' vs 'Highly engaged' Class 2 vs 1 (ref)		'Less engaged' vs 'Highly engaged' Class 3 vs 1 (ref)		'Less engaged' vs 'Somewhat engaged' Class 3 vs 2 (ref)	
	RRR 95% CI	***p*-value**	**RRR 95% CI**	***p*-value**	**RRR 95% CI**	***p*-value**
Age (18-29 ref)						
30-39	0.83 (0.60, 1.14)	0.255	0.63 (0.44, 0.90)	0.011	0.76 (0.53, 1.10)	0.147
40-49	0.74 (0.54, 1.00)	0.048	0.38 (0.27, 0.54)	<0.001	0.52 (0.36, 0.75)	<0.001
50-59	0.67 (0.49, 0.91)	0.010	0.49 (0.35, 0.68)	<0.001	0.73 (0.51, 1.04)	0.079
60-69	0.56 (0.41, 0.76)	<0.001	0.41 (0.29, 0.57)	<0.001	0.73 (0.51, 1.04)	0.084
70+	0.69 (0.51, 0.94)	0.017	0.45 (0.32, 0.64)	<0.001	0.65 (0.46, 0.94)	0.021
Household income per month (£) (less than 1,200 ref)						
1,201 - 2,200	1.05 (0.85, 1.30)	0.634	0.74 (0.58, 0.94)	0.013	0.70 (0.54, 0.91)	0.008
2,201 - 3,700	1.04 (0.85, 1.28)	0.688	0.65 (0.51, 0.82)	<0.001	0.62 (0.48, 0.81)	<0.001
3,701 or more	0.94 (0.77, 1.15)	0.552	0.50 (0.40, 0.64)	<0.001	0.54 (0.41, 0.70)	<0.001
Ethnicity (non-white ref)						
White	0.75 (0.59, 0.97)	0.028	0.89 (0.64, 1.23)	0.464	1.18 (0.84, 1.65)	0.349
Education (below A-levels ref)						
A levels+	0.71 (0.62, 0.81)	<0.001	0.48 (0.41, 0.57)	<0.001	0.68 (0.57, 0.81)	<0.001
BMI (every 1-unit increase)	1.02 (1.01, 1.03)	0.001	1.02 (1.00, 1.03)	0.026	1.00 (0.98, 1.01)	0.590

RRR = relative risk ratio; CI = confidence interval. *The 'highly engaged' class were chosen as the main reference group as it was the largest class.

Experiment outcomes - Ranking Ability

The experimental ranking outcome analyses by predicted latent class are presented in Table 40 (mean or median outcomes) and Table 41 (multinomial logistic regression). The 'less engaged' class participants were significantly more likely to rank fewer product categories correctly at baseline (without FOPL), compared to ‘somewhat’ and ‘highly engaged’. ANOVA analyses and multinomial logistic regression results showed that the ability to rank product categories correctly at follow-up (with FOPLs) differed significantly between all of the classes, with the more engaged classes more likely to rank more product categories correctly (e.g., ‘highly engaged’ compared to ‘somewhat’; $P<0.001$). The mean global food score results show the ‘highly engaged’ class were significantly more likely to have a higher score compared to the other classes ($P<0.001$), but these differences were small (RRRs between 0.86-0.96). To further explore the impact of predicted class membership on ranking ability, the mean global food score was examined by predicted class and experimental group, presented in Table F2**.** Within FOPL schemes there were no significant differences between the predicted classes. Across all classes, N-S led to the largest mean improvement, followed by MTL, then WL and PC (the no label control group led to a decrease). The speed of baseline ranking (without FOPL) was found to be significantly faster for the ‘less engaged’ class participants compared to both of the other classes (median speed of 198 seconds, compared to 209 seconds for ‘somewhat’ and 218 for ‘highly engaged’). For speed of follow-up ranking (with FOPLs), the only observed difference was the ‘less engaged’ class were significantly faster compared to the ‘highly engaged’ (median speed of 110 seconds compared to 118 seconds). The results presented did not require participants ranking the products correctly, for consistency with previous analyses, the results with the requirement for being correct are presented in Table F3 in the Appendix F, showing consistent results.

Table 40: Mean ranking outcomes by predicted class membership

	Total sample	'Highly engaged' Class 1	'Somewhat engaged' Class 2	'Less engaged' Class 3	
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**F (*p*-value)**
Baseline ranking (0-6)	3.0 (1.2)	3.0 (1.2)	3.0 (1.2)	2.8 (1.2)	10.3 (<0.001)
Follow-up ranking (0-6)	4.3 (1.7)	4.4 (1.7)	4.2 (1.7)	3.9 (1.8)	25.7 (<0.001)
Global score (-5-+5)	1.1 (1.5)	1.2 (1.4)	1.0 (1.5)	0.9 (1.5)	10.6 (<0.001)
	Median (IQR)	**Median (IQR)**	**Median (IQR)**	**Median (IQR)**	**χ^2 (*p*-value)**
Baseline speed (secs)	212 (160-291)	218 (165-297)	209 (158-295)	198 (150-267)	17.8 (<0.001)
Follow-up speed (secs)	116 (87-162)	118 (88-167)	113 (87-161)	110 (85-151)	10.7 (0.005)

Secs = seconds *For speed outcomes there was no requirement of being correct. SD = standard deviation.

Table 41: Univariate multinomial logistic regression of ranking outcomes on predicted class membership.

	'Somewhat engaged' vs 'Highly' Class 2 vs 1 (ref)		'Less engaged' vs 'Highly' Class 3 vs 1 (ref)		Less engaged' vs 'Somewhat' Class 3 vs 2 (ref)	
	RRR (95% CI)	***p*-value**	**RRR (95% CI)**	***p*-value**	**RRR (95% CI)**	***p*-value**
Baseline ranking (correct)	0.99 (0.94, 1.04)	0.637	0.86 (0.81, 0.92)	<0.001	0.87 (0.81, 0.94)	<0.001
Follow-up ranking (correct)	0.92 (0.89, 0.96)	<0.001	0.86 (0.82, 0.89)	<0.001	0.93 (0.88, 0.97)	0.002
Global food score	0.93 (0.89, 0.97)	0.001	0.89 (0.85, 0.95)	<0.001	0.97 (0.91, 1.03)	0.247
Baseline speed	1.00 (1.00, 1.00)	0.214	1.00 (1.00, 1.00)	0.050	1.00 (1.00, 1.00)	0.010
Follow-up speed	1.00 (1.00, 1.00)	0.066	1.00 (1.00, 1.00)	0.050	1.00 (1.00, 1.00)	0.557

RRR = relative risk ratio; CI = confidence interval; BMI = body mass index. Class 2, the 'highly engaged' class is the reference group compared to 'somewhat' and 'less engaged' and 'somewhat' compared to 'less engaged'.

Experiment Outcomes - Label Perception Outcomes

The label perception outcomes analyses by predicted class are presented in Table 42 (frequency and proportion). The perception questions were specific to the allocated label condition; therefore, the no label control group was excluded. This focus of this analysis was overall labels perceptions by predicted class; therefore, responses are not presented by FOPL group here (see Table F4 in Appendix F for label perception outcomes by predicted class and FOPL group). Overall, each of the label perceptions differed significantly between predicted classes. The 'highly engaged' class had the highest proportion of participants who reported: seeing and using the label in all ranking tasks; that the labels were easy to understand; quick enough to use; and supported having labels on all food and drink products in the UK. The 'somewhat engaged' class had the highest proportion of participants who reported the labels were helpful in making purchasing decisions, but otherwise followed 'highly engaged' in proportion. The 'less engaged' class had the lowest proportions of participants positively perceiving the labels for each of the label perception outcomes.

Table 42: Proportion of participant label perception outcomes, by predicted class membership

	FOPL groups combined *	'Highly engaged' Class 1	'Somewhat engaged' Class 2	'Less engaged' Class 3	χ^2 (*p*-value)
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	
Saw label					38.7 (P<0.001)
Yes	2,870 (74)	1,619 (77)	834 (72)	417 (67)	
No/not sure/don't know	1014 (26)	482 (23)	330 (28)	204 (33)	
Used label					47.1 (P<0.001)
All	1,992 (51)	1,157 (55)	573 (49)	262 (42)	
Some	587 (15)	314 (15)	182 (16)	91 (15)	
Did not use	289 (7)	147 (7)	78 (7)	64 (10)	
Don't know/not applicable	1,016 (26)	482 (23)	330 (28)	204 (33)	
Understanding labels					21.4 (P<0.001)
Easy	2,391 (83)	1,382 (85)	693 (83)	316 (76)	
Difficult	477 (17)	236 (15)	141 (17)	100 (24)	
Label helpfulness					39.5 (P<0.001)
Helpful	3,415 (88)	1,833 (87)	1,072 (92)	510 (82)	
Not helpful	469 (12)	266 (13)	92 (8)	111 (18)	
Time to use label$^+$					39.0 (P<0.001)
Quick enough	3,622 (94)	1,990 (95)	1,088 (94)	544 (88)	
Too long	250 (6)	104 (5)	72 (6)	74 (12)	
Labels on products in UK^					78.5 (P<0.001)
Yes-all	2,810 (72)	1,593 (76)	838 (72)	379 (61)	
Yes-some	748 (19)	322 (15)	261 (22)	165 (27)	
No-none	322 (8)	181 (9)	64 (6)	77 (12)	

*Control group not included. $^+$ 14 do not know or refused responses; ^ 6 do not know responses. Helpful dichotomised as helpful (very helpful, quite helpful) vs not helpful (not very helpful, not at all helpful). Understanding dichotomised as easy (very or quite) vs difficult (quite or very).

Experiment Outcome – Descriptive Analysis of Ranking Perceptions

The descriptive results of the ranking perceptions by predicated class are presented in Table 43 (see Chapter 5 for these ranking perceptions presented by FOPL group). For the baseline ranking tasks (without FOPLs), the 'less engaged' class had the highest proportions of participants reporting they had enough information to rank the products (across all product categories, 12-27%), followed by the 'somewhat engaged' class (9-19%) and then the 'highly engaged' class with the lowest proportions (7-17%). Confidence in ranking products correctly at baseline was similarly low across all classes. For the follow-up ranking tasks (with FOPLs), participants across all three classes had similar responses to having enough information to complete the rankings (within product categories). Participants' confidence in ranking the products correctly at follow-up varied, the 'less engaged' class had the lowest proportion reporting that they had ranked all or most of them correctly and the highest proportion reporting some of them or 'don't know'; in contrast, the 'highly engaged' class had the highest proportion who reported that they had ranked all or most of them correctly.

Table 43: Frequency and proportion of ranking perceptions, enough information to rank the products and confidence in the ranking tasks, at baseline and follow-up, by predicted class membership

	Total sample *n* **= 4,762**		**'Highly engaged' Class 1** *n* **= 2,516**		**'Somewhat engaged' Class 2** *n* **= 1,457**		**'Less engaged' Class 3** *n* **= 789**	
	BL	**FU**	**BL**	**FU**	**BL**	**FU**	**BL**	**FU**
	n **(%)**	***n*** **(%)**	***n*** **(%)**	***n*** **(%)**	***n*** **(%)**	***n*** **(%)**	***n*** **(%)**	***n*** **(%)**
Reported 'yes' to having enough information to rank (*n* = 4,863)								
Pizza	763 (16)	1,982 (41)	355 (14)	1,043 (41)	243 (16)	608 (41)	164 (20)	331 (41)
Drink	406 (8)	1,889 (39)	178 (7)	1,001 (39)	127 (9)	572 (38)	101 (12)	316 (39)
Cake	800 (16)	2,082 (43)	359 (14)	1,088 (42)	262 (18)	637 (43)	179 (22)	357 (44)
Crisps	949 (20)	2,130 (44)	444 (17)	1,120 (44)	290 (19)	652 (44)	215 (27)	358 (44)
Yoghurt	543 (11)	1,912 (39)	234 (9)	1,009 (39)	180 (12)	580 (39)	129 (16)	323 (40)
Cereal	595 (12)	1,943 (40)	259 (10)	1,027 (40)	197 (13)	593 (40)	139 (17)	323 (40)
Confidence in ranking (how many of them do you think you ranked correctly?)								
All or most	691 (14)	2,153 (45)	376 (14)	1207 (47)	198 (13)	623 (42)	117 (15)	323 (39)
Some	2,586 (53)	1,862 (38)	1,335 (52)	922 (36)	829 (56)	613 (41)	422 (52)	327 (40)
None	498 (10)	205 (4)	251 (10)	105 (4)	158 (11)	62 (4)	89 (11)	38 (5)
Don't know	1,086 (22)	641 (13)	601 (23)	329 (13)	304 (20)	191 (13)	181 (22)	121 (15)

BL = baseline; FU = follow-up.

Chapter 7 Discussion

This LCA examined motivation and engagement characteristic patterns in a sample from a randomised controlled experiment testing the effectiveness of FOPLs. Three latent classes were identified, described as 'highly engaged, knowledgeable and motivated', 'somewhat motivated, knowledgeable and motivated', and 'less engaged, knowledgeable and motivated', based on differing levels of label engagement, knowledge of healthy eating and motivation to use labels. The sociodemographic composition of the predicted classes indicated that participants classified as 'less engaged' were more likely to have disadvantaged SES characteristics (lower education and income levels), compared to 'highly engaged' who were more likely to have advantaged SES characteristics (higher education and income levels). The results of the experimental outcome analyses suggest that ability to use FOPLs does differ according to engagement and motivation characteristics, but this is a small impact and disappears when examined by FOPL scheme. Conversely, participants' confidence in correctly using FOPLs and their broader perceptions of FOPLs did differ significantly according to engagement characteristics, with the 'highly engaged' class having the highest confidence in ranking ability and the most positive label perceptions.

The sociodemographic differences observed between the predicted classes is supported in the literature.[43, 45, 47, 51] The finding that ranking and perception outcomes differed by predicted class is also supported in the literature, with evidence suggesting that engagement with FOPLs and food buying behaviours is associated with motivations to eat healthily, nutrition knowledge and household factors.[33, 38, 43, 48, 49, 56, 173, 174] The 'less engaged' class had a lower baseline ranking score (without FOPLs) and the smallest improvement from baseline to follow-up, which could be explained by their lower self-reported nutritional knowledge.

As previously discussed in Chapter 5, N-S and MTL led to the largest improvement in ranking ability, with evidence from the LCA suggesting that there was little difference in this improvement when examined by participants' engagement characteristics. The results from Chapter 6, show that overall perceptions of FOPLs (especially N-S and MTL) were positive, but when examined by predicted classes within FOPL schemes there were differences. The 'less engaged' class, with disadvantaged SES characteristics, had the lowest proportion of participants who were confident in their ability to use FOPLs in the ranking tasks. This is a proxy measure for subjective understanding (if participants *think* they understand the label; whereas objective understanding is if they *do* understand it correctly), a key element in the conceptual framework for label use developed by Grunert and Wills.[37] Other factors in the framework include seeing and liking the label, for which the 'less engaged' class also had the lowest proportions. The results suggest that FOPLs work regardless of predicted class (i.e., improve the ability to correctly understand product healthiness) but the likelihood they will be used differs (i.e., confidence in using and liking of labels).

As discussed in Chapter 1, FOPLs are a downstream intervention, focusing on individual behaviour change (in addition to reformulation via industry), rather than an upstream intervention, which focuses on changing social factors.[26] Downstream interventions can exacerbate health inequalities, due to the reliance on individual agency. Therefore, it is crucial that the chosen FOPL is suitable for use across the population, specifically those with lower education levels or nutritional knowledge (who would be classified as 'less engaged') and does not widen inequalities. There is evidence from this LCA that N-S may be more suitable for use across the UK population. N-S users across all classes showed no differences in perceptions for seeing the label or reporting it was quick enough to use; but there were significant differences for MTL users (see Table F4). This is supported by results from an earlier study and Chapter 6.[73] Time, in addition to perceptions, has been identified as a barrier for label use.[38, 42, 46, 50] The 'less engaged' class participants completed all ranking tasks significantly faster than the other classes, which may be reflective of their real food purchasing behaviours. The results presented in Chapter 5 and 6, show that participants using N-S were the fastest at ranking products correctly and also suggested that N-S may work better across income groups and education levels. Evidence from another experiment testing the ability of participants to correctly rank products supports this, with N-S appearing to work better in those with lower education levels, compared to the other labels including MTL.[152] A cross-sectional study testing perceptions of four labels (including N-S, MTL) found N-S had the widest reach, appealing to subjects with lower adherence to nutritional recommendations, compared to the other labels.[73] Secondary analysis of a large international study (18 countries, n = 18,393) examined the effectiveness of five FOPLs according to income status, and found comparable effects for all five labels; with N-S leading to the largest improvement across all income categories.[175]

To improve the perceptions of FOPLs and confidence in using them correctly, consumer awareness and public education campaigns could be conducted, aimed at improving these factors. Trust in the labels and subjective understanding, where consumers think they understand the labels, are two key factors for engagement with labels. This is especially vital for participants who would be classified as ‘less engaged’. For example, this could include targeted media campaigns or in-aisle information explaining how the FOPL was created, how the information/score is determined and how to use them. An experimental study that examined the provision of an in-aisle explanation of FOPLs were found to be somewhat helpful to consumers in making healthier decisions.[126] Another experimental study examining an early iteration of N-S tested the impact of a communication leaflet that explained the logic behind N-S and how to use it to compare products found that recall and understanding of N-S was higher compared to other conditions without the leaflet.[176]

Potential limitations of the analysis include choosing the three-class model, as the decreasing BIC indicates more classes may have been a better model; but this was balanced by the other indicators. Participants only answered perception questions about the FOPL scheme they were randomly allocated to and used in the ranking tasks, so it is not possible to make comparisons within participants. Although, this does mean that their responses to perception questions were based on their actual use of the label. The timing data was taken from the survey software, so care should be taken with the interpretation of the findings. All participants were given the same instructions: “We would like to know how long it takes for you to rank the different foods, so please click ‘next’ as soon as you have completed each task”. The study was powered for the primary aims, and so formal statistical testing was limited. It is also not possible to determine causality from LCA as it uses cross-sectional data. This analysis suggests possible clusters but needs to be confirmed using longitudinal data. Although, the measures of food habits, label use and motivation factors were related to their baseline/current habits, which I assume precede the experimental outcomes and therefore may suggest causality.

Chapter 7 Conclusions

The results presented in this Chapter suggest that regardless of predicted class, FOPLs improve the ability to correctly understand product healthiness. But the likelihood they will be used differs according to engagement and motivation characteristics. Therefore, the practical implementation of FOPLs in the UK will be crucial to encourage engagement across the population and not widen health inequalities. Consumer awareness and public education campaigns (such as media campaigns or in-aisle information) could target consumer confidence in using FOPLs and their perceptions of FOPLs, especially for participants who would be classified as 'less engaged'. In the next Chapter, I will present the findings of the Patient and Public Involvement and Engagement session that I conducted with children and young people, to better understand and discuss their views on FOPLs and UK food policy. This session builds on the findings from Chapter 5-6, with the aim to discuss the findings of the experiment, their real-world experience of FOPLs (including the impact of motivation and other factors on real FOPL use) and FOPL policy suggestions.

Chapter 8: Patient and Public Involvement and Engagement Session

Chapter 8 Background

In this Chapter, I will outline the importance of involving the public in research, describe the aim and methods of this session and discuss the findings and views of the children and young people (CYP). The topic of this session relates to the main experiment, as presented in Chapters 4-7.

Patient and Public Involvement and Engagement (PPIE) is research that is carried out 'with' or 'by' members of the public, instead of 'to', 'about' or 'for' them. By integrating the public voice, PPIE can enhance the quality of research (including the design), improve the appropriateness and ethics of research, and lead to research that is more relevant and impactful.[177-179] The importance of PPIE is recognised by research funding bodies, including NIHR, with many requiring and prioritising PPIE throughout all stages of research.[178] Beyond benefitting research output, there is also a benefit for researchers, who can gain increased understanding and insight into their field, and for PPIE contributors in terms of self-confidence, self-worth and skill development.[177, 180] PPIE with CYP requires additional care to ensure the topic is not too challenging or sensitive, but still beneficial for the research projects and the CYP.[181, 182]

The best practices for public involvement in research have been outlined in the six UK standards for Public Involvement.[183] The six standards are: inclusive opportunities, to ensure the PPIE opportunities are accessible and reach the relevant groups and people; working together, to ensure that the PPIE session is collaborative, that all contributions are valued and that the relationships are mutually respectful and productive; support and learning, to allow the necessary support and training to build confidence and skills needs for PPIE involvement; governance, facilitating involvement in the management, regulation, leadership and decision making; communications, ensure the communication is relevant and appropriate, during the planning and PPIE activities; and impact, to ensure that the impact that the PPIE made to the research is identified and shared. Other best practice principles for using the findings from PPIE in terms of impact and dissemination include involving PPIE participants in the research outputs and identifying their involvement; providing feedback and communicating the research findings throughout the process, the final results and outputs, in accessible formats and language.[177-179, 184] Specific strategies for involving CYP include showing them respect for their contributions and making them feel included; giving them feedback on the outputs and impact of the PPIE sessions, to illustrate the value of their contributions and that their involvement is valued; and involving them in as many aspects of the research as possible, throughout the entire process.[182]

Rationale

In 2021, I organised a PPIE session with the National Children's Bureau (NCB), who are one of the OPRU public engagement partners. The NCB is focused on work that promotes the wellbeing of CYP, by promoting their active participation in research and the policies that affect them.[185] The session was conducted with the Young Research Advisors (YRA) group, who are trained on research methods, ethics and key policy issues.[185] This session was useful to better understand the views of CYP as the NatCen panel is 18 years and older, so their perspectives were missing from the main experiment.[153] The main experiment had no qualitative component or opportunity for feedback, due to budget and time constraints, so this session also provided a learning opportunity to explore this topic and any issues in more detail, provide greater insight into the findings and thought-processes, and discuss real-world application of FOPLs. Although the CYP are unlikely to be the main household grocery shopper, their insights as adolescents with autonomy who are becoming food purchasers is an interesting perspective. The aim of the PPIE session was to examine the YRA's ability to rank products through an exploratory adapted mini experiment, to present and discuss the findings of the main experiment (as presented in Chapter 5 and 6) and discuss their use of FOPLs and their views towards UK FOPL policy.

Chapter 8 Methods

I conducted the session on the 19th of August 2021, with facilitation by two researchers from NCB. Seven YRAs accepted the invitation from NCB to take part. The group all identified as female, they were aged between 12-18 years and were from across England. The session was conducted entirely online (because of COVID-19 restrictions) using Zoom and lasted two hours. The session was composed of three sections, an online mini experiment involving a ranking task and survey, a brief presentation of the main experiment results, and a facilitated discussion session.

Before discussing or explaining FOPLs, the YRAs completed an online mini experiment. The experiment was conducted on Mentimeter,[186] which is a free interactive presentation software, where participants enter a code to access an online presentation and are able to vote/interact anonymously. The mini experiment was adapted from the main experiment and included a ranking task, followed by a survey about their perceptions of the labels and their current food habits. The ranking task involved three product categories (cake, pizza, cereal), each with three products of varying healthiness ('most healthy', 'in between' and 'least healthy'). Only three of the original six products were chosen due to time restraints, and they were selected because they were representative of the different product types (yogurt /cereal were hardest for people to rank without labels, hot chocolate was excluded from global food score). The participants navigated to the webpage on their own devices, which displayed the product images and they were asked to rank the products according to the healthiness of the products, firstly with no labels and then again with a FOPL (cake products paired with MTL; pizza products paired with N-S; cereal products paired with PC), see Figure 26 for an example. Following the ranking of each product category, participants were asked if they felt they had enough information to rank the products (yes or no). Once all ranking tasks were completed, they were then asked their views on the labels (which they liked most; which they thought was the quickest; which was the easiest; their view on mandatory labelling in the UK) and their food habits (food shopping responsibility; food label use; knowledge/interest in healthy eating) (see Appendix G for the detailed survey and Figure G1 for a ranking task example). The directions for the ranking task and survey were provided on the presentation, but I also read instructions out aloud, to ensure that the tasks were understood.

Figure 26: Example of the ranking task from Mentimeter: (a) without labels, (b) with Multiple Traffic Light

(a)

(b)

Following the mini experiment, I presented some background information on FOPLs, including their aims and policy relevance. I then explained the methods and results of the main experiment and presented the results of the mini experiment they had just completed, including how their results compared to the main findings.

This was then followed by a short break and an ice-breaker activity led by the NCB researchers, before the discussion session. The facilitated discussion lasted 40 minutes and the group was split into two smaller groups to allow for better engagement. The discussions were guided by the following questions, to ensure the same topics were covered by each group:

- Do you agree with the research findings?
- What are your experiences of front of pack labels?
- What do you think about front of pack labels?
- What do you think about the policy options? What should be done in the real-world?

Chapter 8 Results

Mini Experiment

Ranking Task

Six of the seven CYP took part in the online ranking tasks and mini experiment (due to internet issues faced by one of the CYP). For the baseline tasks (with no FOPL), the ability to correctly rank the three products according to healthiness varied between product categories. For the cereal products, only one of the CYP were correct, and for pizza and cake products three were correct. For the follow-up ranking tasks (with FOPL), the ability to correctly rank the products improved overall. For the cake products displaying MTL and the cereal products displaying PC, four of the CYP were correct; and all six of the CYP were correct when ranking the pizza products displaying N-S. None of the CYP responded that they had enough information to complete the rankings for any of the three product categories at baseline; for the follow-up rankings, three CYP said they had enough information with MTL; four said they had enough information with N-S, and none of the CYP said they had enough information with PC.

Label Perceptions and Food Habits Survey

In response to the label perception questions, five of the six CYP liked the MTL label most but all responded that N-S was both the easiest and the quickest to use. Regarding support for mandatory labelling, all CYP agreed that there should be labelling, either on all products (four CYP) or on some products (two CYP). For the food shopping questions, all CYP reported being responsible for some food shopping; majority read the labels quite often when buying food or drink (four CYP) and the rest reported reading labels occasionally (two CYP). The CYP had high levels of self-reported interest in healthy eating (five CYP) and having some knowledge of healthy eating (five CYP).

Discussion Session

Agreement with Main Experiment Findings

In response to the main experiment research findings, the CYP agreed with the findings and reported having the same experience during the mini experiment (as they also found FOPLs helped them rank products according to their healthiness). In regards to their experiences with

FOPLs, some of the group had previous knowledge of MTL due to studying nutrition at school. Some of the group reported previous use of MTLs, and discussed how they found the MTL label helpful for making decisions. They also mentioned that price is an important factor when making choices about what food to buy, along with how healthy it is. Some of the CYP mentioned experiences of trying to find information on food and finding it difficult.

Views on FOPL Use and Policy Suggestions

The CYP views on the FOPL options were that, if FOPLs are going to be used then everyone (i.e., manufacturers) needs to use them and they should not be voluntary because then some will choose not to display them. They reiterated that FOPLs help to understand the healthiness of a product but the fact that they are not on every pack, means they are not that helpful. The CYP liked MTL as it was familiar to them and it gave more information, so they felt that they could use that information and then decide for themselves. They agreed that N-S was a good label because it does not require a lot of knowledge, it is intuitive and they all agreed it was easiest to understand "*Easiest for people that don't have a clue. Colours and letters are universal, you don't need much knowledge to put it into perspective*." The CYP agreed that the WL and PC were not helpful as they do not provide much information. See Table 44 for more detailed YRA's responses by FOPL type.

Table 44: Young Research Advisor views by FOPL type

Label	**Young Research Advisor views**
Multiple Traffic Light (MTL)	- They agreed MTL was useful because it allows people to make quick decisions. - They commented that MTL displays exactly what is in the food. They thought this may be especially helpful for people with specific dietary requirements, such as people with high blood pressure who want to avoid high salt. - They discussed that the recommended serving sizes shown on MTL were not helpful as the recommended serving size is different for each person. They elaborated that the recommended daily intake for calories, fats and protein varies depending on multiple factors such as gender, height, activity level and age. - They also commented that the red traffic light could be potentially triggering for people with eating disorders or issues with food.
Nutri-Score (N-S)	-The CYP all reported that they had never seen this label but they liked it.

	- They stated that N-S does not require a lot of knowledge to understand. They liked the grading system, as they stated that it is possible to compare it to school marks, which they agreed most people would find it easy to understand. - All CYP agreed that the N-S label was the easiest to understand and quick to use. - Some of the CYP were unsure about the rating and wanted to form their own judgement, because they did not know how the rating was decided.
Warning Label (WL)	- They agreed that WL was not very helpful because it does not give enough information and is difficult to compare the foods together, which is more helpful. - Some CYP thought it might be a trigger for someone with an eating disorder. - The CYP did not understand the threshold for warnings and did not trust it.
Positive Choice Tick (PC)	- The CYP thought that the PC label was unhelpful because it did not give any explanation and does not allow people to make their own decisions between 'healthy foods'.

During the facilitated discussion regarding their FOPL policy suggestions, the CYP were engaged and had numerous ideas regarding the FOPL selection, design and implementation. Their ideas for the selection of a FOPL in the UK were that MTL and N-S should be used because they are easier to understand than the other options. Both groups suggested that a combined FOPL should be developed using the best elements of MTL and NS, as they stated it would make the label even clearer. Regarding the design of a label, they discussed that the labels should be bigger, to make them easier to see. For the policy implementation in the UK, they said that FOPLs should be on all packaging and that a mobile application could be created, where a user could scan the label on the food to compare the healthiness of products and the information could be tailored to the individual "*meaning that the traffic light system could adapt to your personal recommended daily intake*". It was also discussed that this might even reduce the amount of packaging needed, which could have positive environmental benefits. One CYP added that they thought it would be more helpful for the label to represent the whole pack of food not just a portion.

The CYP also had wider food policy suggestions for policymakers. They suggested that the UK government should continue to implement other policy interventions to help address obesity and improve the healthiness of diets, including regulations on the placement of food items in supermarkets (i.e., no 'junk food' being placed next to the tills at the supermarket). They stated that there should be limitations on what food manufacturers can put into food, for example sugar limits. They also mentioned that currently healthy food is often more expensive compared to fast food, and therefore policies need updating or should be developed to make healthy food cheaper and more accessible.

Following the session, the CYP filled in an online evaluation of the PPIE session on what they enjoyed and what they did not enjoy. Their feedback showed that they enjoyed the session, particularly the online mini experiment and the smaller breakout rooms. There was no negative feedback, just a comment from one young person reiterating that the PC and WL labels should be used less.

Chapter 8 Discussion

The aim of this PPIE session was to understand the ability of CYP to use FOPLs, to present the findings of the main experiment, and to discuss their views, experiences and policy suggestions regarding FOPLs. The key findings of the mini experiment showed FOPLs improved the ability of CYP to correctly rank food items according to their perceived healthiness, with N-S and MTL showing the biggest improvement, and they agreed with the findings of the main experiment. The discussion session showed CYP strongly supported the implementation of a mandatory FOPL policy in the UK, and led to suggestions for a combined MTL and N-S label. The CYP were highly engaged throughout the session and with the topic. The findings from the PPIE session were presented in a poster for policymakers at DHSC and submitted to the UCL Great Ormond Street Institute of Child Health Postgraduate Research Showcase (see Figure G2 in Appendix G).

The results of the mini experiment were supportive of the main experiment, and the discussion session provided greater insight into the results. In the survey, MTL was found to be the most liked label in the survey following the ranking tasks, while N-S was found to be the easiest and quickest to use. In the discussion session, the CYP said they liked MTL as it was familiar, which is known to impact the liking and use of FOPLS, and likely impacted the survey results as it was the only label that was familiar to the group at that stage. Following the presentation with background information provided for the other labels, the discussion session showed that N-S and MTL were the most popular FOPLs. All CYP agreed that N-S was the easiest to understand, in both the survey (before it being explained to them) and in the discussion session. One of the discussion points raised was that N-S is a label that would be helpful for people with lower nutritional knowledge, as it does not require a lot of knowledge to understand, due to the alphabet-based system being comparable to the school mark system and therefore most people would find it easy to understand. This has also been theorised in the analysis of the main experiment (Chapter 5), latent class analysis (Chapter 7) and is also supported in the literature.[69] The use of colour in MTL and N-S was also popular with CYP, as they stated it was easy to understand, this has also been discussed in earlier Chapters (Chapters 3, 5 and 6) and is supported in the literature.[33, 59, 150]

The FOPL policy ideas discussed by the CYP were thought-provoking but some of the practicalities are limiting. For example, their suggestion of creating a combined MTL and N-S label would require extensive research and investment, including the design, legal and copyright issues, in addition to an experiment to test if this hypothetical label would be effective in a UK sample (similar to the main experiment). Their suggestions for making the label bigger would potentially face pushback from industry, such as delaying progress (e.g., by pushing for more research or longer consultation periods), a common tactic as described by the World Cancer Research Fund (WCRF).[187] The mobile application idea is interesting and there are similar schemes that exist. For example, the FoodSwitch application which allows consumers to scan the barcode of a product to see how healthy the product is and to see alternative healthier options.[188] Their suggestion of this shows that they are not aware of this application and that government endorsement or promotion would be helpful in increasing the use. Their comments on issues with the portion sizes informing the MTL FOPL have been raised by consumers in qualitative studies, which I will discuss more in Chapter 9.[50]

The CYP were highly engaged with this topic and had recommendations for policymakers in regards to FOPL, and broader health and food policies to improve diet quality and target obesity. They were strongly supportive of a mandatory FOPL policy and insinuated that voluntary policies are not observed by all manufacturers, which therefore weakens the impact of FOPLs. This is a point that has been raised by researchers in policy analyses, with calls for mandatory labelling to overcome the static uptake of voluntary labels in the UK, to have the desired impact on product reformulation.[33, 189] The World Health Organization have repeatedly called for mandatory FOPLs and evidence from simulation studies show that the impact on disability-adjusted life years (DALYs) would be 10 times greater if the Australian HSR scheme was made mandatory.[75, 160, 190-192] They discussed the importance of the FOPL policy in respect to broader policies that could improve the wider food environment, including restrictions on HFSS product location and advertising, which have been approved but only partially implemented (location restrictions since October 2022, advertising restrictions pushed back from January 2023 until October 2025).[193] They also requested for policies that would make healthier food and drinks more accessible and affordable, a point that has been raised in other OPRU PPIE sessions with YRAs.[194]

Evaluating the session in line with the UK standards for Public Involvement, the PPIE session was planned and conducted following the best practice principles. Firstly, the session was inclusive, this was facilitated by the NCB, as their recruitment and participation involves children and young people from across England. They are able to provide support for different developmental requirements and assess the sessions and materials before the session, to ensure all elements will be accessible to all participants. The session was held online and at a time that worked for the CYP. They were renumerated for their time and input in the form of a voucher. Secondly, for working together, the PPIE session was designed with collaboration in mind, allowing enough time in each section to allow valuable contributions. Communication from myself as the lead researcher and facilitators was respectful at all times, allowing the CYP to voice their opinions and encouraging members to contribute. No jargon was used during the session. Thirdly, for support and learning, NCB ensures that all CYP receive ongoing training on research methods, ethics and on key policy issues. Topics that are relevant to each session are recapped and support throughout the session was offered. Fourthly, for governance, this is more applicable to adults/ other research stages but when discussing the outputs with the CYP, I involved the CYP in the discussions about how they would like the findings from the session to be communicated (e.g., poster or blogpost). Fifthly, for communication, the session was delivered to the NCB facilitators a week in advance, to ensure that the content was pitched at the correct level and that no jargon was included. Lastly, for impact, the findings of the session were included in the final reports submitted to DHSC and a poster was created and submitted to a UCL Great Ormond Street Institute of Child Health Postgraduate Research Showcase conference and to DHSC for reporting purposes. The poster and a summary of the impact of the session, including delivery and feedback from policymakers, was also sent to NCB to share with the YRA group.

The limitations of this PPIE session include that all participants were females and there was a small number, although this did allow for strong engagement and in-depth discussions. All of the CYP reported helping with the food shopping, but not as the main shopper, so their motivations and access to money may be different to the adult NatCen panel, and provide a complementary view to adults in the main experiment. A limitation of the main experiment was that due to budget and time constraints, it was not feasible to add a qualitative component. The ability to run this PPIE session with the CYP was beneficial to further understand the quantitative results. The CYP were highly aware of nutrition, they had high levels of knowledge and interest in healthy eating, which may be a reflection of the changing school curriculums. Also, important to contextualise the findings, that the YRAs are trained on research and policy issues, so are likely to be more informed than other CYP on research and policy and able to verbalise their thoughts.

Chapter 8 Conclusions

The CYP were highly engaged with the FOPL topic, and through the exposure to using the FOPLs in the mini experiment they agreed with the findings of the main experiment. In discussions about their lived experience with food shopping and using FOPLs, they provided useful insights into their real-world application of FOPLs. The CYP I engaged with thought that FOPLs that used colour were easier to understand and quicker to use. They recommended that policy makers ensure the updated FOPL policy is mandatory and broader policies are supportive of healthy eating. In the next Chapter, I will present and discuss the main findings from the thesis, discussing the mechanisms, the original contribution to knowledge, the policy implications, limitations and next steps.

Chapter 9: Discussion

Summary of Key Findings

The results presented in this thesis illustrate the effectiveness and positive perceptions of FOPLs in a British sample. They provide relevant findings for the UK evidence-base, which will assist policymakers in understanding which FOPL scheme may be best for the countries in the UK. The results identify key participant characteristics that will be crucial to target through public education and media campaigns, to ensure an effective FOPL policy that will be suitable for use across the UK population.

The key learnings of the thesis are that FOPLs do work as intended, by improving understanding of product healthiness (presented in Chapter 5 and 6). There is evidence to suggest that N-S and MTL, both classified as interpretive labels, work best of the labels tested in a British sample for objectively and subjectively measured understanding outcomes, and for participant perceptions of the labels (presented in Chapter 5 and 6), which is supported by the wider literature (as evidenced in Chapter 2). The systematic review on the impact of FOPLs on consumer behaviours also indicated that interpretive labels that used colour were more effective at impacting the healthiness of purchases than non-interpretive or non-colour coded labels (presented in Chapter 3). The Patient and Public Involvement and Engagement (PPIE) session with children and young people (CYP) also supported these findings, with the N-S and MTL FOPLs viewed as easiest to use and most effective at communicating the nutritional quality of products (presented in Chapter 8). There is further evidence to suggest that N-S in particular could be the most suitable FOPL for the UK, as it led to the highest proportion of correct ranking for all three products and the healthiest product only (presented in Chapter 5 and 6). Additionally, participants in the N-S condition were the quickest at correctly ranking the products (presented in Chapter 6) and had the highest proportion of participants who saw and used the label, and thought it was easy and quick to use (presented in Chapter 6). Furthermore, N-S was the most stable FOPL for ranking outcomes and speed across the measures of SES categories (dichotomised education and income, presented in Chapter 5 and 6); and participants in the 'less engaged' predicted class were found to be fastest at completing the rankings, indicating they may benefit most for the N-S FOPL (presented in Chapter 7). These are the population groups that ideally should be most targeted by food and health policies. The discussion from the PPIE session indicated that MTL may have been relatively successful in this experiment due to the familiarity in the British sample, but N-S was agreed by all CYP to be the quickest and easiest to understand (presented in Chapter 8), which are the main aims of FOPLs (presented in Chapter 1). The background literature presented and evidence generated in this thesis show that there is public support for FOPLs in the UK, specifically for the mandatory implementation in the UK, even with CYP (presented in Chapter 1, 6, 8). The analyses that assessed which individuals and groups use FOPLs most effectively, show that these participants were most likely to be older, more educated, female, interested and knowledgeable in healthy eating, have a higher income, and use labels more often (presented in Chapter 5, 6 and 7). The CYP also discussed that N-S would be most effective for people with lower nutritional knowledge (presented in Chapter 8).

Policy Considerations – Meaning, Mechanisms and Policy Options

This thesis provides evidence that the ability to correctly understand FOPLs is not the only factor for label effectiveness and real-world engagement. Both N-S and MTL appear to be effective and suitable in a representative UK population. However, other factors related to engagement are key and my findings indicate that these socially patterned, where more disadvantaged groups have lower positive perceptions and engagement characteristics. The finding that all participants could correctly understand the fundamental message is encouraging, but there were differences between population groups for the other outcomes which indicate that the way it is communicated i.e., the format, is an important aspect for impact on wider engagement. Examining the results in the context of the Grunert and Wills conceptual framework for consumer use of FOPLs, the experiment showed all FOPLs improved the objective understanding when compared to control; but the factors of liking, subjective understanding and perception of the FOPL schemes did differ and therefore difference in label use is anticipated.

The Role of FOPLS in Supporting Equitable Policies

Population-wide engagement is important so that the updated FOPL policy itself does not widen health inequalities. This is raised in the literature, with an analysis of government strategies that aimed to improve nutrition and health, summarising that labels shown at the point of purchasing (including FOPLs) might lead to the exacerbation of disparities, as smaller effects of improving the nutrition of purchases might be seen with disadvantaged groups.[30] The pattern of lower FOPL engagement with more disadvantaged groups, as raised in the government analysis and seen with the baseline food habit and label engagement results (see Chapter 7), is also seen with other downstream policies, as the responsibility is placed on the individual and relies on individual agency.[26] As introduced in Chapter 1, the hierarchy of preventive interventions by effectiveness describes how upstream interventions are more effective as they tend to be comprehensive, reach all parts of the population and are not reliant on individual agency/responses.[28] Another study concluded that population interventions that do not rely on individual agency are more likely to be equitable and effective, but interventions that rely on high individual agency tend to be favoured by governments globally.[27] A systematic review examining the impact of socio-economic position on interventions that promote healthy eating more broadly, found upstream/top down interventions (e.g., focusing on price) appeared to decrease inequalities, whereas downstream interventions (e.g., focusing on person, such as dietary counselling) appeared to increase inequalities.[29] A top down intervention is more likely to be equitable as it is less likely to rely on the ability of consumers to actively choose a healthier product, e.g., the accessibility of healthier foods for low income families, regardless of if they understand guidance and FOPLs when shopping and eating. This also relates to the argument that to encourage healthier lifestyles, these actions need to be the easier choice, or the default option and not require individual agency.[195, 196] These studies collectively indicate that for FOPLs to have the desired impact, the impact on reformulation to improve the overall healthiness of food products available is key (to reduce the reliance on individual agency); therefore, mandatory implementation and monitoring and enforcement is important. This difference is further illustrated by the Nuffield Intervention Ladder presented in Chapter 1, which indicates that FOPL policy can impact level 3 (enabling choice), but can also impact level 4 through reformulation (guiding choice through changing the default by making healthier choices the default option), which may be more impactful.[41] The European Commission (EC) called for FOPLs that offer support to low SES consumer groups in particular and the promotion of education and information campaigns targeted at improving consumer understanding.[197] Academics have similarly suggested that subgroups with lower SES and poorer nutritional knowledge (who have related increased risk of diet-related illnesses) should be efficiently targeted, further suggesting this should be done through simpler formats, such as N-S.[198] Before the N-S label was developed, a cross-sectional study in a large French cohort (NutriNet-Sante, *n* = 38,763) examined understanding and acceptability outcomes of five labels (two endorsement logos: green tick and logo; MTL; simple traffic light; colour range logo). They found that poorly educated individuals favoured the simple label formats, compared to individuals with substantial knowledge who favoured the MTL label; concluding that although the group most favourable to MTL was the largest, the simple formats may be most suitable to target population groups with lower nutritional knowledge or interest in the nutritional quality of pre-packaged foods.[198]

Specifications

The specifications of the FOPL policy will be key for the overall effectiveness. A policy analysis of global FOPL regulations summarised that the following elements are important for effective policies: strategic objectives, clear label and display specifications, a valid scoring mechanism (i.e., label information is based on a validated nutrient profiling criteria), and justification for product category inclusion.[66] The specifications will need to be clear and standardised, which will assist with enforcement of the regulations, improving consumers' ability to see the FOPL and improve their familiarity (which can lead to trust).[37] The standardised placement of the label on product packaging has been shown to assist consumers with engaging with the label, improving their confidence in using, noticing and familiarity,[49, 199] specifically when positioned in the top-right corner.[43, 54] An example of placement specifications is the Chilean WL policy, which suggests the location of the descriptor/s is the upper right corner of the front of the packaging.[146] Logistically, the placement of the FOPL at the top of packaging makes sense, as sometimes in supermarkets the products are not taken fully out of the transport boxes (see

Figure 27 image A for an example). Standard sizing of the label and the colours are also important for recognition and legibility. The size will be particularly important if MTL is chosen, as the text will need to be legible (which will be challenging on some products and for some consumers). The online specifications should also be considered, as more households in the UK are completing their food shopping online – £12.3 billion spent in 2018 (a 7% share of total grocery retail sales, forecast to account for 10% in 2023) and increased since COVID-19;[88, 89] moreover, the UK government stated that the UK FOPL should remain "effective however and wherever we shop".[81] The current presentation of FOPLs online is inconsistent, both within retailers' websites and between online and in-store products.[90] The authors suggested that the colours of the labels need to be consistent and that websites should be designed to ensure the visibility of labels (by raising their vertical position).[90] Another study assessing FOPL presence in online UK supermarkets found that often more than one FOPL was shown, with RI labels or endorsement logos displayed alongside MTL labels.[200] The presence of other labels and claims (as discussed later) is also important to consider, as the variety may cause consumer confusion, thereby defeating the purpose of FOPLs.[197] A real-world example of this is a product from an online Chilean supermarket shown in Figure 27 image B.

Figure 27: Real-world of examples of FOPL issues – a) Supermarket shelves showing interference with products left in transport packaging b) Product with multiple FOPL schemes presented

a)

b)

Implementation

Evidence suggests that the mandatory implementation of the chosen FOPL will be vital for policy effectiveness, both to impact consumer choice and to drive reformulation.[189] Mandatory FOPLs are repeatedly recommended by WHO;[75, 160, 190, 191] and research, including my results, indicate there is public support for a mandatory FOPL in the UK.[201] Results from a modelling study indicated a 2.5% reduction in obesity in Europe with the implementation of mandatory labelling (-4.1% for 25 year olds and -2.1% for 65 year olds).[65] It is important to contextualise this finding by noting that the assumptions for the model were that: labels help consumers choose healthier products (specifically lower fat and increased fruit and vegetable consumption), labels could encourage manufacturers to reformulate products with healthier nutrients/decrease serving size; only two thirds of consumers actively read the labels; they would only be applied to food sold in stores; that retailers would post information on how to use the labels and the benefits of a healthy diet (no other communication); and that the label accuracy would be verified. The cost per capita was also estimated at $2.16, with assumed costs being planning/administration, implementation/evaluation, poster preparation/issuing; but would not include the additional packaging costs to manufacturers.

An additional argument for the mandatory implementation of MTL in the UK is so they can be used as originally intended and outlined in the UK government guidance. The guidance states that they are useful for comparing products within categories to assist the selection of the healthier product, which is not possible if both products do not have the label.[85] The Grunert and Wills framework outlines that you first need to see the label to engage with it.[37] Content analyses indicate that current UK FOPLs are more commonly displayed on healthier products and supermarket brand products, which is a major disadvantage of voluntary FOPL schemes.[87] Mandatory labelling is important for overcoming suboptimal uptake, such as the static uptake of voluntary labels in the UK, and for having the desired impact on product reformulation.[33, 189] A simulation study from Australia estimated the positive impact of the voluntary HSR label on disability-adjusted life years (DALYs) and indicated the impact would be 10 times greater if the scheme was made mandatory.[192] A real-world example of how FOPLs encourage reformulation is Arran Dairies, a company that reformulated their base ice-cream recipe with support of the

Food and Drink Federation Scotland, stating it would "bring it into amber for FOPL."[202] This action removed an estimated 42 million calories yearly from Scottish diets and the company stated "anything we can do to help on obesity and health is a positive outcome for us. We also expect sales benefits from meeting consumer trends in the sector illustrating industry support."[202] Mandatory FOPL legislation would also create a level playing field between manufacturers, as it would no longer be a business decision of whether or not to display them, a sentiment echoed by Action on Salt.[203] The current voluntary system disincentivises participation, e.g., ice cream brand A not wanting to label their ice creams with red MTLs if ice cream brand B and their competitor does not. Reformulation encouraged by labelling may also impact longer term health outcomes, beyond the immediate consumer behaviour outcomes that labels directly impact.[111] The key elements for successful reformulation policies have been summarised as having strong incentives, a tight implementation strategy and a focus on overall nutritional quality of food products, rather than individual nutrients (relevant for MTL, which is a nutrient-specific FOPL).[204]

Policy development

A policy analysis of global FOPL regulations highlighted the need for a transparent and accountable policy cycle including the development, implementation, evaluation and enforcement of regulations.[66] The policy monitoring and evaluation is vital for understanding the adoption, success and areas for improvement.[160] A study that reviewed nutrition and health government strategies highlighted the need for systematic surveillance and evaluation to accompany policy actions, assessing progress and guiding further developments.[30] The evaluation could also be used to assess compliance if the policy was mandatory. The study also highlighted the value of embedding policies that focus on individual responsibility (described as soft policies) within broader, multicomponent food and nutrition strategies, as discussed by CYP in the PPIE session (presented in Chapter 8).[30] Chile's food laws are an example of this as they cover FOPL, marketing and sales restrictions. Academic insight into the integral components for the success of the Chilean WL policy included the minimal involvement of industry in the initial regulation.[205] In academic analyses of the Chilean WL implementation, areas were identified that could be strengthened, including preventing the use of any health claims on 'high in' products.[205] Common industry tactics have been categorised by the World Cancer Research Fund (WCRF) into four main types: delay (e.g., pushing for longer consultation periods, more research, threatening litigation), divide (e.g., arguing for voluntary over mandatory, development and promotion of own labels), deflect (e.g., reframing the issue, claiming they scare consumers, self-regulation works) and deny (e.g., claiming a lack of evidence, discrediting FOPL effectiveness).[187] The co-creator of N-S addressed some of the issues they faced from the food lobby during FOPL development, who called for an unrealistic large and expensive trial (delaying).[206] A Global Health Advocacy Incubator 2021 report titled on 'Big Food's War on Healthy Food Policies' outlines the actions from the ultra-processed food (UPF) industry in this area.[207] They monitored the industry attempts to weaken FOPL policy in 20 geographies, summarising the five main strategies: protecting brand reputation through corporate washing; using multilateral bodies to delay policy implementation and threaten countries with legal or economic issues; shifting blame to individuals; implying their products are healthy and contribute to society while blocking healthy food policies; and finding regulation loopholes to continue the promotion of UPFs.

Educational or Public Awareness Campaigns

The findings of my thesis indicate that the perceptions of FOPLs, including confidence in using (i.e., subjective understanding), liking, seeing, and trusting the label, differ across population groups. Research shows that these perceptions are vital for engaging with the label.[37] A WHO Europe report highlighted the importance of public education initiatives to improve awareness and understanding, which will be vital for any updated or new FOPL policy.[33] The NOURISHING framework developed by WCRF outlines the key areas for food policies to promote healthy diets, where the 'ING' covers the domain of behaviour change communication. A paper by Hawkes et al. examined smart food policies aimed at obesity prevention, and outlined that for the overall success of nutrition labelling the complementary/reinforcing elements should include accompanying public awareness and education campaigns regarding the label.[208] This includes informing people through public awareness (e.g., education about food-based dietary guidelines, mass media, social marketing; community and public information campaigns).[209] The current UK FOPL guidance document outlines some communication options to assist consumer awareness, including: information on websites, mobile applications and social media; instore promotion and information; recipes, booklets and articles; individual or group sessions within wider educational programmes and diet or health counselling.[83] These communication tactics should be combined with a public awareness, marketing or education campaign to target the label perceptions. This would improve FOPL awareness, familiarity, and trust by explaining what the FOPL is, how to use it, the underlying logic, the government endorsement, and supply links to further information and resources. The aim of these campaigns should be to empower, familiarise and educate, but not to shame consumers as this would be counterproductive.[210] PPIE sessions to co-develop the overall messaging of any campaigns would be beneficial, to ensure that the messaging and campaigns are not stigmatising.

This need for education has also been expressed by consumers, with survey results following the implementation of the WL in Chile finding that 34% of consumers noted more education on the new label was needed;[211] and a qualitative study with mothers and children four years post-implementation found reinforcement campaigns were needed to maintain their effectiveness.[212] Results from experimental studies illustrate the potential effectiveness of communication campaigns. The provision of in-aisle signs explaining FOPLs was found to be somewhat helpful in assisting consumers to make more healthful dietary decisions.[126] Another experimental study found that the recall and understanding of an early iteration of N-S was highest in the group that received a communication leaflet (explaining the introduction of the FOP nutrition label, the way the FOP nutrition label could be used to compare products, and the way the colour is attributed). This group also had a significantly higher nutritional quality for the sweet biscuit category compared to the control group, in purchases from an experimental supermarket.[176]

Enhancing FOPL success

To enhance the effectiveness of FOPLs, researchers have suggested the assistance of technology, such as the use of mobile applications and improving online supermarket design. The rollout of a complementary mobile application could help with identifying healthier options and the tailoring/personalisation of nutritional information for consumers (e.g., nutrients to avoid due to medical conditions). The FoodSwitch application developed in the UK by Action on

Salt, Action on Sugar and The George Institute allows consumers to scan barcodes to see how healthy a product is and opt to switch to a healthier alternative in the same product category.[188] Another mobile application is a French application called Yuka, which is independently owned and allows consumers to scan food and cosmetic barcodes to get a Yuka score (with colour-coding), also presenting alternative healthier options.[213] The Yuka score correlates to N-S, additionally accounting for additives and if the product is organic. An online news article suggests that Yuka is popular with consumers and shows that it led directly to the reformulation of 900 products, by Intermarché (a French supermarket retailer), who wanted to achieve better Yuka scores.[214] For online supermarket design improvements, studies have suggested the aggregation and integration of nutritional information to prompt healthier choices.[89] For example, online supermarket designs could present the overall healthiness of online shopping baskets or present alternative healthier options throughout the shopping experience, which research has shown to be effective.[215, 216]

Additional Nutritional or Health Information On Product Packaging

Health and nutrition claims can impact consumers' understanding of product healthiness and are commonly used as shortcuts to identify healthier options.[217] For the online experiment described in this thesis, the steering group decided to exclude any health or nutritional claims on the packaging, allowing the standardisation of product characteristics and ensuring that the ranking ability was testing the FOPL only (there was also no mandatory back of pack information). In real-world settings, these claims are present, therefore, their interaction with FOPLs is important to understand. A systematic review examining consumer responses to health claims with other nutritional information found that FOPLs can reduce the bias caused by health claims.[52] A qualitative study exploring the motivations and contexts for using nutrition information with UK-based females found that some participants reported using nutritional claims to bypass the cognitive effort of interpreting nutrition information, such as 'reduced fat'.[50] The nutrition information can also reduce bias but only if it is viewed and correctly interpreted (which is rare). In the UK, regulations exist for some nutrition claims (the presence or absence of nutrients e.g., 'low fat', or comparison with another product 'reduced fat'), health claims (a link between the product and health e.g., "reducing consumption of saturated fat contributes to the maintenance of normal blood cholesterol levels"), and nutrition labelling for fortified foods (e.g., 'fortified with vitamins and iron').[218-220] A register of all nutrition (n = 30) and health claims (n = 2550) is maintained by the UK government, with the wording and authorisation overseen by the UK Nutrition and Health Claims Committee.[221] However, not all claims are covered, including dietary suitability claims (e.g., 'V' for vegan) and nutrition claims such as 'one of your five a day', which was viewed as out of scope (see Figure 3 for real-world examples).[218, 222] There is a government developed '5 A Day logo' which requires a licence from PHE and for products to meet the recommended portion size of 80g vegetable or fruit (see Figure 3 image B).[223] However, manufacturers also make their own claims, which are not covered (see Figure 3 Image A) and research shows most UK products with these claims do not meet the requirements.[224, 225] Having competing claims is understandably confusing for consumers and a systematic review found that health claims can lead to a 'health halo' effect, where a food appears healthier than it is and consumers then consume more of the product.[204] To my knowledge there are no specifications on the physical size of health claims on packaging in proportion with FOPL, meaning it is possible they could detract attention away from FOPL. Two real-world examples of this are shown in Figure 28. Image C shows a product presenting MTL, nutrition labelling ('fortified with vitamins and iron') and a dietary suitability claim ('V' for vegan). The MTL text is illegible, but the nutrition labelling of the fortification is clearly presented. The label uses the same colour green as the MTL and it resembles a tick/happy figure despite it being high in sugar. Image D shows an example of a product with a green dietary suitability claim, while the MTL indicates the product has 50% of the daily recommended intake of sodium.

Figure 28: Real-world example of a) product packaging with MTL and manufacturer made 'HEALTHY CHOICE – 1 OF YOUR 5 A DAY' b) UK government created '5 A Day logo' c) product packaging with MTL, nutrition labelling 'fortified with vitamins and iron', and dietary suitability claim 'vegan' d) product packaging with MTL showing and dietary suitability claim 'vegan'

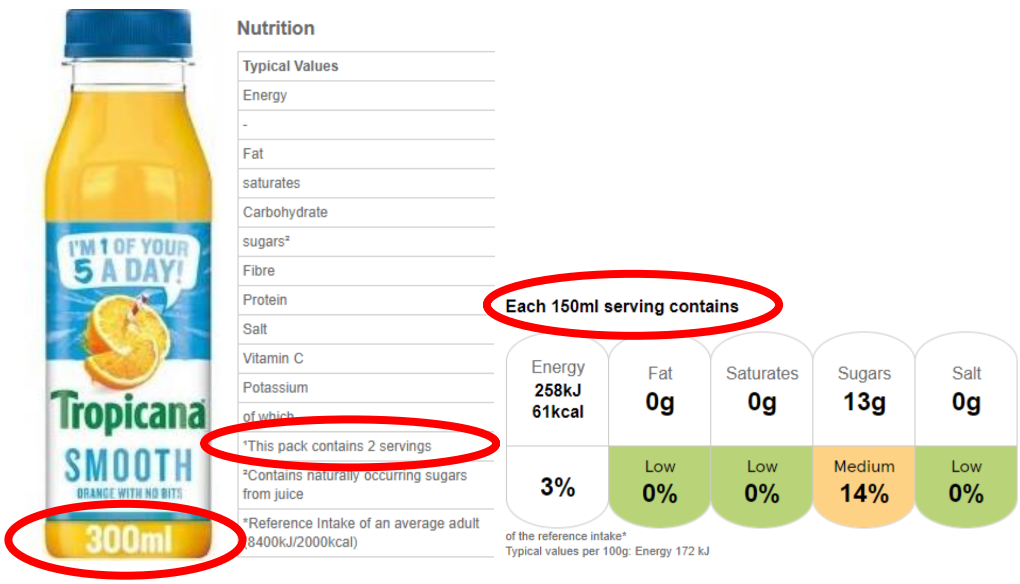
Label Schemes Considerations

Multiple Traffic Light

The MTL is a nutrient-specific interpretive label, indicating if a product is low (green), medium (amber) or high (red) in fat, saturated fat, salt and sugar via text and corresponding colour.[81] The MTL was designed to convey this information 'at a glance', to enable comparisons between the same product types and assist the identification of a healthier option.[86] Arguably these aims are not being achieved as the voluntary implementation means it is not on all products and therefore, it is neither possible for consumers to compare products nor can it be interpreted 'at a glance'. Additionally, the prevalence of colour-blindness is important to consider (approximately 4.5% in the UK, mostly males), as the text specifying the nutrient level and percentage of recommended daily intake is small and not legible for all consumers, thereby relying more heavily on the information being communicated correctly through the colours.[226] The current awareness of MTL is low considering that it has been the endorsed label in the UK since 2013, with research from 2021 showing only 79% of participants reported knowing what MTL is for.[91] The consultation document stated "seven years on, the MTL has proven to be a success with 9 in 10 shoppers agreeing it helps them to make informed decisions when purchasing food" using results from a 2018 online market research study (n = 2,121 UK adults).[227] This is not consistent with other research, with a tracking study from the Institute of Grocery Distribution finding that use of MTL dropped from 58% in 2019 to 54% in 2020.[228]

MTL was not designed to "demonise foods with lots of reds" but research shows that consumers do feel pressure to avoid the reds.[86, 229, 230] A negative/avoidant label may not be the best option if it contributes to the weight stigma cycle (positive feedback loop where weight stigma causes weight gain).[210] A study examining the experience of people living with Type 2 diabetes using MTL showed that red traffic lights appeared to over emphasise the need to avoid those foods and their link with danger (e.g., feta as a specific example), and a sentiment that advanced maths skills were needed to interpret information about portion.[229] There is also ambiguity with nutrient-specific interpretive FOPLs that consumers face, for example if a product is high in saturated fat but low in sugar, they may then need to decide if that is healthier or less healthy than the opposite. Portion size issues were also raised in the Eatwell Guide booklet food labelling section, with a common misalignment between suggested portions and what is actually consumed (e.g., food and drinks viewed as single servings having nutritional information presented per half pack used on labels), a point raised by the CYP in the PPIE session.[84] A real-world example of this is a 300ml bottle of orange juice that is commonly included in meal deals from major retailers (i.e., single-serve implication), but presents the serving size as 150ml (see Figure 29). Qualitative research further describes the difficulties experienced by consumers with portion sizes "I don't like [it when] they say a quarter of a packet of [crisps]...contain[s]this many calories, I [have to then] stand there and calculate how many in the whole bag and work out how [many calories] that would mean per [real-world] serving".[50] These issues with portion sizes on MTL could negatively impact overall trust in the labels and therefore impact engagement. The information presented by MTL focuses on nutrients that are of concern or to avoid, rather than nutrients to target or increase such as protein, fibre or fruit and vegetables. Evidence suggests that focusing on the composition of the whole product, rather than specific nutrients, is important to ensure an improvement in food composition, as nutrient-specific reformulation can lead to negative change in other nutrients (such as increases in additive or fat to counter-act decreases in sugar).[204, 231, 232]

Figure 29: Real-world example of a product commonly consumed as a single serving but labelled as 2 servings[233]



Nutri-Score

The N-S is a summary indicator label, providing an aggregated score on a scale of A (green, higher nutritional quality) to E (red, lower nutritional quality).[145] It assesses the overall nutritional quality of the product by offsetting negative/unfavourable nutrients (sugar, calories, saturated fat, sodium) by accounting for positive/favourable nutrients (fruit, vegetables, nuts, fibre, protein, walnut, rapeseed and olive oils). The score is standardised per 100g/100ml across all products, compared to MTL which can display the information per portion (which, as discussed, is not always representative). Research suggests focusing on the composition of the whole product, to ensure an improvement in food composition rather than nutrient-specific focus.[204] This could be useful for industry reformulation as fibre and protein could be increased to improve the overall score rather than just focusing on decreasing unfavourable nutrients, such as fat, salt or sugar. The results presented in this thesis show that N-S performed the best across most outcomes including objective understanding and speed of ranking. The latent class analysis and descriptive analyses indicate that this label may work better across the entire population as there were minor differences observed between the highest and the lowest education and income groups. It is also an intuitive label, a point raised during discussions with CYP (presented in Chapter 8) and observed from the speed of ranking results (presented in Chapter 6). The original goal of the UK FOPL policy is to provide nutritional information "at a glance," and there is evidence to suggest that N-S is the most suitable FOPL based on this criterion. International academics have described N-S as an evidence-based health policy tool, due to the strength of the evidence.[234] [235]

Warning Label

The WL is a nutrient-specific interpretive warning label that applies 'high in' warnings for salt, sugar, fat, and calories, if products exceed a specified threshold. Again, there is ambiguity with this FOPL as it is nutrient-specific interpretive, for example, consumers must decide if a product that is high in sugar is healthier than a product that is high in sodium. The results of this thesis showed it was the third best performing label, but the Chilean WL still provides useful lessons for policy, including the development, implementation, and impact. A strength of their labelling policy is that it is a part of a broader Chilean food law that applies to labelling, sales restrictions and advertising regulations, a recommendation raised by the EC.[197] In Chile, products that exceed the thresholds for any of the nutrients are banned from sales at schools and subject to advertising regulations between 6am and 10pm on TV or movies and regulations for product placement, online, radio, billboards (but not sporting events, Article 110).[146] The implementation of the policy was staggered, with three threshold stages over 36 months that became increasingly stricter. The label design underwent extensive research and development (including a literature review, qualitative studies with public and experts and quantitative survey).[205] The specifications are clearly set out in government documents, relating to the components (the inclusion of 'Ministry of Health'), design (font, colour, shape, size) and presentation (order of labels and standardised distance between, with suggested placement in top right corner of the front of the packaging), see Figure 30.[146] The inclusion of 'Ministry of Health' is important to make clear it was a government health warning and not from the manufacturer (can be replaced by "MINSAL" if descriptor is the minimum size, 1.5x1.5cm).

Figure 30: Images from the Chilean government specifications a) Example of the elements in the descriptor "HIGH IN SUGARS" (page 52) b) Examples of label location on different containers (page 56)

a) b)

A key aspect of the labelling policy roll-out was the enforcement on retailers rather than manufacturers.[236] The sanitary regulatory office of the Ministry of Health (SEREMI) was responsible for the monitoring of the policy, which included visiting retail stores and issuing fines to any who stocked non-compliant products.[205] The compliance in retail environments was high (88% in 2019 and 89% 2020) and the positive impact on product reformulation was significant.[237] Evaluations showed that between 15-18% of products were reformulated and the proportion of products with 'high in' labelling dropped from 52% pre-implementation (2015-16) to 42% post-implementation (2017).[67, 68] Calorie content was reduced in 23% of products, with an average reduction of 2.8%.[238] Sugar content was reduced in 11.5-25% of products (depending on the timepoint), impacting up to 33% of products, specifically dairy products, conserves, sugary beverages and breakfast cereals.[237-239] Sodium content was reduced in 5-10% of products, and products with 'high sodium' labels dropped from 74% pre-implementation to 27% post-implementation, mostly for spreads, cheeses, ready-meals, soups and sausages.[237, 240] The impact on consumer behaviours showed that household purchasing of 'high in' beverages decreased significantly from 2015 to 2017 by 23.7% or 22.8 mL/capita/day.[141] Encouragingly, the absolute reduction was similar in household purchases by education level (more vs less educated, but less educated purchasers had a higher baseline; and therefore, higher percentage decrease). Household food purchasing over the same period showed an overall decrease in purchases for calories (3.5%), sugar (10.2%), saturated fat (3.9%), and sodium content (4.7%).[241] The impact on food purchasing before and after showed the likelihood of purchasing 'high in' breakfast cereal decreased by 12.5%, but there was no difference for chocolate or cookies.[132] Research also indicates consumers may be willing to pay more for processed meat products without 'high in' labels.[242] Qualitative research evaluating the impact on consumer behaviour found that consumers with higher incomes used the labels more (as a quick shortcut); and mothers were aware of the regulations and thought foods without the WL were healthier.[243] A qualitative study following implementation in Uruguay showed that 37% of participants reported that WL had not influenced their purchasing decisions, with reasons including: disinterest, attitudes, lack of perceived self-efficacy, cognitive biases and limited rationality when making purchase decisions.[244]

Positive Choice Tick

The PC is an endorsement logo label that applies to products that meet a pre-specified threshold. The thesis results show this was the worst performing label across all outcomes, even when accounting for the binary nature (the healthiest product outcome analysis in Chapter 6). Due to the binary nature, it is limited in the information it can portray and the product categories it can represent. These limitations impacted the study design in two ways: firstly, the label did not provide sufficient information to rank three products; secondly, for two of the food categories (crisps, cakes) none of the products qualified for use, which was reflected in the comparatively low proportion of participants in the PC group who reported seeing the label (compared to the other FOPL groups). These issues are reflective of the real-world application and limitations of the PC label, as many products would not qualify and, therefore, would not have a label or provide any information to consumers. The PC label was included as endorsements logos are favoured by the food and beverage industry, due to the focus on the promotion of healthier products rather than discouraging less healthy products. Researchers have warned that they may lead to price premiums, which would negatively impact low SES groups,[245] and 'health halo' impacts, leading to the overestimation of product healthiness and over-consumption.[33] The PPIE findings showed that CYP did not find the PC label helpful, as it did not provide much information (no explanation) and they did not think it allowed the consumer to make their own decision about 'healthy foods'.

Future Thinking

Beyond the nutrients typically displayed on FOPLs, there have been suggestions for the labelling of the ultra-processed content/level of food processing.[246] In the UK, consumption of UPFs is the highest in Europe (and increasing).[247] High UPF intake is associated with weight gain, increased CVD risk and all-cause mortality.[248-250] Analyses of a large Italian prospective cohort (Moli-sani) assessing mortality risk with diet quality (as indicated by N-S classification) and UPF intake (as indicated by NOVA classification) found that adults with the lowest quality diet and highest intake of UPF had the highest mortality risk; this is mostly explained by the high UPF rather than lower N-S scores/poor diet quality.[246, 251] Vital Strategies, a public health foundation, suggested adding UPF warnings to existing or new FOPLs (see Figure 31).[252] There is real-world evidence that FOPLs can lead to reformulation that increases the ultra-processed content, such as decreasing the amount of sugar by replacing it with low calorie/artificial sweeteners.[253, 254] Other swaps include fat with modified starches, addition of extrinsic fibre, vitamins and minerals, which is not seen as a solution, especially if these products are then promoted as 'healthy' or premier products.[255] This nutrient-specific reformulation may be counter-productive if ultra-processing is itself harmful.[256] The preferred solution is a whole food approach to reformulation, or the development/promotion of less-processed options (e.g., rolled oats rather than heavily processed breakfast cereals).[256] Following the publication of two related articles, the BMJ suggested that policy makers should shift their priorities away from food reformulation that could position UPF as a solution to dietary problems and instead focus on the availability, affordability and accessibility of unprocessed/minimally processed foods.[248, 249, 257] This echoes the opinion of the CYP on the importance of improving the accessibility and availability of healthy food.

Figure 31: Example provided by Vital Strategies of an integrated UPF label with existing WL FOPL

FOPL Limitations

The use of FOPLs as a policy intervention to improve population health has limitations. Firstly, improving consumer understanding of product healthiness (one of the two FOPL mechanisms), is a downstream action which is less effective as it puts the responsibility on consumers and therefore low SES groups are less likely to engage, which risks the widening of health inequalities. The other FOPL mechanism is product reformulation, which is an upstream action. Regardless of the FOPL displayed, this is just one piece of information that is used by consumers to make a food purchasing decision. There are other frameworks outlining factors influencing behaviour and COM-B is one such model. Applying the COM-B system to FOPL use, the labels can improve the *capability* of consumers and having consistent labelling can provide the *opportunity* to use them, but *motivation* is still needed.[39] This is especially important for low SES groups, where cost is a major determinant, so choosing the label that appeals to most consumers, especially low SES groups is important.[258] Other factors of the food environment should also be addressed to support healthier food purchase and consumption decisions. Looking at expected longer term outcomes gives some insight into how FOPLs perform in real life conditions over the longer term, therefore taking into account the broader context in which FOPLs are operating in. A macro-simulation study using estimates from a laboratory experimental economics test on the nutritional quality of household purchases, estimated the impact of diet modification due to FOPLs (N-S, HSR, MTL, Reference Intakes (RI) and SENS, similar) on mortality from non-communicable diseases (NCDs). The addition of FOPLs decreased diet-related mortality, with the largest effects seen for N-S (3.4%), HSR (2.8%), RI (1.9%), MTL (1.6%), SENS (1.1%).[259] Modelling by Sassi et al. indicated that the impacts of labelling on decreasing obesity were comparable to fiscal measures (specifically fiscal measures modelled to increase/decrease the price of high fat foods/fruit and vegetables by 10%) and individual treatment (general practitioner led dietitian counselling), but moderate compared to food advertising restrictions.[65] In the UK, HFSS advertising and promotions regulations have been approved, but the implementation delayed. These scheduled regulations will likely assist in making any FOPL policy more effective and improve the obesogenic environment more broadly.[193] The removal of advertising and price promotions would be beneficial for FOPL effectiveness as they both of these factors impact on consumer food purchasing and consumption decisions. Research shows that price is a major determinant of product choice, especially for low SES groups.[33] Analysis of the impact of price promotions found that they are effective at influencing preferences and purchases, encouraging consumers to buy more than they intended, which increases consumption as the products are not stockpiled.[260-262]

Recent Developments

Since the experiment was conducted, other studies have been conducted that corroborate the findings from this work. These include a secondary analysis of data from the FOP-ICE (Front-Of-Pack International Comparative Experimental) across 18 countries (including data from the Egnell et al. study)[69] that showed, in a product choice task and a ranking task, that N-S performed the best (in all 12 tasks), followed by MTL, (in 6/12 tasks), compared to the non-interpretive RI label.[263] A recent review of the evidence found that from 134 articles that N-S was effective in 83% of studies (110 original articles including modelling studies; experimental studies testing understanding, impact on food selection/purchases; and 24 reviews); the probability of an article finding non-favourable results was 21 times higher if the authors had declared a conflict of interest or if it was funded by the food industry.[264] This illustrates the importance of transparency and assessing conflicts of interest and funding, to prevent private sector interests in diminishing scientific findings. In France, the N-S scoring system was recently updated based on recent evidence, to improve the appropriateness of scoring for non-water drinks, a process that was reported transparently.[265] This revision is important to consider for the nutrient-profiling scoring system chosen in the UK, to ensure the score and information provided by the FOPL makes sense and can be trusted (including the scientific process), which will be paramount to its success.

Other work has further explored the impact of literacy. An online survey tested objective understanding and perceptions of FOPLs (N-S, MTL and RI) in tertiary education students with varying health literacy. The N-S label performed best for objective understanding across all health literacy levels. Cluster analysis of FOPL perceptions by participant characteristics showed that the 'N-S preferred' cluster had the highest percentage of disadvantaged students (low estimated nutrition knowledge, diet quality, inadequate HL).[266] An online RCT tested N-S, RI and no label, low-income adults asked to perform a shopping task in an experimental online supermarket – assessing overall nutritional quality of the shopping cart and nutrient content (n = 336). N-S led to the highest overall nutritional quality (significantly compared to RI, but not significantly lower than no label), also significantly lower calories and saturated fat than RIs.[267] Another online RCT tested the impact of N-S on the nutritional quality of individuals' hypothetical weekly grocery shopping (they were asked to carry out their weekly grocery shopping and entered a lottery to win the product they chose during the task), using a Chrome web browser extension to display the N-S next to each product. Compared to control, the N-S group selected products with a less sugar and saturated fat and higher nutritional quality (-3.3%, -7.5%, +8% on the healthy trolley index, HETI). Additionally, they found that users with low food literacy (assessed with a validated questionnaire), seemed to particularly benefit from the N-S label (-10.5% sugar, -5.5%, +11% HETI).[268]

Thesis Strengths and Limitations

This thesis has a number of strengths, including the close collaboration with policymakers for the experiment design and the appropriately powered and high-quality experiment with a pre-specified protocol and analysis plan. It involved a large representative sample of adults in GB, included a control group and was broadly comparable to the full NatCen panel and the British population.[143, 144] The strengths of using the NatCen panel include the experience the panel members have with taking surveys, which limited the non-response and missing data rates. The survey design and assistance with the survey questions also improved the data collection and the reliability of the findings by reducing bias. However, there were some limitations including the sample and design decisions. The NatCen panel is nationally representative, and weights are applied to adjust the representation and diversity of the panel but it is still not completely representative of the UK population (e.g., it does not include Northern Ireland). The methodology of the survey also meant that participants needed to complete the experiment online, have internet access, have smartphone/computer access and be able to speak English. Therefore, participants without this access or ability were automatically excluded. NatCen conducts approximately 15% of their surveys with participants over the phone, meaning that the study design would have excluded some of the older panel members or who are less experienced with technology or inclined to use it. Approximately 1.3% of the UK population report that they can speak English but not well and 0.3% say not at all, and research indicates they are more likely to have poor health.[269, 270] To complete the survey, a reasonable grasp of English and literacy was needed, which would have excluded some people. I assume that for consumers with English as a second language, that they would most likely struggle with MTL or WL, due to the reliance on text. Mock images were created for all products to reduce bias and variability, but this may have reduced ecological validity. No back of pack information was included, in contrast to some experimental studies and actual products, as the experiment was designed so the ranking was exclusively informed by FOPL. In a real-life context, people would have access to additional information (both on the front and back of pack), so the limited information provided may have underestimated their ability to rank products. However, this is unlikely to be substantial, as evidence shows that people rarely read the back of pack.[33, 37] Evidence from an observational study in the UK showed only 12% of observed consumers looked elsewhere than the front of packaging.[38] The proportion of participants correctly ranking the products at baseline varied across food products and was markedly low for yoghurt. This led to large increases in correct ranking at follow-up for this category and indicates that yoghurts were harder to rank without a FOPL. Despite all images being designed to allow correct ranking, it is still possible that the yoghurt images may have had some limitations as indicated by the difficulty participants had in ranking the products at baseline. This difficulty could also be reflective of the health and nutrition claims (e.g., ‘low in fat’) that normally feature in this product category.

For the perception questions, participants only responded regarding the FOPL condition they were randomised to and had just used in the ranking task. Therefore, it was not possible to compare preferences between FOPL schemes in the experiment, but it was possible to explore this in the PPIE session. This was partly to reduce the time burden on the participants, but to ensure responses were based on the actual use of the label. In the experiment, no FOPL schemes were explained to participants, but we assume MTL was familiar to most participants (79% according to recent survey data), thereby giving an advantage compared to the other labels. The familiarity and trust in the label would have helped participants in the ranking task and their perceptions.

The overall limitations of this thesis include that the primary data generated were from an online experimental study, testing the objective understanding of comparative product healthiness and was not based on objectively measured consumer behaviours. The experiment was also limited, in that we were unable to conduct qualitative interviews with participants who took part in the experiment, due to cost limitations. A logical extension to the current work would be an experimental study using a laboratory mock supermarket to allow better data on consumer interactions with the FOPLs, but this was not possible with the resources and is addressing a different research question. To more robustly test the impact of FOPLs on the likelihood of improving the nutritional quality of purchase and consumption would require a real-world randomised controlled trial, or a before and after observational study. For example, an ideal experiment would test FOPLs applied to real products with qualitative interviews with consumers immediately following the purchase, to understand their thought-process; and delayed interviews, allowing time for the consumption of products. Results from this type of experiment would improve the confidence in the findings and real-world validity, as they would be objectively measured outcomes (purchase and consumption) with the inclusion of external factors (including price and health claims etc.). Although, a ranking task study testing objective understanding as requested by DHSC in this real-world or research laboratory setting may not have made sense. Additionally, since this study was conducted during COVID-19, in-person research was limited, and real-world studies are typically costly and impractical.

The primary outcome analysis was conducted blinded by an independent statistician (DR). The analysis sample for global food score was limited to 44% (*n* = 1,976), as it was pre-specified that only participants who had consumed or purchased the product within the last 12 months would be included. Due to the low level of instant hot chocolate consumption or purchase, this category was excluded from the global food score. Sensitivity analyses showed that results did not change when all participants were considered. The study was powered on the primary aims, and so the statistical testing of the secondary aims was limited. The formal comparison between FOPL schemes was limited to N-S and MTL due to the policy relevance (this was pre-specified). Data from the software were used for the timing outcomes and the participants were given the following instructions: "we would like to know how long it takes for you to rank the different foods, so please click 'next' as soon as you have completed each task." These directions were not provided to rush participants, but to make them aware that this data would be collected and therefore to encourage them to move to the next page once they had finished ranking the products. Timing was not the primary outcome and, therefore, care needs to be taken with the interpretation of the findings. BMI was calculated using self-reported height and weight, so care needs to be taken as the self-reported BMI is often inaccurate, with height overestimated and weight underestimated.[271] The general limitations of the thesis are that the experiment was not testing actual label use but instead the ability to understand FOPLs correctly when comparing the healthiness of products. However, participant perceptions provided insight into factors that will impact engagement. The PPIE session with CYP provided further insights, as the experiment findings and thoughts were discussed and explained, but conducting it with CYP instead of the NatCen panel who completed the main experiment is a limitation. Although, the 2-hour session with CYP did allow an in-depth discussion and a complementary view of FOPLs and UK policy options, that would not have been possible with the NatCen panel due to time and budget constraints.

Next Steps and Future Research Directions

The results of this thesis will be used to inform future FOPL policy developments in the UK. They have been delivered through online presentations to the steering group members (including DHSC, PHE and devolved administrations) and through three briefing papers written specifically for policy makers/ non-academic audience (preliminary, exploratory analyses, final report). The UK government conducted a public consultation at the same time as the experiment was conducted, but the results of the consultation have not yet been published. The next steps include waiting for the results of the consultation and the policy decisions, specifically the decision on the implementation (mandatory or remaining voluntary) and the FOPL scheme (staying with MTL, changing to N-S, or an alternate option not assessed). During the PPIE session with the CYPs, following the presentation of the experiment findings, they suggested a combined MTL and N-S label. This potentially warrants more thought but identifying the 'perfect' FOPL would require more research, thereby delaying any action, therefore this needs to be balanced pragmatically. One option would be incorporating the MTL interpretation on the back of pack labels, so that this information was still available to the UK consumers that currently like and use this label, while adopting a summary indicator style as the FOPL. The choice of the label needs to be carefully considered, as highlighted by an academic involved in the HSR label in the WCRF report on FOPL, stating "It's that real risk of if you implement a system that's not effective, it's really hard to undo it."[187] The monitoring/evaluation and public education campaigns will be crucial for the success.

Reflecting on the logic model by Crockett et al.,[32] this thesis has addressed the short-term outcome of 'greater understanding of healthiness of food items' (Chapter 5) and examined factors that will impact on use outlined in the Grunert and Wills framework (Chapter 6, 7, 8).[37] The areas that this thesis has not addressed are: the actual use of FOPLs to improve the healthier food purchasing and consumption, the impact on reformulation and the impact on healthiness of diets and on obesity. All of these areas require more research, such as large-scale randomised controlled trials applying FOPLs to products and tracking the purchases with follow-up or qualitative studies assessing the impact of the FOPLs on consumption. Alternatively, and more realistically, all of these outcomes could also be assessed and evaluated following implementation of the updated FOPL policy, through before and after studies focusing on uptake/implementation, product reformulation and purchases/sales. Thinking about other future research areas and useful next steps that could aid in improving the effectiveness of FOPL are the following main themes: a PPIE session or co-development study focusing on the messaging on any FOPL campaigns; research assessing and informing the policy specifics; research understanding the public/media response; and research understanding the industry response and product reformulation to ensure that any FOPL policy is improving the overall healthiness of foods.

Thesis Conclusions

The results of this thesis provide original and policy relevant findings for the UK evidence base. The findings indicate that FOPLs are effective at improving the understanding of healthiness in a British sample. FOPLs were positively perceived and there was strong support for a mandatory label in the UK. Interpretive labels that use colour were most effective at improving understanding, and there is evidence to suggest that a summary indicator label may be best overall, especially for low SES groups. The results indicate that improving participant perceptions of FOPLS (including liking, familiarity, awareness and trust) will be crucial for their success in the UK, which could be achieved through communication/education campaigns. FOPLs will not independently improve the nutritional quality of diets and by extension obesity, but it is an important aspect of reshaping the food environment into a supportive environment for healthier choices and lifestyles. This work was completed collaboratively with DHSC and will be important for the UK-evidence base and any future FOPL policy decisions.

References

1. Vos, T., et al., *Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019.* The Lancet, 2020. **396**(10258): p. 1204-1222.
2. WHO. *Noncommunicable diseases*. 2022 19 January 2023]; Available from: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.
3. Foundation, B.H. *UK Factsheet*. 2022 19 January 2023]; Available from: https://www.bhf.org.uk/-/media/files/research/heart-statistics/bhf-cvd-statistics---uk-factsheet.pdf.
4. Institute for Health Metrics and Evaluation. *Global burden of disease, cause of death, England.* 2019; Available from: http://ihmeuw.org/5gqe.
5. House of Commons, *Obesity statistics - Research Briefing*. 2023. p. 22.
6. NHS Digital, *Health Survey for England, 2021*. 2022.
7. Public Health England, *Guidance, Health matters: obesity and the food environment*. 2017.
8. Bennett, J.E., et al., *Contributions of diseases and injuries to widening life expectancy inequalities in England from 2001 to 2016: a population-based analysis of vital registration data.* The Lancet Public Health, 2018. **3**(12): p. e586-e597.
9. Marmot, M., et al., *Fair Society, Healthy Lives – The Marmot Review: Strategic review of health inequalities in England post-2010*. 2010: UK. p. 242.
10. Office for Health Improvement & Disparities, *Guidance: Health disparities and health inequalities: applying All Our Health*, UK Government, Editor. 2022.
11. Office for National Statistics, *Health state life expectancies by national deprivation deciles, England: 2018 to 2020*. 2022.
12. NHS Digital, *National Child Measurement Programme, England, 2021/22 school year*. 2022: London.
13. NHS Digital, *Statistics on Obesity, Physical Activity and Diet, England 2021*. 2021.
14. Adams, J., *Addressing socioeconomic inequalities in obesity: Democratising access to resources for achieving and maintaining a healthy weight.* PLOS Medicine, 2020. **17**(7): p. e1003243.
15. Public Health England, *Government Dietary Recommendations: Government recommendations for energy and nutrients for males and females aged 1 –18 years and 19+ years.* 2016: London. p. 12.
16. Food Standards Agency and Public Health England, *NDNS: results from years 7 and 8 (combined)*. 2018.
17. Public Health England, *National Diet and Nutrition Survey: assessment of dietary sodium, adults in England 2014*. 2016.
18. Public Health England, *Calorie reduction: The scope and ambition for action*. 2018.
19. Scarborough, P., et al., *Modelling the impact of a healthy diet on cardiovascular disease and cancer mortality.* J Epidemiol Community Health, 2012. **66**(5): p. 420-6.
20. Department of Health and Social Care. *Tackling obesity: empowering adults and children to live healthier lives*. 2020 February 5, 2021]; Available from: https://www.gov.uk/government/publications/tackling-obesity-government-strategy/tackling-obesity-empowering-adults-and-children-to-live-healthier-lives.
21. Romieu, I., et al., *Energy balance and obesity: what are the main drivers?* Cancer Causes & Control, 2017. **28**(3): p. 247-258.
22. Lake, A. and T. Townshend, *Obesogenic environments: exploring the built and food environments.* J R Soc Promot Health, 2006. **126**(6): p. 262-7.
23. House of Commons, *Obesity - Research Briefing*, H.o.C. Library, Editor. 2021. p. 73.

24. Rose, G., *The objectives of preventive medicine*, in *Rose's Strategy of Preventive Medicine*, G. Rose, K.-T. Khaw, and M. Marmot, Editors. 2008, Oxford University Press. p. 0.
25. Hall, K.D., et al., *Quantification of the effect of energy imbalance on bodyweight.* Lancet, 2011. **378**(9793): p. 826-37.
26. Williams, O. and S. Fullagar, *Lifestyle drift and the phenomenon of 'citizen shift' in contemporary UK health policy.* Sociology of Health & Illness, 2019. **41**(1): p. 20-35.
27. Adams, J., et al., *Why Are Some Population Interventions for Diet and Obesity More Equitable and Effective Than Others? The Role of Individual Agency.* PLoS Med, 2016. **13**(4): p. e1001990.
28. Capewell, S. and A. Capewell, *An effectiveness hierarchy of preventive interventions: neglected paradigm or self-evident truth?* J Public Health (Oxf), 2018. **40**(2): p. 350-358.
29. McGill, R., et al., *Are interventions to promote healthy eating equally effective for all? Systematic review of socioeconomic inequalities in impact.* BMC Public Health, 2015. **15**: p. 457.
30. Mozaffarian, D., et al., *Role of government policy in nutrition-barriers to and opportunities for healthier eating.* Bmj, 2018. **361**: p. k2426.
31. The Behavioural Insights Team, *Putting health in the spotlight: Quantifying the impact of obesity prevention policies in the UK*. 2022, Impact on Urban Health,. p. 14.
32. Crockett, R.A., et al., *Nutritional labelling for healthier food or non-alcoholic drink purchasing and consumption.* Cochrane Database of Systematic Reviews, 2018(2).
33. Kelly, B. and J. Jewell. *What is the evidence on the policy specifications, development processes and effectiveness of existing front-of-pack food labelling policies in the WHO European Region?* Health Evidence Network synthesis report 61 2018 February 5, 2021]; 66]. Available from: https://www.euro.who.int/__data/assets/pdf_file/0007/384460/Web-WHO-HEN-Report-61-on-FOPL.pdf.
34. Department of Health and Social Care, *Technical guidance on nutrition labelling* 2017: London.
35. van der Bend, D.L.M. and L. Lissner, *Differences and Similarities between Front-of-Pack Nutrition Labels in Europe: A Comparison of Functional and Visual Aspects.* Nutrients, 2019. **11**(3).
36. Kanter, R., L. Vanderlee, and S. Vandevijvere, *Front-of-package nutrition labelling policy: global progress and future directions.* Public Health Nutrition, 2018. **21**(8): p. 1399-1408.
37. Grunert, K.G. and J.M. Wills, *A review of European research on consumer response to nutrition information on food labels.* Journal of Public Health, 2007. **15**(5): p. 385-399.
38. Grunert, K.G., et al., *Use and understanding of nutrition information on food labels in six European countries.* Journal of Public Health, 2010. **18**(3): p. 261-277.
39. Atkins, L. and S. Michie, *Changing eating behaviour: What can we learn from behavioural science?* Nutrition Bulletin, 2013. **38**(1): p. 30-35.
40. UNICEF, *Policy brief: Front-of-pack nutrition labelling of foods and beverages*. 2021. p. 19.
41. Nuffield Council on Bioethics, *Public health: ethical issues.* 2007, Nuffield Council on Bioethics.: London.
42. Blitstein, J.L., J.F. Guthrie, and C. Rains, *Low-Income Parents' Use of Front-of-Package Nutrition Labels in a Virtual Supermarket.* Journal of Nutrition Education and Behavior, 2020. **52**(9): p. 850-858.
43. Hersey, J.C., et al., *Effects of front-of-package and shelf nutrition labeling systems on consumers.* Nutrition Reviews, 2013. **71**(1): p. 1-14.
44. Ikonen, I., et al., *Consumer effects of front-of-package nutrition labeling: an interdisciplinary meta-analysis.* Journal of the Academy of Marketing Science, 2019.
45. Lima, M., G. Ares, and R. Deliza, *How do front of pack nutrition labels affect healthfulness perception of foods targeted at children? Insights from Brazilian children and parents.* Food Quality and Preference, 2018. **64**: p. 111-119.

46. Machín, L., et al., *The habitual nature of food purchases at the supermarket: Implications for policy making.* Appetite, 2020. **155**: p. 104844.
47. Malam, S., M.M. Raats, and M. Dean. *Comprehension and use of UK nutrition signpost labelling schemes*. 2009.
48. van Herpen, E. and H.C.M. van Trijp, *Front-of-pack nutrition labels. Their effect on attention and choices when consumers have varying goals and time constraints.* Appetite, 2011. **57**(1): p. 148-160.
49. van Kleef, E. and H. Dagevos, *The growing role of front-of-pack nutrition profile labeling: a consumer perspective on key issues and controversies.* Crit Rev Food Sci Nutr, 2015. **55**(3): p. 291-303.
50. Wahlich, C., B. Gardner, and L. McGowan, *How, when and why do young women use nutrition information on food labels? A qualitative analysis.* Psychology & Health, 2013. **28**(2): p. 202-216.
51. Campos, S., J. Doxey, and D. Hammond, *Nutrition labels on pre-packaged foods: a systematic review.* Public Health Nutrition, 2011. **14**(8): p. 1496-1506.
52. Talati, Z., et al., *Consumers' responses to health claims in the context of other on-pack nutrition information: a systematic review.* Nutr Rev, 2017. **75**(4): p. 260-273.
53. Cowburn, G. and L. Stockley, *Consumer understanding and use of nutrition labelling: a systematic review.* Public Health Nutrition, 2005. **8**(1): p. 21-28.
54. Hawley, K.L., et al., *The science on front-of-package food labels.* Public Health Nutrition, 2012. **16**(3): p. 430-439.
55. Vyth, E.L., et al., *Front-of-pack nutrition label stimulates healthier product development: a quantitative analysis.* International Journal of Behavioral Nutrition and Physical Activity, 2010. **7**(1): p. 65.
56. Hung, Y., et al., *Motivation outweighs ability in explaining European consumers' use of health claims.* Food Quality and Preference, 2017. **58**: p. 34-44.
57. Drichoutis, A.C., P. Lazaridis, and R.M. Nayga, Jr., *Nutrition knowledge and consumer use of nutritional food labels.* European Review of Agricultural Economics, 2005. **32**(1): p. 93-118.
58. Cecchini, M. and L. Warin, *Impact of food labelling systems on food choices and eating behaviours: a systematic review and meta-analysis of randomized studies.* Obes Rev, 2016. **17**(3): p. 201-10.
59. Ikonen, I., et al., *Consumer effects of front-of-package nutrition labeling: an interdisciplinary meta-analysis.* Journal of the Academy of Marketing Science, 2020. **48**(3): p. 360-383.
60. Temple, N.J., *Front-of-package food labels: A narrative review.* Appetite, 2020. **144**: p. 104485.
61. Becker, M.W., et al., *Front of pack labels enhance attention to nutrition information in novel and commercial brands.* Food Policy, 2015. **56**: p. 76-86.
62. Feunekes, G.I.J., et al., *Front-of-pack nutrition labelling: Testing effectiveness of different nutrition labelling formats front-of-pack in four European countries.* Appetite, 2008. **50**(1): p. 57-70.
63. He, F.J., H.C. Brinsden, and G.A. MacGregor, *Salt reduction in the United Kingdom: a successful experiment in public health.* J Hum Hypertens, 2014. **28**(6): p. 345-52.
64. Bablani, L., et al., *The impact of voluntary front-of-pack nutrition labelling on packaged food reformulation: A difference-in-differences analysis of the Australasian Health Star Rating scheme.* PLoS Med, 2020. **17**(11): p. e1003427.
65. Sassi F, C.M., Lauer J, Chisholm D. , *Improving Lifestyles, Tackling Obesity: The Health and Economic Impact of Prevention Strategies. OECD Health Working Papers, No. 48,* . 2009, OECD Publishing: Paris.
66. Jones, A., et al., *Front-of-pack nutrition labelling to promote healthier diets: current practice and opportunities to strengthen regulation worldwide.* BMJ Global Health, 2019. **4**(6): p. e001882.

67. Ministerio de Salud, *INFORME DE EVALUACIÓN DE LA IMPLEMENTACIÓN DE LA LEY SOBRE COMPOSICIÓN NUTRICIONAL DE LOS ALIMENTOS Y SU PUBLICIDAD [Translated: Evaluation report of the implementation of the law on nutritional composition of food and its advertising].* 2017: p. 77.
68. Corvalán, C., et al., *Impacto de la ley chilena de etiquetado en el sector productivo alimentario [Impact of the law labeling chilean in the productive sector food]*. 2021, FAO e INTA: Santiago de Chile. p. 87.
69. Egnell, M., et al., *Objective Understanding of Front-of-Package Nutrition Labels: An International Comparative Experimental Study across 12 Countries.* Nutrients, 2018. **10**(10).
70. Egnell, M., et al., *Objective understanding of the Nutri-score front-of-pack label by European consumers and its effect on food choices: an online experimental study.* International Journal of Behavioral Nutrition and Physical Activity, 2020. **17**(1): p. 146.
71. Vanderlee, L., et al., *The efficacy of 'high in' warning labels, health star and traffic light front-of-package labelling: an online randomised control trial.* Public Health Nutrition, 2020: p. 1-13.
72. Taillie, L.S., et al., *An evaluation of Chile's Law of Food Labeling and Advertising on sugar-sweetened beverage purchases from 2015 to 2017: A before-and-after study.* PLoS medicine, 2020. **17**(2): p. e1003015-e1003015.
73. Julia, C., et al., *Perception of different formats of front-of-pack nutrition labels according to sociodemographic, lifestyle and dietary factors in a French population: cross-sectional study among the NutriNet-Santé cohort participants.* BMJ Open, 2017. **7**(6): p. e016108.
74. WHO. *Policy brief: producing and promoting more food products consistent with a healthy diet*. 2014; Available from: https://www.who.int/nmh/ncd-coordination-mechanism/Policybrief32.pdf.
75. WHO, *Final report of the Commission on ending childhood obesity.* 2016: Geneva.
76. Department of Health. *Front of pack nutrition labelling: joint response to consultation*. 2013 February 5, 2021]; Available from: http://www.foodstandards.gov.scot/downloads/Front_of_pack_nutrition_labelling_joint_responses.pdf.
77. Stockley, R., E. Jordan, and A. Hunter, *Citizens' Forums on Food: Front of Pack (FoP) Nutrition Labelling*. 2009: London.
78. Food Standards Agency, *Concept Testing of Alternative Labelling of Healthy/Less Healthy Foods.* 2004.
79. European Commission, *Regulation (EU) No 1169/2011 of the European Parliament and of the Council on the provision of food information to consumers.* Official Journal of the European Union, 2011. **54**(L304): p. 46.
80. Care, D.o.H.a.S., *Childhood obesity: a plan for action, chapter 2*, D.o.H.a.S. Care, Editor. 2018.
81. Department of Health and Social Care, *Building on the success of front-of-pack nutrition labelling in the UK: a public consultation* 2020.
82. Food Standards Agency, *Consultation on Front of Pack Nutrition Labelling*. 2012.
83. UK Government. *Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets.* 2016 February 5, 2021]; Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/566251/FoP_Nutrition_labelling_UK_guidance.pdf.
84. Public Health England, *The Eatwell Guide - Booklet*. 2016.
85. NHS. *Food labels*. 2022 29 January 2023]; Available from: https://www.nhs.uk/Live-well/eat-well/food-guidelines-and-food-labels/how-to-read-food-labels/.
86. UK Government, *Press release: Final design of consistent nutritional labelling system given green light*. 2013.

87. Moore, S. and A. Hall, *P24 Penetration and presentation of front-of-pack nutrition labelling in UK supermarket websites: preliminary survey results.* Journal of Epidemiology and Community Health, 2021. **75**(Suppl 1): p. A53.
88. Mintel. *Brits Spent £12.3 Billion On Online Groceries In 2018.* 2019 29 January 2023]; Available from: https://www.mintel.com/press-centre/retail-press-centre/brits-spent-12-3-billion-on-online-groceries-in-2018
89. Cummins, S., et al., *COVID-19: impact on the urban food retail system and dietary inequalities in the UK.* Cities & Health, 2021. **5**(sup1): p. S119-S122.
90. Stones, C., *Online food nutrition labelling in the UK: how consistent are supermarkets in their presentation of nutrition labels online?* Public Health Nutrition, 2016. **19**(12): p. 2175-2184.
91. Food Standards Agency, *Eating Well Choosing Better: Wave 7*. 2022. p. 55.
92. Draper, A.K., et al., *Front-of-pack nutrition labelling: are multiple formats a problem for consumers?* European Journal of Public Health, 2013. **23**(3): p. 517-521.
93. Ares, G., et al., *Immediate effects of the implementation of nutritional warnings in Uruguay: awareness, self-reported use and increased understanding.* Public Health Nutrition, 2021. **24**(2): p. 364-375.
94. Andreeva, V.A., et al., *Bulgarian consumers' objective understanding of front-of-package nutrition labels: a comparative, randomized study.* Arch Public Health, 2020. **78**: p. 35.
95. Fialon, M., et al., *Effectiveness of Different Front-of-Pack Nutrition Labels among Italian Consumers: Results from an Online Randomized Controlled Trial.* Nutrients, 2020. **12**(8).
96. Vandevijvere, S., et al., *Consumers' food choices, understanding and perceptions in response to different front-of-pack nutrition labelling systems in Belgium: results from an online experimental study.* Arch Public Health, 2020. **78**: p. 30.
97. Raats, M.M., et al., *Reference amounts utilised in front of package nutrition labelling; impact on product healthfulness evaluations.* European Journal of Clinical Nutrition, 2015. **69**(5): p. 619-625.
98. Ducrot, P., et al., *Effectiveness of Front-Of-Pack Nutrition Labels in French Adults: Results from the NutriNet-Santé Cohort Study.* PLOS ONE, 2015. **10**(10): p. e0140898.
99. Maubach, N., J. Hoek, and D. Mather, *Interpretive front-of-pack nutrition labels. Comparing competing recommendations.* Appetite, 2014. **82**: p. 67-77.
100. Arrúa, A., et al., *Warnings as a directive front-of-pack nutrition labelling scheme: comparison with the Guideline Daily Amount and traffic-light systems.* Public Health Nutr, 2017. **20**(13): p. 2308-2317.
101. Vargas-Meza, J., et al., *Acceptability and understanding of front-of-pack nutritional labels: an experimental study in Mexican consumers.* BMC Public Health, 2019. **19**(1): p. 1751.
102. Khandpur, N., et al., *Are Front-of-Package Warning Labels More Effective at Communicating Nutrition Information than Traffic-Light Labels? A Randomized Controlled Experiment in a Brazilian Sample.* Nutrients, 2018. **10**(6).
103. Liu, D. and M. Juanchich, *Conceptual understanding and quantity inferences: a new framework for examining consumer understanding of food energy.* Public Health Nutr, 2018. **21**(17): p. 3168-3177.
104. Feunekes, G.I., et al., *Front-of-pack nutrition labelling: testing effectiveness of different nutrition labelling formats front-of-pack in four European countries.* Appetite, 2008. **50**(1): p. 57-70.
105. Grunert, K.G., J.M. Wills, and L. Fernández-Celemín, *Nutrition knowledge, and use and understanding of nutrition information on food labels among consumers in the UK.* Appetite, 2010. **55**(2): p. 177-89.

106. Hodgkins, C.E., et al., *Guiding healthier food choice: systematic comparison of four front-of-pack labelling systems and their effect on judgements of product healthiness.* Br J Nutr, 2015. **113**(10): p. 1652-63.
107. Roseman, M.G., H.-W. Joung, and E.I. Littlejohn, *Attitude and Behavior Factors Associated with Front-of-Package Label Use with Label Users Making Accurate Product Nutrition Assessments.* Journal of the Academy of Nutrition and Dietetics, 2018. **118**(5): p. 904-912.
108. Aschemann-Witzel, J., et al., *Effects of nutrition label format and product assortment on the healthfulness of food choice.* Appetite, 2013. **71**: p. 63-74.
109. Goodman, S., et al., *The Impact of Front-of-Package Label Design on Consumer Understanding of Nutrient Amounts.* Nutrients, 2018. **10**(11).
110. Forouzanfar, M.H., et al., *Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013.* The Lancet, 2015. **386**(10010): p. 2287-2323.
111. Shangguan, S., et al., *A Meta-Analysis of Food Labeling Effects on Consumer Diet Behaviors and Industry Practices.* American Journal of Preventive Medicine, 2019. **56**(2): p. 300-314.
112. Thomas, J., J. Brunton, and S. Graziosi, *EPPI-Reviewer 4.0: software for research synthesis.*, in *EPPI-Centre Software*. 2010, London: Social Science Research Unit, Institute of Education, University of London.
113. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS Med, 2009. **6**(7): p. e1000097.
114. Higgins, J.P.T., et al., *A revised tool for assessing risk of bias in randomized trials In: Chandler J, McKenzie J, Boutron I, Welch V (editors).* Cochrane Database of Systematic Reviews, 2016(Issue 10 (Suppl 1)).
115. Sterne, J.A.C., et al., *ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions.* BMJ, 2016. **355**: p. i4919.
116. Moodie, C., K. Angus, and M. Stead, *A systematic review of research exploring the response of consumers, retailers and tobacco companies to standardised packaging in the United Kingdom*, P.H.R. Consortium., Editor. 2019: Stirling, Scotland: Institute for Social Marketing, University of Stirling. .
117. O'Mara-Eves, A. and D. Kneale, *A critical appraisal tool for assessing research on purchasing (point of sales) data. Unpublished manuscript*, S.S.R.U. London: EPPI-Centre, UCL Institute of Education, Editor. 2019.
118. Health, N.R.C.U.C.o.D.a., *Diet and Health: Implications for Reducing Chronic Disease Risk.* 1989, National Academies Press (US): Washington (DC).
119. Scientific Advisory Committee on Nutrition, *Salt and Health.* 2003.
120. Higgins, J.P.T. and S. Green. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].* The Cochrane Collaboration, 2011. 2011; Available from: http://handbook.cochrane.org.
121. Egger, M., et al., *Bias in meta-analysis detected by a simple, graphical test.* Bmj, 1997. **315**(7109): p. 629-34.
122. Acton, R.B. and D. Hammond, *The impact of price and nutrition labelling on sugary drink purchases: Results from an experimental marketplace study.* Appetite, 2018. **121**: p. 129-137.
123. Acton, R.B., et al., *Taxes and front-of-package labels improve the healthiness of beverage and snack purchases: a randomized experimental marketplace.* International Journal of Behavioral Nutrition and Physical Activity, 2019. **16**(1): p. 46.
124. Neal, B., et al., *Effects of different types of front-of-pack labelling information on the healthiness of food purchases-a randomised controlled trial.* Nutrients, 2017. **9 (12) (no pagination)**(1284).

125. Tangari, A.H., et al., *That's Not So Bad, I'll Eat More! Backfire Effects of Calories-per-Serving Information on Snack Consumption.* Journal of Marketing, 2019. **83**(1): p. 133-150.
126. Graham, D.J., et al., *Impact of explained v. unexplained front-of-package nutrition labels on parent and child food choices: a randomized trial.* Public health nutrition, 2017. **20**(5): p. 774-785.
127. Ang, F.J.L., S. Agrawal, and E.A. Finkelstein, *Pilot randomized controlled trial testing the influence of front-of-pack sugar warning labels on food demand.* BMC Public Health, 2019. **19**(1): p. 164.
128. Crosetto, P., et al., *Modifications of food purchases in response to five nutrition simplified labelling.* Cahiers de Nutrition et de Dietetique, 2017. **52**(3): p. 129-133.
129. Machin, L., et al., *Can nutritional information modify purchase of ultra-processed products? Results from a simulated online shopping experiment.* Public Health Nutrition, 2018. **21**(1): p. 49-57.
130. Ni Mhurchu, C., et al., *Effects of interpretive nutrition labels on consumer food purchases: the Starlight randomized controlled trial.* Am J Clin Nutr, 2017. **105**: p. 695-704.
131. Zandstra, E.H., A.A. Willems, and R. Lion, *Making salt-reduced products more appealing to consumers: impact of front-of-pack messages on liking and table salt use over time.* Public Health Nutrition, 2018. **21**(15): p. 2762-2772.
132. Araya, S., et al., *Identifying Food Labeling Effects on Consumer Behavior*. 2018, Federal Reserve Bank of St Louis: St. Louis.
133. Elshiewy, O. and Y. Boztug, *When Back of Pack Meets Front of Pack: How Salient and Simplified Nutrition Labels Affect Food Sales in Supermarkets.* Journal of Public Policy & Marketing, 2018. **37**(1): p. 55.
134. Peñaherrera, V., et al., *Efecto del etiquetado de semáforo en el contenido nutricional y el consumo de bebidas gaseosas en Ecuador.* Revista Panamericana de Salud Publica, 2018. **42**: p. 1-8.
135. Kliemann, N., et al., *Reliability and validity of a revised version of the General Nutrition Knowledge Questionnaire.* Eur J Clin Nutr, 2016. **70**(10): p. 1174-1180.
136. Barbosa, L.B., et al., *Estudos de avaliação do conhecimento nutricional de adultos: uma revisão sistemática.* Ciência & Saúde Coletiva, 2016. **21**: p. 449-462.
137. Andrea, S.B., et al., *Does the association between early life growth and later obesity differ by race/ethnicity or socioeconomic status? A systematic review.* Ann Epidemiol, 2017. **27**(9): p. 583-592.e5.
138. Cameron, A.J., et al., *A Review of the Relationship Between Socioeconomic Position and the Early-Life Predictors of Obesity.* Curr Obes Rep, 2015. **4**(3): p. 350-62.
139. Fichera, E. and S. von Hinke, *The response to nutritional labels: Evidence from a quasi-experiment.* 2017.
140. Smed, S., A.K. Edenbrandt, and L. Jansen, *The effects of voluntary front-of-pack nutrition labels on volume shares of products: the case of the Dutch Choices.* Public Health Nutrition, 2019: p. 1-12.
141. Taillie, L.S., et al., *An evaluation of Chile’s Law of Food Labeling and Advertising on sugar-sweetened beverage purchases from 2015 to 2017: A before-and-after study.* PLOS Medicine, 2020. **17**(2): p. e1003015.
142. Vepsäläinen, H., et al., *Do we eat what we buy? Relative validity of grocery purchase data as an indicator of food consumption in the LoCard study.* British Journal of Nutrition, 2022. **128**(9): p. 1780-1788.
143. Jessop, C., *The NatCen Panel: developing an open probability-based mixed-mode panel in Great Britain.* Social Research Practice 2018(6): p. 2-14.

144. Office for National Statistics, National Records of Scotland, and Northern Ireland Statistics and Research Agency. *2011 Census aggregate data*. 2016 February 5, 2021]; June 2016:[

145. Santé Publique France. *Nutri-Score: Charte Graphique*. 2020 February 5, 2021]; Available from: https://partage.santepubliquefrance.fr/public/folder/MMGg-LaV6UiJ0BAm1Bb5PA/Pack_logo_hzt.

146. Republica de Chile Ministerio de Salud. *Reglamento sanitario de los alimentos*. 2019 February 5, 2021]; Available from: https://dipol.minsal.cl/wp-content/uploads/2019/06/DECRETO_977_96_2019-2.pdf.

147. Department of Health. *Nutrient Profiling Technical Guidance*. 2011 February 5, 2021]; Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216094/dh_123492.pdf.

148. Open Science Framework. *Experimental Study Examining the Impact of Front of Pack Labelling Skills on Knowledge.* 2020; Available from: https://osf.io/k9v2p/.

149. StataCorp., *Stata Statistical Software: Release 16*. 2019, StataCorp LLC.: College Station, TX.

150. Croker, H., et al., *Front of pack nutritional labelling schemes: a systematic review and meta-analysis of recent evidence relating to objectively measured consumption and purchasing.* Journal of Human Nutrition and Dietetics, 2020. **33**(4): p. 518-537.

151. Slavin, R. and D. Smith, *The Relationship Between Sample Sizes and Effect Sizes in Systematic Reviews in Education.* Educational Evaluation and Policy Analysis, 2009. **31**(4): p. 500-506.

152. Ducrot, P., et al., *Objective Understanding of Front-of-Package Nutrition Labels among Nutritionally At-Risk Individuals.* Nutrients, 2015. **7**(8): p. 7106-7125.

153. Packer, J., et al., *Assessing the Effectiveness of Front of Pack Labels: Findings from an Online Randomised-Controlled Experiment in a Representative British Sample.* Nutrients, 2021. **13**(3).

154. Arrúa, A., et al., *Warnings as a directive front-of-pack nutrition labelling scheme: comparison with the Guideline Daily Amount and traffic-light systems.* Public Health Nutrition, 2017. **20**(13): p. 2308-2317.

155. Dubois, P., et al., *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, 2021. **49**(1): p. 119-138.

156. Talati, Z., et al., *Consumers' Perceptions of Five Front-of-Package Nutrition Labels: An Experimental Study Across 12 Countries.* Nutrients, 2019. **11**(8).

157. Egnell, M., et al., *Consumers' Responses to Front-of-Pack Nutrition Labelling: Results from a Sample from The Netherlands.* Nutrients, 2019. **11**(8): p. 1817.

158. Vandevijvere, S., et al., *Consumers' food choices, understanding and perceptions in response to different front-of-pack nutrition labelling systems in Belgium: results from an online experimental study.* Archives of Public Health, 2020. **78**(1): p. 30.

159. European scientists and health professionals for Nutri-Score, *WHY THE EUROPEAN COMMISSION MUST CHOOSE THE NUTRI-SCORE NUTRITION LABEL* 2023.

160. WHO, *Guiding principles and framework manual for front-of-pack labelling for promoting healthy diet*. 2019. p. 46.

161. Collins, L.M. and S.T. Lanza, *Latent class and latent transition analysis.* . 2010: Wiley.

162. Weller, B.E., N.K. Bowen, and S.J. Faubert, *Latent Class Analysis: A Guide to Best Practice.* Journal of Black Psychology, 2020. **46**(4): p. 287-311.

163. Lanza, S.T. and B.L. Rhoades, *Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment.* Prev Sci, 2013. **14**(2): p. 157-68.

164. Clark, S. and B. Muthén, *Relating Latent Class Analysis Results to Variables not Included in the Analysis.* 2009.

165. James, R.J.E., C. O'Malley, and R.J. Tunney, *Sociodemographic predictors of latent class membership of problematic and disordered gamblers.* Addictive Behaviors Reports, 2016. **3**: p. 61-69.
166. MacDonald, K., *Latent Class Analysis (LCA) in Stata*. 2018: London Stata Conference.
167. Nagin, D.S., *Group-based modeling of development*. 2005, Cambridge, MA: Harvard University Press.
168. Nylund, K.L., T. Asparouhov, and B.O. Muthén, *Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study.* Structural Equation Modeling: A Multidisciplinary Journal, 2007. **14**(4): p. 535-569.
169. Lanza, S.T., X. Tan, and B.C. Bray, *Latent Class Analysis With Distal Outcomes: A Flexible Model-Based Approach.* Structural Equation Modeling: A Multidisciplinary Journal, 2013. **20**(1): p. 1-26.
170. Larsen, F.B., et al., *A Latent Class Analysis of Multimorbidity and the Relationship to Socio-Demographic Factors and Health-Related Quality of Life. A National Population-Based Study of 162,283 Danish Adults.* PLOS ONE, 2017. **12**(1): p. e0169426.
171. Vermunt, J.K. and J. Magidson, *Latent GOLD 5.0 Upgrade Manual*. 2013, Belmont, MA: Statistical Innovations Inc.
172. Saville, D.J., *Multiple Comparison Procedures: The Practical Solution.* The American Statistician, 1990. **44**(2): p. 174-180.
173. Vyth, E.L., et al., *Actual use of a front-of-pack nutrition logo in the supermarket: consumers' motives in food choice.* Public Health Nutrition, 2010. **13**(11): p. 1882-1889.
174. Hung, Y. and W. Verbeke, *Consumer evaluation, use and health relevance of health claims in the European Union.* Food Quality and Preference, 2019. **74**: p. 88-99.
175. Pettigrew, S., et al., *Front-of-pack nutrition labels: an equitable public health intervention.* European Journal of Clinical Nutrition, 2023. **77**(1): p. 135-137.
176. Julia, C., et al., *Impact of the front-of-pack 5-colour nutrition label (5-CNL) on the nutritional quality of purchases: an experimental study.* International Journal of Behavioral Nutrition and Physical Activity, 2016. **13**(1): p. 101.
177. Brett, J., et al., *Mapping the impact of patient and public involvement on health and social care research: a systematic review.* Health Expectations, 2014. **17**(5): p. 637-650.
178. NIHR, *Briefing notes for researchers - public involvement in NHS, health and social care research*. 2021.
179. Health Research Authority and INVOLVE. *Impact of public involvement on the ethical aspects of research*. 2016 29/04/23]; Available from: https://s3.eu-west-2.amazonaws.com/www.hra.nhs.uk/media/documents/impact-public-involvement-ethical-aspects-research-updated-2016.pdf.
180. Thompson, J., et al., *Exploring the Impact of Patient and Public Involvement in a Cancer Research Setting.* Qualitative Health Research, 2013. **24**(1): p. 46-54.
181. Abrehart, N., et al., *"A little (PPI) MAGIC can take you a long way" : involving children and young people in research from inception of a novel medical device to multi-centre clinical trial Roald Dahl, James and the Giant Peach (1961).* Research Involvement and Engagement, 2021. **7**(1): p. 2.
182. NIHR. *Involving children and young people as advisors in research*. 2021 29/04/23]; Available from: https://www.learningforinvolvement.org.uk/content/resource/nihr-involving-children-and-young-people-as-advisors-in-research/.
183. NIHR, et al., *UK Standards for Public Involvement - Better public involvement for better health and social care research*. 2021.
184. Telford, R., J.D. Boote, and C.L. Cooper, *What does it mean to involve consumers successfully in NHS research? A consensus study.* Health Expectations, 2004. **7**(3): p. 209-220.

185. National Children's Bureau. *Involving children & young people in research*. 2022 01/01/2022]; Available from: https://www.ncb.org.uk/what-we-do/involving-children-young-people/involving-children-young-people-research.
186. AB, M. *Mentimeter*. 2021 01/06/2021]; Version 3.2.5, 2021-12-15 [Available from: https://www.mentimeter.com.
187. World Cancer Research Fund International, *Building momentum: lessons on implementing a robust front-of-pack food label*. 2019.
188. Action on Salt. *FoodSwitch*. 2022 01/03/23]; Available from: https://www.actiononsalt.org.uk/foodswitch/
189. Song, J., et al., *Impact of color-coded and warning nutrition labelling schemes: A systematic review and network meta-analysis.* PLOS Medicine, 2021. **18**(10): p. e1003765.
190. WHO Regional Office for Europe. *European Food and Nutrition Action Plan 2015–2020*. Regional Committee for Europe 64th session 2014 February 5, 2021].
191. World Health Organization, *State of play of WHO guidance on Front-of-the-Pack labelling*. 2021.
192. Mantilla Herrera, A.M., et al. *Cost-Effectiveness of Product Reformulation in Response to the Health Star Rating Food Labelling System in Australia*. Nutrients, 2018. **10**, DOI: 10.3390/nu10050614.
193. UK Government, *Government delays restrictions on multibuy deals and advertising on TV and online*. 2022.
194. Russell, S.J., J. Packer, and R.M. Viner, *Obesity Policy Research Unit – Patient and Public Involvement Portfolio 2017-2022*. 2022. p. 21.
195. Volpp, K.G. and D.A. Asch, *Make the healthy choice the easy choice: using behavioral economics to advance a culture of health.* QJM: An International Journal of Medicine, 2017. **110**(5): p. 271-275.
196. Frieden, T.R., *A framework for public health action: the health impact pyramid.* Am J Public Health, 2010. **100**(4): p. 590-5.
197. European Commission, *REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL regarding the use of additional forms of expression and presentation of the nutrition declaration*. 2020: Brussels. p. 25.
198. Méjean, C., et al., *Perception of front-of-pack labels according to social characteristics, nutritional knowledge and food purchasing habits.* Public Health Nutr, 2013. **16**(3): p. 392-402.
199. Institute for Grocery Distribution. *Front Of Pack Labelling: What Consumers Want.* . 2019; Available from: https://www.igd.com/articles/article-viewer/t/front-of-pack-labelling-what-consumers-want/i/22181.
200. Ogundijo, D.A., A.A. Tas, and B.A. Onarinde, *An assessment of nutrition information on front of pack labels and healthiness of foods in the United Kingdom retail market.* BMC Public Health, 2021. **21**(1): p. 220.
201. Pettigrew, S., et al., *An examination of public support for 35 nutrition interventions across seven countries.* Eur J Clin Nutr, 2023. **77**(2): p. 235-245.
202. Food and Drink Federation Scotland, *Reformulation for Health: Guide to Reformulation*. 2021. p. 20.
203. New Food Magazine. *Protecting consumers with front of pack labelling*. 2020 28/02/23]; Available from: https://www.newfoodmagazine.com/article/106906/protecting-consumers-with-front-of-pack-labelling/.
204. Gressier, M., F. Sassi, and G. Frost, *Healthy Foods and Healthy Diets. How Government Policies Can Steer Food Reformulation.* Nutrients, 2020. **12**(7).

205. Corvalán, C., et al., *Structural responses to the obesity and non-communicable diseases epidemic: Update on the Chilean law of food labelling and advertising.* Obesity Reviews, 2019. **20**(3): p. 367-374.
206. Food Navigator Europe. *5-c creator Serge Hercberg on nutrition logos, lobbies and conflicts of interest*. 2016 04/03/23]; Available from: https://www.foodnavigator.com/Article/2016/11/16/5-c-creator-Serge-Hercberg-on-nutrition-logos-lobbies-and-conflicts-of-interest
207. Global Health Advocacy Incubator (GHAI) - Campaign for Tobacco-Free Kids (CTFK), *Behind the labels: Big Food's War on Healthy Food Policies*. 2021. p. 30.
208. Hawkes, C., et al., *Smart food policies for obesity prevention.* The Lancet, 2015. **385**(9985): p. 2410-2421.
209. Hawkes, C., J. Jewell, and K. Allen, *A food policy package for healthy diets and the prevention of obesity and diet-related non-communicable diseases: the NOURISHING framework.* Obes Rev, 2013. **14 Suppl 2**: p. 159-68.
210. Tomiyama, A.J., *Weight stigma is stressful. A review of evidence for the Cyclic Obesity/Weight-Based Stigma model.* Appetite, 2014. **82**: p. 8-15.
211. Crovetto, M., M. Acosta, and Y. Rocco, *Ley 20.606: Efectos en el conocimiento de etiquetado nutricional en consumidores de un supermercado en Valparaíso de Chile: estudio descriptivo, cuanticualitativo, antes y después de 5 meses de la implementación de la ley.* Revista Española de Nutrición Humana y Dietética, 2020. **24**(4): p. 311-323.
212. Correa, T., et al. *Why Don't You [Government] Help Us Make Healthier Foods More Affordable Instead of Bombarding Us with Labels? Maternal Knowledge, Perceptions, and Practices after Full Implementation of the Chilean Food Labelling Law*. International Journal of Environmental Research and Public Health, 2022. **19**, DOI: 10.3390/ijerph19084547.
213. Yuka. *Food analysis*. 2023 01/03/23]; Available from: https://help.yuka.io/l/en/category/zlf9yu9htc-food-scoring.
214. Business, B. *INTERMARCHÉ CHANGES THE RECIPE OF 900 POORLY RATED PRODUCTS ON YUKA*. 2019 01/03/23]; Available from: https://www.bfmtv.com/economie/consommation/intermarche-change-la-recette-de-900-produits-mal-notes-sur-yuka_AV-201909160092.html.
215. Forwood, S.E., et al., *Offering within-category food swaps to reduce energy density of food purchases: a study using an experimental online supermarket.* International Journal of Behavioral Nutrition and Physical Activity, 2015. **12**(1): p. 85.
216. Marty, L., et al. *Effects of Labelling and Increasing the Proportion of Lower-Energy Density Products on Online Food Shopping: A Randomised Control Trial in High- and Low-Socioeconomic Position Participants*. Nutrients, 2020. **12**, DOI: 10.3390/nu12123618.
217. Institute for Grocery Distribution. *Motivators and barriers to use*. 2017 6/03/23]; Available from: https://www.igd.com/Social-impact/Health/Engaging-consumers/Labelling/Our-research

https://www.igd.com/Portals/0/Downloads/Motivators_and_barriers.pptx.
218. Department of Health and Social Care. *Guidance: Nutrition and health claims: guidance to compliance with Regulation (EC) 1924/2006*. 2021; Available from: https://www.gov.uk/government/publications/nutrition-and-health-claims-guidance-to-compliance-with-regulation-ec-1924-2006-on-nutrition-and-health-claims-made-on-foods/nutrition-and-health-claims-guidance-to-compliance-with-regulation-ec-19242006.
219. Department of Health and Social Care. *Guidance: Fortified foods: guidance to compliance on European Regulation (EC) No. 1925/2006 on the addition of vitamins and minerals and certain other substances to food*. 2021 01/03/23]; Available from:

https://www.gov.uk/government/publications/fortified-foods-guidance-to-compliance-with-european-regulation-ec-no-1925-2006-on-the-addition-of-vitamins-and-minerals-and-certain-other-substances-to-food/fortified-foods-guidance-to-compliance-on-european-regulation-ec-no-19252006-on-the-addition-of-vitamins-and-minerals-and-certain-other-substance#section-4.

220. UK Government. *Regulation (EC) No 1924/2006 of the european parliament and of the council*. 2020 05/03/23]; Available from: https://www.legislation.gov.uk/eur/2006/1924/annex
221. UK Government. *Guidance: Great Britain nutrition and health claims (NHC) register*. 2022; Available from: https://www.gov.uk/government/publications/great-britain-nutrition-and-health-claims-nhc-register.
222. Food and Drink Federation, *FDF Guidance on 'Allergen'-Free & Vegan Claims*. 2023. p. 2.
223. UK Government, *Guidance: Government 5 A Day logo*, Public Health England, Editor. 2016. p. 19.
224. British Heart Foundation. *Can I trust 5-a-day food labels?* 2015; Available from: https://www.bhf.org.uk/informationsupport/heart-matters-magazine/nutrition/ask-the-expert/five-a-day-labels.
225. García, A.L., et al., *Confused health and nutrition claims in food marketing to children could adversely affect food choice and increase risk of obesity.* Archives of Disease in Childhood, 2019. **104**(6): p. 541.
226. Colour Blind Awareness. *About Colour Blindness*. 2016 01/01/23]; Available from: https://www.colourblindawareness.org/colour-blindness/.
227. Diabetes UK. *People with diabetes say all companies should use the same food labelling*. 2018 2/03/23]; Available from: https://www.diabetes.org.uk/about_us/news/consistent-food-labelling-for-all.
228. Institute for Grocery Distribution. *Three things to know about front of pack nutrition labels in the UK*. 2021; Available from: https://www.igd.com/articles/article-viewer/t/three-things-to-know-about-front-of-pack-nutrition-labels-in-the-uk/i/27499.
229. McClinchy, J., A. Dickinson, and W. Wills, *'Feta is obviously very dangerous stuff looking at all those reds': Can nutrition labelling help people living with type 2 diabetes to self-manage their condition?* J Hum Nutr Diet, 2023.
230. Scarborough, P., et al., *Reds are more important than greens: how UK supermarket shoppers use the different information on a traffic light nutrition label in a choice experiment.* Int J Behav Nutr Phys Act, 2015. **12**: p. 151.
231. Public Health England, *Sugar reduction and wider reformulation programme: Report on progress towards the first 5% reduction and next steps*. 2018.
232. Savio, S., et al., *A survey of the reformulation of Australian child-oriented food products.* BMC Public Health, 2013. **13**(1): p. 836.
233. TESCO. *Tropicana Orange Juice Smooth 300Ml*. 2023 02/02/23]; Available from: https://www.tesco.com/groceries/en-GB/products/287438170.
234. Aguenaou, H., et al. *Comment on Muzzioli et al. Are Front-of-Pack Labels a Health Policy Tool? Nutrients 2022, 14, 771*. Nutrients, 2022. **14**, DOI: 10.3390/nu14102165.
235. International Agency for Research on Cancer, *The Nutri-Score: A Science-Based Front-of-Pack Nutrition Label. Helping Consumers Make Healthier Food Choices*. 2021: Lyon, France.
236. Dillman Carpentier, F.R., et al., *Evaluating the impact of Chile's marketing regulation of unhealthy foods and beverages: pre-school and adolescent children's changes in exposure to food advertising on television.* Public Health Nutrition, 2020. **23**(4): p. 747-755.
237. Ministerio de Salud, *INFORME SOBRE EVALUACIONES DE LA LEY Nº 20.606 SOBRE COMPOSICIÓN NUTRICIONAL DE LOS ALIMENTOS Y SU PUBLICIDAD [Translated: Report on evaluations of law nº 20,606 on nutritional composition of foods and its advertising]*. 2021. p. 20.
238. Barahona, N., C. Otero, and S. Otero, *Equilibrium Effects of Food Labeling Policies.* SSRN, 2020.

239. Quintiliano Scarpelli, D., et al. *Changes in Nutrient Declaration after the Food Labeling and Advertising Law in Chile: A Longitudinal Approach*. Nutrients, 2020. **12**, DOI: 10.3390/nu12082371.
240. Reyes, M., et al., *Changes in the amount of nutrient of packaged foods and beverages after the initial implementation of the Chilean Law of Food Labelling and Advertising: A nonexperimental prospective study.* PLoS Med, 2020. **17**(7): p. e1003220.
241. Taillie, L.S., et al., *Changes in food purchases after the Chilean policies on food labelling, marketing, and sales in schools: a before and after study.* The Lancet Planetary Health, 2021. **5**(8): p. e526-e533.
242. Schnettler, B., et al., *Are consumers willing to pay more for reformulated processed meat products in the context of the implementation of nutritional warnings? Case study with frankfurters in Chile.* Meat Sci, 2019. **152**: p. 104-108.
243. Correa, T., et al., *"Responses to the Chilean law of food labeling and advertising: exploring knowledge, perceptions and behaviors of mothers of young children".* Int J Behav Nutr Phys Act, 2019. **16**(1): p. 21.
244. Ares, G., et al., *Qualitative exploration of the reasons for not using nutritional warnings after policy implementation in Uruguay.* Health Promotion International, 2023. **38**(1): p. daac174.
245. Health], H.N.D.o., *Anbefaling av en sunnhetsmerkeordning [Recommendation for a health mark scheme]*. 2008: Oslo. p. 68.
246. Bonaccio, M., et al., *Ultra-processed food consumption modifies the association of Nutri-Score with all-cause mortality.* European Journal of Public Health, 2021. **31**(Supplement_3): p. ckab164.448.
247. Mertens, E., C. Colizzi, and J.L. Peñalvo, *Ultra-processed food consumption in adults across Europe.* European Journal of Nutrition, 2022. **61**(3): p. 1521-1539.
248. Rico-Campà, A., et al., *Association between consumption of ultra-processed foods and all cause mortality: SUN prospective cohort study.* BMJ, 2019. **365**: p. l1949.
249. Srour, B., et al., *Ultra-processed food intake and risk of cardiovascular disease: prospective cohort study (NutriNet-Santé).* BMJ, 2019. **365**: p. l1451.
250. Popkin, B.M., et al., *Towards unified and impactful policies to reduce ultra-processed food consumption and promote healthier eating.* Lancet Diabetes Endocrinol, 2021. **9**(7): p. 462-470.
251. Bonaccio, M., et al., *Joint association of food nutritional profile by Nutri-Score front-of-pack label and ultra-processed food intake with mortality: Moli-sani prospective cohort study.* BMJ, 2022. **378**: p. e070688.
252. Cotter, T., et al., *'Warning: ultra-processed' — A call for warnings on foods that aren't really foods.* BMJ Global Health, 2021. **6**(12): p. e007240.
253. Zancheta Ricardo, C., et al., *Changes in the Use of Non-nutritive Sweeteners in the Chilean Food and Beverage Supply After the Implementation of the Food Labeling and Advertising Law.* Front Nutr, 2021. **8**: p. 773450.
254. Mialon, M., *Front of pack nutrition labeling: ultra-processed foods are the elephant in the room.* The Lancet Regional Health – Americas, 2023. **18**.
255. Monteiro, C.A. and G. Cannon, *The trouble with ultra-processed foods.* BMJ, 2022. **378**: p. o1972.
256. Adams, J., et al., *Public health response to ultra-processed food and drinks.* BMJ, 2020. **369**: p. m2391.
257. BMJ. *New evidence links ultra-processed foods with a range of health risks*. 2019 01/03/23]; Available from: https://www.bmj.com/company/newsroom/new-evidence-links-ultra-processed-foods-with-a-range-of-health-risks/.

258. Darmon, N. and A. Drewnowski, *Contribution of food prices and diet cost to socioeconomic disparities in diet quality and health: a systematic review and analysis.* Nutrition Reviews, 2015. **73**(10): p. 643-660.
259. Egnell, M., et al., *Modelling the impact of different front-of-package nutrition labels on mortality from non-communicable chronic disease.* International Journal of Behavioral Nutrition and Physical Activity, 2019. **16**(1): p. 56.
260. UK Government, *Restricting promotions of products high in fat, sugar and salt by location and by price: government response to public consultation*. 2021.
261. Chandon, P. and B. Wansink, *When Are Stockpiled Products Consumed Faster? A Convenience–Salience Framework of Postpurchase Consumption Incidence and Quantity.* Journal of Marketing Research, 2002. **39**(3): p. 321-335.
262. Public Health England, *Sugar Reduction: The evidence for action.* 2015.
263. Pettigrew, S., et al., *An 18-country analysis of the effectiveness of five front-of-pack nutrition labels.* Food Quality and Preference, 2023. **104**: p. 104691.
264. Stephane, B., B. David, and B. Malek, *A study is 21 times more likely to find unfavourable results about the nutrition label Nutri-Score if the authors declare a conflict of interest or the study is funded by the food industry.* BMJ Global Health, 2023. **8**(5): p. e011720.
265. Scientific Committee of the Nutri-Score, *Update of the Nutri-Score algorithm for beverages: Second update report*. 2023. p. 104.
266. Hoge, A., et al. *Health Literacy and Its Associations with Understanding and Perception of Front-of-Package Nutrition Labels among Higher Education Students*. International Journal of Environmental Research and Public Health, 2022. **19**, DOI: 10.3390/ijerph19148751.
267. Manon, E., et al., *Randomised controlled trial in an experimental online supermarket testing the effects of front-of-pack nutrition labelling on food purchasing intentions in a low-income population.* BMJ Open, 2021. **11**(2): p. e041196.
268. Fuchs, K.L., et al. *Effects of Digital Food Labels on Healthy Food Choices in Online Grocery Shopping*. Nutrients, 2022. **14**, DOI: 10.3390/nu14102044.
269. UK Government. *English language skills*. 2018 01/01/23]; Available from: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/english-language-skills/latest.
270. Office for National Statistics. *People who cannot speak English well are more likely to be in poor health*. 2015 01/01/23]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/articles/peoplewhocannotspeakenglishwellaremorelikelytobeinpoorhealth/2015-07-09.
271. Drieskens, S., et al., *Correction of self-reported BMI based on objective measurements: a Belgian experience.* Archives of Public Health, 2018. **76**(1): p. 10.

Appendix A: Literature Review (Chapter 2)

Table A1: Search strategy

Search terms
((((Food packaging/[Title/Abstract] OR Product labeling/[Title/Abstract] OR "front of pack*"[Title/Abstract] OR "food label*"[Title/Abstract] OR (nutri* AND label)[Title/Abstract] OR "FOPL"[Title/Abstract])) AND ((knowledge[Title/Abstract] OR understanding[Title/Abstract] OR ranking*[Title/Abstract]))) AND ((RCT[Title/Abstract] OR "Random* controlled trial*"[Title/Abstract] OR experiment*[Title/Abstract] OR trial[Title/Abstract]))) NOT ((review[Title/Abstract]))

Appendix B: Systematic Review (Chapter 3)

Screening Criteria

Table B1: Screening criteria

Participants	Any age
Exposure (intervention)	Front of pack nutritional labelling schemes- this could be implemented in a real-life setting (community or national level) or an experimental setting. Applied to pre-packaged foods only. Age restrictions (e.g., caffeine) Purchasing for individual/ family, not food supply
Comparison	Randomised or quasi-randomised controlled trials (RCTs/Q-RCTs) Controlled before-and-after studies Interrupted time series (ITS) studies Compared a labelled product (with information on nutrients or energy) with the same product without a nutritional label Assessed objectively measured purchasing or consumption of foods or non-alcoholic drinks in real-world or laboratory settings. Studies where both groups were exposed to FOPL, and an intervention group was exposed to additional info, were excluded
Outcome measure	Purchasing behaviour (individual/ family) Dietary intake (individual) Sales data (higher level)
Study designs	Experimental; intervention (with a control group)
Other Geography Languages Time	 All All From 2017 start

Rationale for Meta-Analysis Inclusion and Data Processing

Table B2: Meta-analysis rationale

Author	Rationale / data processing
Acton 2018	**Comparison groups combined for the overall meta-analysis**: HSR and 'High in sugar' symbol. **Comparison groups included in the separate meta-analyses:** HSR; 'High in sugar' symbol. Notes: Text health warning intervention excluded due to lack of real-world FOPL scheme. The 'no tax' condition results were used, as per advice from authors when data was provided and a final decision made by our team. Sugar and energy outcomes were converted to sugar or energy per 100g/mL, by using the calculated mean of the beverage options, presented in Table 1 in their paper.
Acton 2019	**Comparison groups combined for the overall meta-analysis:** 'High in' symbol, MTL, HSR and nutrition grade (equivalent to Nutri-Score) were all combined for labelling vs no labelling meta-analysis. **Comparison groups included in the separate meta-analyses:** MTL; 'High in' symbol; HSR; nutrition grade (equivalent to Nutri-Score). Notes: The 'no tax' condition results were used, as per advice from authors when data was provided and a final decision made by our team. Sugar and energy outcomes were converted to grams or kilocalorie per 100g/mL, by using the calculated mean of the beverage and food options, presented in the additional file 1, Table S3 of their paper. The sodium and saturated fat results for the beverage purchase task were not included in the meta-analyses, as very few included beverages had either nutrient.
Machin 2017	**Comparison groups combined for the overall meta-analysis:** MTL and 'High in' symbol. **Comparison groups for the separate meta-analyses:** MTL; 'High in' symbol. Notes: Sodium outcomes were converted from presented salt outcomes using 1 mg sodium = 2.55 mg salt (SACN, 2003). Standard deviations were calculated from 95% confidence interval using Cochrane methods (Higgins & Green, 2011).
Neal 2017	**Comparison groups combined for the overall meta-analysis:** HSR, MTL and DIG. **Comparison groups included in the separate meta-analyses:** MTL; HSR; DIG. Notes: Recommended warnings intervention was excluded as it was a back of pack intervention. Energy outcomes were converted from kJ to kcal using 4.184 kJ= 1 kcal (Health NRCUCoDa, 1989)
Ni Mhurchu 2017	**Comparison groups combined for the overall meta-analysis:** Traffic light label and HSR. **Comparison groups for the separate meta-analyses:** Traffic light label; HSR. Notes: Energy outcomes were converted from kJ to kcal using 4.184 kJ= 1 kcal (27). The UK MTL scheme was referenced in the paper, but an example of the scheme was not explicitly provided.

MTL= Multiple Traffic Light; HSR = Health Star Rating; DIG = daily intake guide.

Bias Assessments

Figure B1: Bias assessment for experimental studies

	Randomization process	Deviations from intended interventions (adherence)	Deviations from intended interventions (assignment)	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall bias
Acton 2018	!	+	+	+	+	!	!
Acton 2019	!	+	+	+	+	!	!
Ang 2019	+	+	+	+	+	!	!
Crosetto 2017	!	+	+	+	+	!	!
Graham 2017	+	+	+	+	+	!	!
Machin 2017	+	+	+	+	+	!	!
Neal et al 2017	+	+	+	!	+	!	!
Ni Mhurchu 2017	+	+	+	+	+	+	+
Tangari S1 2019	!	!	+	+	+	!	–
Tangari S5 2019	!	!	+	+	!	!	–
Zandstra 2018	!	+	+	+	+	!	!

Risk of bias: green indicates low risk; yellow indicates some concerns; red indicates high risk.

Table B3: Quantitative sales data bias and quality assessment

	Araya et al, 2018	Elshiewy & Boztug, 2018	Peñaherrera, 2018
AIMS			
Was there a clear statement of the aims of the study?	Yes	Yes	Yes
DATA			
Was a commercial data set used?	Yes	Yes	Yes
What was the data collection method?	Loyalty card record	Scanner data	Kantar data
Was the dataset complete?	Unclear	Unclear	Unclear
Were there apparent, unintended restrictions in the dataset in terms of tobacco products or sales points?	No	No	No
Was the sampling method appropriate to the question/ inference being made?	Yes	Yes	Yes
Was the data sample representative of intended population?	Yes	Yes	Yes
Did the study report a priori power calculations (where appropriate)?	Unclear	Unclear	Unclear
Was the timing of the data collection appropriate for the aims of the study?	Yes	Yes	Yes
Duration of study: Were the data collected over a sufficient time period for the intended analysis?	Yes	Yes	Yes
Were all plausible variables of interest measured in the dataset?	Unclear	Unclear	Unclear
Are the measures/ variables adequately described?	Yes	Yes	Yes
Overall, are there concerns about the soundness of the data for the purposes of the study?	No	No	No
ANALYSIS			
Were the analyses appropriate given the stated aims?	Yes	Yes	Yes
Was the size of the dataset sufficient for the analyses being conducted?	Yes	Yes	Yes
Have the data been analysed appropriately?	Unclear	Unclear	Unclear
Overall, are there concerns about the analyses?	No	No	No
INFERENCES AND CONCLUSIONS			
Are the inferences drawn from analyses appropriate given the sample relative to population?	Yes	Yes	Yes
Are the inferences drawn appropriate given the analyses and results?	Yes	Yes	Yes
Is there an over-emphasis on statistical significance rather than magnitude/ direction of effect?	No/unclear	No/unclear	No/unclear
Is there an appropriate emphasis of the real-world importance of the statistical results?	Yes	Yes	Yes
OVERALL			
Are there concerns about the soundness of the study?	No	No	No

Full Search Strategies for Each Database

Table B4: Search history

ABI Inform Global (Proquest) 8/4/2019 165 records Search date from 2016 onwards Exclude: Newspapers, trade journals, wire feeds, magazines, blogs, podcasts and websites S1 ((TI,AB("food*" OR drink* OR beverage*) AND TI,AB(preference* OR habit*" OR behavio* OR choice* OR decision* OR decid* OR consump* OR "purchas*" OR "buy*" OR "sale*" OR "sell* OR "shop*" OR "store*" OR "supermarket*" OR "point of purchas*")) OR (MAINSUBJECT.EXACT("Diet") OR MAINSUBJECT.EXACT("Eating") OR MAINSUBJECT.EXACT("Eating behavior") OR MAINSUBJECT.EXACT("Product Choice"))) S2 MAINSUBJECT.EXACT("Food packaging") OR MAINSUBJECT.EXACT("Descriptive Labeling") S3 ((TI,AB("food law") AND TI,AB(label* OR pack* OR "pre-pack*" OR "pre pack*")) OR TI,AB(Food* N/1 label*) OR TI,AB("food content*" AND (label* OR sign OR signs OR signage OR symbol OR symbols)) OR TI,AB(("guideline daily amount*" OR "nutrient reference*" OR "reference intake*" OR "nutrient intake*" OR "daily value*") AND (label* OR "content* sign*" OR symbol* OR ticket* OR sticker* OR vignette*)) OR TI,AB(("keyhole" OR "key hole") AND (nordic* OR label* OR "sign" OR "signs" OR "signage" OR symbol* OR vignette*)) OR TI,AB(nyckelh* OR "ley del S*" OR "Ley Super 8") OR TI,AB(("nutrition information" OR "nutrient information" or "nutritional information") AND (pack* OR label* OR prepack* OR "pre pack*" OR symbol* OR vignette*))) S4 ((TI,AB(Nutritive OR Nutrient OR Nutrition OR Nutritional OR Nutrients OR calorific OR calorie OR calories OR caloric OR kilojoule OR kilojoules OR kilocalorie OR kilocalories OR kcal OR kcals OR kj OR "energy" OR fat OR fats or "recommended dietary allowance*" or salt OR Sugar OR sugars) AND TI,AB(label OR labels OR labeling OR labelling OR labeled OR labelled OR symbol or symbols OR warning OR warnings OR vignette OR vignettes)) OR TI,AB((warning AND octagon*) OR "octagonal black system*" OR wobs OR "ley de etiquetado de alimentos" OR "ley de alimentos" OR (law AND "20.606") OR "Super 8" OR "Nutri-score*" OR "Nutri score*" OR "Nutriscore*" OR "Health star*" OR (hsr AND system*) OR ("traffic light*" N/5 (label* OR food* OR nutri* OR diet* OR pack*))) OR TI,AB("Warning label*" N/5 (food* OR nutri* OR diet*)) OR TI,AB((pack OR packed OR package OR packages OR label OR labels OR labeling OR labelling OR labeled OR labelled OR prepack* OR "pre-pack*") AND (drink OR drinks OR beverage OR beverages))) S1 AND (S2 OR S3 OR S4)
ASSIA and Sociological Abstracts (Proquest) 8/4/2019 167 results ASSIA, 88 results (SocAbs) (TI,AB(("Nutriti*" OR "Nutrient*") AND (label* OR "content* sign*" OR symbol* OR ticket* OR sticker* OR warning* OR vignette*)) OR TI,AB(((calorific OR calorie* OR caloric OR kilojoule* OR kilocalorie* OR kcal* OR kj OR "energy") AND information) AND (pack* OR label* OR prepack* OR "pre-pack*" OR "pre pack*")) OR TI,AB(((product N/2 label*) AND food*) OR "front of pack*" OR "food label*" OR ((prepack* AND food*) AND label*) OR (("pre-pack*" AND food*) AND label*) OR (("pre pack*" AND food*) AND label*)) OR TI,AB((calorific OR calorie* OR caloric OR kilojoule* OR kilocalorie* OR kcal* OR kj OR "energy") AND (label* OR "content* sign*" OR symbol* OR ticket* OR sticker* OR vignette*)) OR TI,AB((drink* OR beverage*) AND (label OR labelling OR labelling OR labels)) OR TI,AB((fat OR fats) AND (label* OR "content* sign*" OR symbol* OR tag* OR ticket* OR sticker* OR vignette*)) OR TI,AB((label* AND (legislation* OR regulation* OR policies OR policy OR law

OR laws)) AND (food* OR diet* OR nutri*)) OR TI,AB((soda OR sodas OR "flavored water*" OR "flavoured water*" OR "fruit water*" OR cordial OR cordials OR squash OR squashes OR juice OR juices OR smoothie* OR milkshake* OR tea OR teas OR coffee*) AND (label OR labelling OR labelling OR labels)) OR TI,AB((warning AND octagon*) OR "octagonal black system*" OR wobs OR "ley de etiquetado de alimentos" OR "ley de alimentos" OR (law AND "20.606") OR "Super 8" OR "Nutri-score*" OR "Nutri score*" OR "Nutriscore*" OR "Health star*" OR (hsr AND system*) OR ("traffic light*" AND (label* OR food* OR nutri* OR diet* OR pack*)) OR ("Warning label*" AND (food* OR nutri* OR diet*))) OR TI,AB(food N/1 pack*) OR TI,AB("recommended dietary allowance*" AND (label* OR "content* sign*" OR symbol* OR ticket* OR sticker* OR vignette*)) OR TI,AB(salt AND (label* OR "content* sign*" OR symbol* OR tag* OR ticket* OR sticker* OR vignette*)) OR TI,AB(sugar* AND (label* OR "content* sign*" OR symbol* OR tag* OR ticket* OR sticker* OR vignette*)))Limits applied 2016-2019

Cochrane Library
08/04/2019
No of records: 351 from Cochrane Central Database of Controlled Trials, and 6 from Cochrane Database of Systematic Reviews
ID Search Hits
#1 MeSH descriptor: [Food Packaging] explode all trees 130
#2 (((product NEAR/2 label*) and food*) or "front of pack*" or "food label*" or ((prepack* NEAR/1 food*) and label*) or (("pre-pack*" NEAR/1 food*) and label*) or (("pre pack*" NEAR/1 food*) and label*)):ti,ab,kw 122
#3 (food NEAR/1 pack*):ti,ab,kw 153
#4 MeSH descriptor: [Product Labeling] explode all trees 70
#5 (food* or fat* or sugar* or salt or diet* or health* or calori* or nutritio* or "daily amount*" or "recommended daily amount*" or "reference value*" or "reference intake*" or "daily value*" or "reference nutrient intake*" or snack* or eat*):ti,ab,kw 424113
#6 #4 AND #5 56
#7 MeSH descriptor: [Food Labeling] explode all trees 107
#8 ((Nutriti* or Nutrient*) NEAR/5 (label* or "content* sign*" or symbol* or ticket* or sticker* or warning* or vignette*)):ti,ab,kw 178
#9 ((nutrition* information or nutrient* information) and (pack* or label* or prepack* or "pre pack*" or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)):ti,ab,kw 484
#10 ("Food* label*" or "food* content* label*" or "food* content* sign*" or "food* content symbol*" or "food* content* tag*" or "food* content* ticket*" or "food* content* sticker*"):ti,ab,kw 36
#11 ((warning NEAR/2 octagon*) or "octagonal black system*" or WOBS or "ley de etiquetado de alimentos" or "ley de alimentos" or (Law NEAR/2 "20.606") or "Super 8" or "Nutri-score*" or "Nutri score*" or "Nutriscore*" or "Health star*" or (HSR NEAR/3 system*) or ("traffic light*" and (label* or food* or nutri* or diet* or pack*)) or ("Warning label*" and (food* or nutri* or diet*))):ti,ab,kw 197
#12 ((keyhole or "key hole") NEAR/5 (Nordic* or label* or sign* or symbol* or vignette*)):ti,ab,kw 2
#13 (N?kkelhullet or N?glehullet or Nyckelh?let or "ley del S?per Ocho" or "Ley S?per 8"):ti,ab,kw 0
#14 (("guideline daily amount*" or "nutrient reference*" or "reference intake*" or "nutrient intake*" or "daily value*") NEAR/5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)):ti,ab,kw 36

#15 (recommended dietary allowance* NEAR/5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)):ti,ab,kw 5
#16 ((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) NEAR/5 (label* or "content* sign*" or symbol* or ticket* or sticker* or vignette*)):ti,ab, 166
#17 (((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) NEXT information) and (pack* or label* or prepack* or "pre-pack*" or "pre pack*")):ti,ab,kw 27
#18 ((fat or fats) NEAR/5 (label* or "content* sign*" or symbol* or tag* or ticket* or sticker* or vignette*)):ti,ab,kw 119
#19 (salt NEAR/5 (label* or "content* sign*" or symbol* or tag* or ticket* or sticker* or vignette*)):ti,ab,kw 16
#20 (sugar* NEAR/5 (label* or "content* sign*" or symbol* or tag* or ticket* or sticker* or vignette*)):ti,ab,kw 37
#21 ((Label* NEAR/2 (legislation* or regulation* or policies or policy or law or laws)) and (food* or diet* or nutri*)):ti,ab,kw 14
#22 ("food law" and (label* or pack* or "pre-pack*" or "pre pack*")):ti,ab,kw 0
#23 (drink or drinks or beverage or beverages or soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice* or smoothie* or milkshake* or tea or teas or coffee*):ti,ab,kw 18030
#24 #4 AND #23 6
#25 ((drink* or beverage*) NEAR/2 (label or labelling or labelling or labels)):ti,ab,kw 17
#26 ((soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice or juices or smoothie* or milkshake* or tea or teas or coffee*) NEAR/2 (label or labelling or labelling or labels)):ti,ab,kw 5
#27 #1 OR #2 OR #3 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #24 OR #25 OR #26 1217
#28 ("Food Preferences" or "Food Habits" or "Feeding Behavior" or "Eating"):kw 9442
#29 MeSH descriptor: [Diet] explode all trees 16917
#30 MeSH descriptor: [Choice Behavior] explode all trees 1370
#31 (intak* or consume or consumes or consumption or consumed or eat* or diet*):ti,ab,kw 138710
#32 ((drink* or beverage*) NEAR/5 (preference* or habit* or behavio* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)):ti,ab,kw 2667
#33 (food NEAR/5 (preference* or habit* or behavio* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)):ti,ab,kw 3528
#34 (shop* or store* or supermarket* or market* or outlet* or retail* or "point of purchas*" or purchas* or buy* or sale* or sell*):ti,ab,kw 16826
#35 #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 156563
#36 #27 AND #35 877
#37 #36 with Cochrane Library publication date in The last 2 years, in Cochrane Reviews 6
#38 #36 with Publication Year from 2016 to 2019, in Trials 351

EMBASE
Database: Embase <1980 to 2019 Week 14>
Search Strategy:
----
1 Food packaging/ and (label* or content* sign* or symbol* or ticket* or sticker* or diet* or health* or calori* or nutri* or "daily amount*" or "recommended daily amount*" or "reference value*" or "reference intake*" or "daily value*" or "reference nutrient intake*").ti,ab. (3959)

2 food packaging/ and (nutritional value/ or nutritional requirement/) (729)
3 (((product adj2 label*) and food*) or "front of pack*" or "food label*" or ((prepack* adj1 food*) and label*) or (("pre-pack*" adj1 food*) and label*) or (("pre pack*" adj1 food*) and label*)).ab,ti. (2514)
4 (food adj1 pack*).ab,ti. (2395)
5 ((Nutriti* or Nutrient*) adj5 (label* or content* sign* or symbol* or ticket* or sticker* or warning* or vignette*)).ab,ti. (1773)
6 ((nutrition* information or nutrient* information) and (pack* or label* or prepack* or "pre pack*" or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ti,ab. (552)
7 ("Food* label*" or "food* content* label*" or "food* content* sign*" or "food* content symbol*" or "food* content* tag*" or "food* content* ticket*" or "food* content* sticker*").ab,ti. (1665)
8 ((warning adj2 octagon*) or "octagonal black system*" or WOBS or "ley de etiquetado de alimentos" or "ley de alimentos" or (Law adj2 "20.606") or "Super 8" or "Nutri-score*" or "Nutri score*" or "Nutriscore*" or "Health star*" or (HSR adj3 system*) or ("traffic light*" and (label* or food* or nutri* or diet* or pack*)) or ("Warning label*" and (food* or nutri* or diet*))).ab,ti. (835)
9 ((keyhole or "key hole") adj5 (Nordic* or label* or sign* or symbol* or vignette*)).ab,ti. (102)
10 (N?kkelhullet or N?glehullet or Nyckelh?let or "ley del S?per Ocho" or "Ley S?per 8").ab,ti. (0)
11 (("guideline daily amount*" or "nutrient reference*" or "reference intake*" or "nutrient intake*" or "daily value*") adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (65)
12 (recommended dietary allowance* adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (1)
13 ((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (1697)
14 (((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) adj information) and (pack* or label* or prepack* or "pre-pack*" or "pre pack*")).ab,ti. (118)
15 ((fat or fats) adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (908)
16 (salt adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (364)
17 (sugar* adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (897)
18 ((Label* adj2 (legislation* or regulation* or policies or policy or law or laws)) and (food* or diet* or nutri*)).ti,ab. (479)
19 ("food law" and (label* or pack* or "pre-pack*" or "pre pack*")).ab,ti. (20)
20 Product labelling/ and (drink or drinks or beverage or beverages or soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice* or smoothie* or milkshake* or tea or teas or coffee*).ab,ti. (0)
21 ((drink* or beverage*) adj2 (label or labelling or labelling or labels)).ab,ti. (75)
22 ((soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice or juices or smoothie* or milkshake* or tea or teas or coffee*) adj2 (label or labelling or labelling or labels)).ab,ti. (9)
23 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 (12397)
24 (intak* or consume or consumes or consumption or consumed or eat* or diet*).ti,ab. (1141118)
25 ((drink? or beverage?) adj5 (preference* or habit* or behavio?r* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)).ab,ti. (4106)

26 (food adj5 (preference* or habit* or behavio?r* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)).ab,ti. (32813)
27 (shop* or store* or supermarket* or market* or outlet* or retail* or "point of purchas*" or purchas* or buy* or sale* or sell*).ab,ti. (426446)
28 food preference/ (12313)
29 feeding behavior/ or drinking behavior/ or eating habit/ or portion size/ (123502)
30 food intake/ or eating/ or energy consumption/ (151262)
31 exp *diet/ (85163)
32 health behavior/ and (food$ or eat$ or diet$).ti,ab. (10547)
33 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 (1631887)
34 23 and 33 (6053)
35 limit 34 to (conference abstracts or embase) (4228)
36 (2017* or 2018* or 2019*).dd. (2156867)
37 (2017* or 2018* or 2019*).dc. (3959164)
38 36 or 37 (3987218)
39 35 and 38 (907)

HMIC
Database: HMIC Health Management Information Consortium <1979 to January 2019>
Search Strategy:
----
1 Food packaging/ and (label* or content* sign* or symbol* or ticket* or sticker* or diet* or health* or calori* or nutri* or "daily amount*" or "recommended daily amount*" or "reference value*" or "reference intake*" or "daily value*" or "reference nutrient intake*").ti,ab,nt. (46)
2 (food packaging/ or product labelling/) and (nutritional value/ or nutritional requirements/) (38)
3 (((product adj2 label*) and food*) or "front of pack*" or "food label*" or ((prepack* adj1 food*) and label*) or (("pre-pack*" adj1 food*) and label*) or (("pre pack*" adj1 food*) and label*)).ab,ti,nt. (81)
4 (food adj1 pack*).ab,ti,nt. (26)
5 Product labelling/ and (food* or fat* or sugar* or salt or diet* or health* or calori* or nutritio* or "daily amount*" or "recommended daily amount*" or "reference value*" or "reference intake*" or "daily value*" or "reference nutrient intake*" or snack* or eat*).ti,ab,nt. (203)
6 ((Nutriti* or Nutrient*) adj5 (label* or content* sign* or symbol* or ticket* or sticker* or warning* or vignette*)).ab,ti,nt. (67)
7 ((nutrition* information or nutrient* information) and (pack* or label* or prepack* or "pre pack*" or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ti,ab,nt. (24)
8 (Food* label* or food* content* label* or food* content* sign* or food* content symbol* or food* content* tag* or food* content* ticket* or food* content* sticker*).ab,ti,nt. (61)
9 ((warning adj2 octagon*) or "octagonal black system*" or WOBS or "ley de etiquetado de alimentos" or "ley de alimentos" or (Law adj2 "20.606") or "Super 8" or "Nutri-score*" or "Nutri score*" or "Nutriscore*" or "Health star*" or (HSR adj3 system*) or ("traffic light*" and (label* or food* or nutri* or diet* or pack*)) or ("Warning label*" and (food* or nutri* or diet*))).ab,ti,nt. (36)
10 ((keyhole or "key hole") adj5 (Nordic* or label* or sign* or symbol* or vignette*)).ab,ti,nt. (0)
11 (N?kkelhullet or N?glehullet or Nyckelh?let or "ley del S?per Ocho" or "Ley S?per 8").ab,ti,nt. (0)
12 (("guideline daily amount*" or "nutrient reference*" or "reference intake*" or "nutrient intake*" or "daily value*") adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti,nt. (4)

13 (recommended dietary allowance* adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti,nt. (0)
14 ((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti,nt. (27)
15 (((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) adj information) and (pack* or label* or prepack* or "pre-pack*" or "pre pack*")).ab,ti,nt. (4)
16 ((fat or fats) adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti,nt. (11)
17 (salt adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti,nt. (1)
18 (sugar* adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (3)
19 ((Label* adj2 (legislation* or regulation* or policies or policy or law or laws)) and (food* or diet* or nutri*)).ti,ab,nt. (16)
20 ("food law" and (label* or pack* or "pre-pack*" or "pre pack*")).ab,ti,nt. (0)
21 Product labelling/ and (drink or drinks or beverage or beverages or soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice* or smoothie* or milkshake* or tea or teas or coffee*).ab,ti,nt. (21)
22 ((drink* or beverage*) adj2 (label or labelling or labelling or labels)).ab,ti,nt. (5)
23 ((soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice or juices or smoothie* or milkshake* or tea or teas or coffee*) adj2 (label or labelling or labelling or labels)).ab,ti,nt. (0)
24 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 (340)
25 exp Food Habits/ or Feeding Behavior/ or Diet/ (2697)
26 (intak* or consume or consumes or consumption or consumed or eat* or diet*).ti,ab,nt. (9397)
27 ((drink? or beverage?) adj5 (preference* or habit* or behavio?r* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)).ab,ti,nt. (98)
28 (food adj5 (preference* or habit* or behavio?r* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)).ab,ti,nt. (443)
29 (shop* or store* or supermarket* or market* or outlet* or retail* or "point of purchas*" or purchas* or buy* or sale* or sell*).ab,ti,nt. (18197)
30 25 or 26 or 27 or 28 or 29 (27576)
31 24 and 30 (192)
32 (2017* or 2018* or 2019*).up. (11038)
33 31 and 32 (8)
34 from 33 keep 1-8 (8)

Medline
9/4/2019
1288 records
Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) <1946 to April 08, 2019>
1 Food packaging/ and (label* or content* sign* or symbol* or ticket* or sticker* or diet* or health* or calori* or nutri* or "daily amount*" or "recommended daily amount*" or "reference value*" or "reference intake*" or "daily value*" or "reference nutrient intake*").ti,ab. (775)
2 (food packaging/ or Product labeling/) and nutritive value/ (120)

3 (((product adj2 label*) and food*) or "front of pack*" or "food label*" or ((prepack* adj1 food*) and label*) or (("pre-pack*" adj1 food*) and label*) or (("pre pack*" adj1 food*) and label*)).ab,ti. (1846)
4 (food adj1 pack*).ab,ti. (2193)
5 Product labeling/ and (food* or fat* or sugar* or salt or diet* or health* or calori* or nutritio* or "daily amount*" or "recommended daily amount*" or "reference value*" or "reference intake*" or "daily value*" or "reference nutrient intake*" or snack* or eat*).ti,ab. (878)
6 Food Labeling/ (3400)
7 ((Nutriti* or Nutrient*) adj5 (label* or content* sign* or symbol* or ticket* or sticker* or warning* or vignette*)).ab,ti. (1344)
8 ((nutrition* information or nutrient* information) and (pack* or label* or prepack* or "pre pack*" or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ti,ab. (391)
9 (Food* label* or food* content* label* or food* content* sign* or food* content symbol* or food* content* tag* or food* content* ticket* or food* content* sticker*).ab,ti. (1174)
10 ((warning adj2 octagon*) or "octagonal black system*" or WOBS or "ley de etiquetado de alimentos" or "ley de alimentos" or (Law adj2 "20.606") or "Super 8" or "Nutri-score*" or "Nutri score*" or "Nutriscore*" or "Health star*" or (HSR adj3 system*) or ("traffic light*" and (label* or food* or nutri* or diet* or pack*)) or ("Warning label*" and (food* or nutri* or diet*))).ab,ti. (650)
11 ((keyhole or "key hole") adj5 (Nordic* or label* or sign* or symbol* or vignette*)).ab,ti. (87)
12 (N?kkelhullet or N?glehullet or Nyckelh?let or "ley del S?per Ocho" or "Ley S?per 8").ab,ti. (1)
13 (("guideline daily amount*" or "nutrient reference*" or "reference intake*" or "nutrient intake*" or "daily value*") adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (54)
14 (recommended dietary allowance* adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (1)
15 ((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (1501)
16 (((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) adj information) and (pack* or label* or prepack* or "pre-pack*" or "pre pack*")).ab,ti. (89)
17 ((fat or fats) adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (857)
18 (salt adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (344)
19 (sugar* adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (923)
20 ((Label* adj2 (legislation* or regulation* or policies or policy or law or laws)) and (food* or diet* or nutri*)).ti,ab. (360)
21 ("food law" and (label* or pack* or "pre-pack*" or "pre pack*")).ab,ti. (11)
22 Product labeling/ and (drink or drinks or beverage or beverages or soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice* or smoothie* or milkshake* or tea or teas or coffee*).ab,ti. (73)
23 ((drink* or beverage*) adj2 (label or labelling or labelling or labels)).ab,ti. (61)
24 ((soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice or juices or smoothie* or milkshake* or tea or teas or coffee*) adj2 (label or labelling or labelling or labels)).ab,ti. (8)
25 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 (12128)

26 Food Preferences/ or Food Habits/ or Feeding Behavior/ or Eating/ or exp Diet/ or exp Choice Behavior/ (407570)
27 (intak* or consume or consumes or consumption or consumed or eat* or diet*).ti,ab. (937209)
28 ((drink? or beverage?) adj5 (preference* or habit* or behavio?r* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)).ab,ti. (3093)
29 (food adj5 (preference* or habit* or behavio?r* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)).ab,ti. (26741)
30 (shop* or store* or supermarket* or market* or outlet* or retail* or "point of purchas*" or purchas* or buy* or sale* or sell*).ab,ti. (335772)
31 (2017* or 2018* or 2019*).ez. (2481506)
32 26 or 27 or 28 or 29 or 30 (1421696)
33 25 and 32 (5233)
34 31 and 33 (879)
35 (2017* or 2018* or 2019*).dt,ed. (3643985)
36 33 and 35 (1288)
37 34 or 36 (1288)

PsycINFO
<1806 to April Week 1 2019>
265 records
1 exp Food/ and (label* or content* sign* or symbol* or ticket* or sticker*).ti,ab. (749)
2 (((product adj2 label*) and food*) or "front of pack*" or "food label*" or ((prepack* adj1 food*) and label*) or (("pre-pack*" adj1 food*) and label*) or (("pre pack*" adj1 food*) and label*)).ab,ti. (437)
3 (food adj1 pack*).ab,ti. (228)
4 ((Nutriti* or Nutrient*) adj5 (label* or content* sign* or symbol* or ticket* or sticker* or warning* or vignette*)).ab,ti. (383)
5 ((nutrition* information or nutrient* information) and (pack* or label* or prepack* or "pre pack*" or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ti,ab. (205)
6 ("Food* label*" or "food* content* label*" or "food* content* sign*" or "food* content symbol*" or "food* content* tag*" or "food* content* ticket*" or "food* content* sticker*").ab,ti. (274)
7 ((warning adj2 octagon*) or "octagonal black system*" or WOBS or "ley de etiquetado de alimentos" or "ley de alimentos" or (Law adj2 "20.606") or "Super 8" or "Nutri-score*" or "Nutri score*" or "Nutriscore*" or "Health star*" or (HSR adj3 system*) or ("traffic light*" and (label* or food* or nutri* or diet* or pack*)) or ("Warning label*" and (food* or nutri* or diet*))).ab,ti. (213)
8 ((keyhole or "key hole") adj5 (Nordic* or label* or sign* or symbol* or vignette*)).ab,ti. (4)
9 (N?kkelhullet or N?glehullet or Nyckelh?let or "ley del S?per Ocho" or "Ley S?per 8").ab,ti. (0)
10 (("guideline daily amount*" or "nutrient reference*" or "reference intake*" or "nutrient intake*" or "daily value*") adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (12)
11 ("recommended dietary allowance*" adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (0)
12 ((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) adj5 (label* or content* sign* or symbol* or ticket* or sticker* or vignette*)).ab,ti. (260)
13 (((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ* or energy) adj information) and (pack* or label* or prepack* or "pre-pack*" or "pre pack*")).ab,ti. (73)
14 ((fat or fats) adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (87)

15 (salt adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (12)
16 (sugar* adj5 (label* or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)).ab,ti. (32)
17 ((Label* adj2 (legislation* or regulation* or policies or policy or law or laws)) and (food* or diet* or nutri*)).ti,ab. (60)
18 ("food law" and (label* or pack* or "pre-pack*" or "pre pack*")).ab,ti. (1)
19 ((drink* or beverage*) adj2 (label or labelling or labelling or labels)).ab,ti. (40)
20 ((soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice or juices or smoothie* or milkshake* or tea or teas or coffee*) adj2 (label or labelling or labelling or labels)).ab,ti. (3)
21 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 (1786)
22 (intak* or consume or consumes or consumption or consumed or eat* or diet*).ti,ab. (159591)
23 ((drink? or beverage?) adj5 (preference* or habit* or behavio?r* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)).ab,ti. (1911)
24 (food adj5 (preference* or habit* or behavio?r* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)).ab,ti. (13375)
25 (shop* or store* or supermarket* or market* or outlet* or retail* or "point of purchas*" or purchas* or buy* or sale* or sell*).ab,ti. (122422)
26 exp Food Preferences/ or exp Eating Behavior/ or exp Food Intake/ or exp Eating/ or exp Diets/ or exp Choice Behavior/ (69457)
27 22 or 23 or 24 or 25 or 26 (309125)
28 21 and 27 (1293)
29 (2017* or 2018* or 2019*).up. (405135)
30 28 and 29 (265)

Web of Science
Social Sciences Citation Index, Science Citation Index, Emerging Sources Citation Index
8/4/2019
902 items

# 31	902	#29 AND #23 Refined by: PUBLICATION YEARS: (2019 OR 2018 OR 2017 OR 2016) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 30	3,091	#29 AND #23 Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 29	796,594	#28 OR #27 OR #26 OR #25 OR #24 Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 28	295,048	TI=(shop* or store* or supermarket* or market* or outlet* or retail* or "point of purchas*" or purchas* or buy* or sale* or sell*) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 27	13,671	TI=(food* NEAR/5 (preference* or habit* or behavi* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 26	3,920	TI=((drink or drinks or beverage or beverages) NEAR/5 (preference* or habit* or behavio* or choice* or decision* or decid* or inclin* or lik* or choos* or select* or pick*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years

# 25	484,996	TI=(intak* or consume or consumes or consumption or consumed or eat* or diet*) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 24	333,718	TI=("food preference*" or "Food Habit*" or "Feeding Behav*" or "Eating" or "Diet*" or "Choice Behav*") Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 23	31,317	#22 OR #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 22	170	TS=((soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice or juices or smoothie* or milkshake* or tea or teas or coffee*) NEAR/2 (label or labelling or labelling or labels)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 21	250	TS=((drink* or beverage*) NEAR/2 (label or labelling or labelling or labels)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 20	56	TS=("product packag*") and TS=(drink or drinks or beverage or beverages or soda or sodas or "flavored water*" or "flavoured water*" or "fruit water*" or cordial or cordials or squash or squashes or juice* or smoothie* or milkshake* or tea or teas or coffee*) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 19	42	TS=("food law" and (label* or pack* or "pre-pack*" or "pre pack*")) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 18	810	TS=((Label* NEAR/2 (legislation* or regulation* or policies or policy or law or laws)) and (food* or diet* or nutri*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 17	1,120	TS=(sugar* NEAR/5 (label* or "content* sign*" or symbol* or tag* or ticket* or sticker* or vignette*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 16	632	TS=(salt NEAR/5 (label* or "content* sign*" or symbol* or tag* or ticket* or sticker* or vignette*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 15	1,100	TS=((fat or fats) NEAR/5 (label* or "content* sign*" or symbol* or tag* or ticket* or sticker* or vignette*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 14	282	TS=(((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ or "energy") NEAR/1 information) and (pack* or label* or prepack* or "pre-pack*" or "pre pack*")) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 13	2,927	TS=((Calorific or calorie* or caloric or kilojoule* or kilocalorie* or kcal* or kJ or "energy") NEAR/5 (label* or "content* sign*" or symbol* or ticket* or sticker* or vignette*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years

# 12	41	TS=(recommended dietary allowance* NEAR/5 (label* or "content* sign*" or symbol* or ticket* or sticker* or vignette*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 11	79	TS=(("guideline daily amount*" or "nutrient reference*" or "reference intake*" or "nutrient intake*" or "daily value*") NEAR/5 (label* or "content* sign*" or symbol* or ticket* or sticker* or vignette*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 10	43	TS=(("keyhole" or "key hole") NEAR/5 (Nordic* or label* or "sign" or "signs" or "signage" or symbol* or vignette*)) OR TS=(N$kkelhullet or N$glehullet or Nyckelh$let or "ley del S$per Ocho" or "Ley S$per 8") Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 9	990	TS=((warning NEAR/2 octagon*) or "octagonal black system*" or WOBS or "ley de etiquetado de alimentos" or "ley de alimentos" or (Law NEAR/2 "20.606") or "Super 8" or "Nutri-score*" or "Nutri score*" or "Nutriscore*" or "Health star*" or (HSR NEAR/3 system*) or ("traffic light*" and (label* or food* or nutri* or diet* or pack*)) or ("Warning label*" and (food* or nutri* or diet*))) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 8	2,307	TS=("Food* label*" or "food* content* label*" or "food* content* sign*" or "food* content symbol*" or "food* content* tag*" or "food* content* ticket*" or "food* content* sticker*") Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 7	877	TS=(("nutriti* information" or "nutrient* information") and (pack* or label* or prepack* or "pre pack*" or content* sign* or symbol* or tag* or ticket* or sticker* or vignette*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 6	2,287	TS=(("Nutriti*" or "Nutrient*") NEAR/5 (label* or "content* sign*" or symbol* or ticket* or sticker* or warning* or vignette*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 5	2,250	TS="Food Label*" Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 4	15,355	TI= (label*) and TS=(food* or fat* or sugar* or salt or diet* or health* or calori* or nutritio* or "daily amount*" or "recommended daily amount*" or "reference value*" or "reference intake*" or "daily value*" or "reference nutrient intake*" or snack* or eat*) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 3	7,388	TS=(food NEAR/1 pack*) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 2	3,849	TS=(((product NEAR/2 label*) and food*) or "front of pack*" or "food label*" or ((prepack* NEAR/1 food*) and label*) or (("pre-pack*" NEAR/1 food*) and label*) or (("pre pack*" NEAR/1 food*) and label*)) Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years
# 1	1,298	TS=("Food packag*") and TS=(label* or content* sign* or symbol* or ticket* or sticker* or diet* or health* or calori* or nutri* or "daily amount*" or "recommended daily amount*" or "reference value*" or "reference intake*" or "daily value*" or "reference nutrient intake*")

Indexes=SCI-EXPANDED, SSCI, ESCI Timespan=All years

SCOPUS
8/4/2019
1,207 records
(TITLE (shop* OR store* OR supermarket* OR market* OR outlet* OR retail* OR "point of purchas*" OR purchas* OR buy* OR sale* OR sell*) OR TITLE (food* W/5 (preference* OR ha bit* OR behavi* OR choice* OR decision* OR decid* OR inclin* OR lik* OR choos* OR select* OR pick*)) OR TITLE ((drink OR drinks OR beverage OR beverages) W/5 (preference* OR ha bit* OR behavio* OR choice* OR decision* OR decid* OR inclin* OR lik* OR choos* OR select * OR pick*)) OR TITLE (intak* OR consume OR consumes OR consumption OR consumed OR eat* OR diet*) OR TITLE ("food preference*" OR "Food Habit*" OR "Feeding
Behav*" OR "Eating" OR "Diet*" OR "Choice Behav*")) AND ((TITLE (label*) AND TITLE-ABS-KEY (food* OR fat* OR sugar* OR salt OR diet* OR health* OR calori* OR nutritio* OR "daily amount*" OR "recommended daily amount*" OR "reference value*" OR "reference
intake*" OR "daily value*" OR "reference nutrient intake*" OR snack* OR eat*)) OR TITLE-ABS-KEY ("Food Label*") OR TITLE-ABS-KEY ("food law" AND (label* OR pack* OR "pre-
pack*" OR "pre pack*")) OR (TITLE-ABS-KEY ("Food packag*") AND TITLE-ABS-
KEY (label* OR "content*
sign*" OR symbol* OR ticket* OR sticker* OR diet* OR health* OR calori* OR nutri* OR "daily amount*" OR "recommended daily amount*" OR "reference value*" OR "reference
intake*" OR "daily value*" OR "reference nutrient intake*")) OR TITLE-ABS-KEY ("Food*
label*" OR "food* content* label*" OR "food* content* sign*" OR "food* content
symbol*" OR "food* content* tag*" OR "food* content* ticket*" OR "food* content*
sticker*") OR (TITLE-ABS-KEY ("product packag*") AND TITLE-ABS-
KEY (drink OR drinks OR beverage OR beverages OR soda OR sodas OR "flavored
water*" OR "flavoured water*" OR "fruit
water*" OR cordial OR cordials OR squash OR squashes OR juice* OR smoothie* OR milkshake * OR tea OR teas OR coffee*)) OR TITLE-ABS-KEY (("guideline daily amount*" OR "nutrient reference*" OR "reference intake*" OR "nutrient intake*" OR "daily
value*") W/5 (label* OR "content*
sign*" OR symbol* OR ticket* OR sticker* OR vignette*)) OR TITLE-ABS-
KEY (("keyhole" OR "key
hole") W/5 (nordic* OR label* OR "sign" OR "signs" OR "signage" OR symbol* OR vignette*)) OR TITLE-ABS-KEY (*kkelhullet OR *glehullet OR nyckelh* OR "ley del Super Ocho" OR "Ley Super 8") OR TITLE-ABS-KEY (("nutriti* information" OR "nutrient*
information") AND (pack* OR label* OR prepack* OR "pre pack*" OR "content*
sign*" OR symbol* OR tag* OR ticket* OR sticker* OR vignette*)) OR TITLE-ABS-
KEY (("Nutriti*" OR "Nutrient*") W/5 (label* OR "content*
sign*" OR symbol* OR ticket* OR sticker* OR warning* OR vignette*)) OR TITLE-ABS-
KEY (((calorific OR calorie* OR caloric OR kilojoule* OR kilocalorie* OR kcal* OR kj OR "energ y") W/1 information) AND (pack* OR label* OR prepack* OR "pre-pack*" OR "pre
pack*")) OR TITLE-ABS-KEY (((product W/2 label*) AND food*) OR "front of pack*" OR "food label*" OR ((prepack* W/1 food*) AND label*) OR (("pre-
pack*" W/1 food*) AND label*) OR (("pre pack*" W/1 food*) AND label*)) OR TITLE-ABS-
KEY ((calorific OR calorie* OR caloric OR kilojoule* OR kilocalorie* OR kcal* OR kj OR "energy ") W/5 (label* OR "content*
sign*" OR symbol* OR ticket* OR sticker* OR vignette*)) OR TITLE-ABS-

KEY ((drink* OR beverage*) W/2 (label OR labelling OR labelling OR labels)) OR TITLE-ABS-KEY ((fat OR fats) W/5 (label* OR "content* sign*" OR symbol* OR tag* OR ticket* OR sticker* OR vignette*)) OR TITLE-ABS-KEY ((label* W/2 (legislation* OR regulation* OR policies OR policy OR law OR laws)) AND (food* OR diet* OR nutri*)) OR TITLE-ABS-KEY ((soda OR sodas OR "flavored water*" OR "flavoured water*" OR "fruit water*" OR cordial OR cordials OR squash OR squashes OR juice OR juices OR smoothie* OR milkshake* OR tea OR teas OR coffee*) W/2 (label OR labelling OR labelling OR labels)) OR TITLE-ABS-KEY ((warning W/2 octagon*) OR "octagonal black system*" OR wobs OR "ley de etiquetado de alimentos" OR "ley de alimentos" OR (law W/2 "20.606") OR "Super 8" OR "Nutri-score*" OR "Nutri score*" OR "Nutriscore*" OR "Health star*" OR (hsr W/3 system*) OR ("traffic light*" AND (label* OR food* OR nutri* OR diet* OR pack*)) OR ("Warning label*" AND (food* OR nutri* OR diet*))) OR TITLE-ABS-KEY (food W/1 pack*) OR TITLE-ABS-KEY ("recommended dietary allowance*" W/5 (label* OR "content* sign*" OR symbol* OR ticket* OR sticker* OR vignette*)) OR TITLE-ABS-KEY (salt W/5 (label* OR "content* sign*" OR symbol* OR tag* OR ticket* OR sticker* OR vignette*)) OR TITLE-ABS-KEY (sugar* W/5 (label* OR "content* sign*" OR symbol* OR tag* OR ticket* OR sticker* OR vignette*)) AND LOAD-DATE > 20170425) AND (LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016))
TRoPHI 137 records 1. Coded with: healthy eating, OR obesity 2. "food* OR nutri* OR drink OR drinks OR beverage OR beverages" (in Title and Abstract) 3. "pack* OR prepack* OR "pre-pack*" " (in Title and Abstract) 4. "label* OR sign OR signs OR symbol OR symbols OR vignette* OR ticket OR tickets OR sticker* " (in Title and Abstract) 5. "2019" (in Year) OR "2018" (in Year) OR "2017" (in Year) OR "2016" (in Year) 6. (1 OR 2) AND (2OR 4) AND 5

Appendix C: Experiment Design and Methodology (Chapter 4)

NatCen Panel Technical Information

NatCen
Social Research that works for society

NatCen Panel – October 2020

Technical information – Front of Package Labelling

Date: December 2020

At **NatCen Social Research** we believe that social research has the power to make life better. By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that can make a difference to everyone. And as an independent, not for profit organisation we're able to put all our time and energy into delivering social research that works for society.

NatCen Social Research
35 Northampton Square
London EC1V 0AX
T 020 7250 1866
www.natcen.ac.uk

A Company Limited by Guarantee
Registered in England No.4392418.
A Charity registered in England and Wales (1091768) and Scotland (SC038454)

Contents

Introduction

In October and November 2020, NatCen conducted a survey amongst its panel members in collaboration with UCL and on behalf of the DHSC to investigate the impact of front of package labelling (FoPL) on knowledge

This document outlines some of the technical information related to the survey and associated dataset, including what data are included, questionnaire wording, the sample design, fieldwork approach, and information on the weighting approach.

If you have any questions about the data, or the information in this document, please get in touch at panel.info@natcen.ac.uk.

Key figures

Fieldwork dates	29/10/20 –15/11/2020
Total complete interviews	4,863
Web interviews	4,863 (100%)
Telephone interviews	0 (0%)
Survey response rate	67%
Overall response rate	12%

1 Survey dataset

This section summarises the information included in the survey dataset.

1.1 Sampling & weighting information

1.1.1 Sampling information

As the sample design for the British Social Attitudes Survey (BSA) (and therefore the NatCen Panel which is recruited via the BSA survey) involves stratification and clustering (see Section 2), these design features affect standard errors and should therefore be taken into account in analysis, and variables are provided to allow for this:

- Panel_PSU indicates the Primary Sampling Unit from which the panellist was recruited in their BSA survey
- Panel_Strata indicates the sampling stratum from which the panellist was selected. This is a simplified version of the strata ID variable used in BSA. Each code indicates a different combination of BSA year and region

In addition, at the sampling stage for this fieldwork wave, panel members selected to be invited to take part in the survey were randomly allocated to five groups which determined which version of the FoPL they would receive. This is recorded in the variable Oct20SampSplit1.

Before being randomly allocated to these groups, the sample was stratified based on their year of recruitment to the Panel, sex, age group, household income, and region. These stratification variables are recorded in StratVar1 to StratVar2.

1.1.2 Non-response weight[1]

As a random probability sample, estimates are affected by non-coverage and non-response. In order to ensure the sample is representative of the population, a set of non-response weights has been computed to account for non-response to the recruitment survey (BSA), refusal to the join the panel, and non-response in the survey of panel members itself.

Oct20_Weight is the product of these three weights and should be used for the general analysis of the survey results.

1.2 Survey paradata

Two pieces of survey paradata are included in the survey dataset:

- **Oct20_IntDate** gives the date on which the survey was completed
- **Oct20_IntMode** gives the mode in which the survey was completed (online or on the telephone). At this wave, all participants completed online

[1] More information about the non-response weight is provided in Section 3

1.3 Questionnaire data

The datasets include all substantive questions carried by the survey and funded by the DHSC. Variables from this wave of the questionnaire are indicated in the variable name with the preface '**Oct200_DHSC_**'. The questionnaire content is shown in Section 4, indicating variable names.

1.4 Fed-forward data

One of the features of the NatCen Panel is that, because it was recruited via the face-to-face British Social Attitudes (BSA) survey, and our panellists are interviewed regularly, we possess a wealth of background information on our panellists. This not only allows us to develop better response propensity models (Section 3), but also allows for analysis by a greater range of background variables[2].

These variables are indicated with the preface 'FF_' where the data is directly from the BSA questionnaire, or 'Cur_' where it has been subsequently updated.

The following variables are included in this dataset as standard:

- Fed-forward sex
- Fed-forward respondent's NS-SEC analytic class (grouped) (DV)
- Fed-forward whether respondent has long-standing condition that affects day-to-day life (DV)
- Latest age category (grouped) (DV)
- Latest highest educational qualification achieved
- Latest class identity
- Latest main economic activity (grouped)
- Latest subjective income
- Latest household income (grouped)
- Latest household income - equivalised (grouped)
- Latest number of people in household (grouped) (DV)
- Latest household type (DV)
- Latest relationship status (grouped)
- Latest whether respondent has any children (0-18) in the household (DV)
- Latest urban/rural indicator 2011 (England & Wales) (grouped) (DV)
- Latest urban/rural indicator 2011 (Scotland) (grouped) (DV)
- Latest government office region (DV)
- Latest frequency of internet use (grouped)
- Latest political party identification (grouped) (DV)
- Latest ethnic group (grouped) (DV)

[2] It should be noted that the data for these variables may have been collected before the panel survey was conducted.

2 Sampling and fieldwork

NatCen's panel is based on a random probability design, with panel members originally selected at random and considerable effort put in to maximise participation in order to minimise bias.

2.1 BSA recruitment

Panel members are recruited from the British Social Attitudes survey (BSA) which interviews those aged 18 and over across Britain (south of the Caledonian canal). The BSA is a high-quality, random probability face-to-face survey: this means that households and individuals are selected at random, and then considerable effort is expended by field interviewers to achieve an interview, including visiting the selected addresses multiple times.[3]

Those interviewed as part of the BSA were asked to join the Panel at the end of the BSA interview. For this survey, all panel members recruited from BSA 2015 and 2019 who had not subsequently left the panel or become 'inactive' were approached to participate (no quotas were used) and the random probability design was therefore maintained.

2.2 Panel fieldwork

2.2.1 Fieldwork design

At this wave, the panel survey was conducted using a web-only fieldwork approach, with panel members invited to participate in the research online (using multiple points of contact by post, email and text). A £5 love to shop gift card was sent as a 'thank you' to those who participated.

At this wave we have targeted our design somewhat to improve the sample quality. Effort was balanced away from those who are typically over-represented in the sample and that take part regularly, and targeted at those who are under-represented in the sample and take part less regularly.

Reflecting the short time-frame for this project and the lack of telephone fieldwork, the fieldwork period was shorter than a 'typical' panel wave, lasting for 17 days – from 29th October to 15th November. This ensured that we could make repeated attempts to contact the selected individuals to try to secure their participation, rather than only including those that are 'readily' available.

2.2.2 Response rates

The probability design allows us to apply statistical theory to the study, including tests of statistical significance or the 'margin of error'. Response rates are a simple indicator of quality for surveys of this sort and are provided in Table 2:1. This survey achieved a 67% response rate among those panellists invited to participate. When taking account of non-response at the BSA interview and then also at the point of recruitment to the panel, our overall response rate was 12%. This is slightly lower than a typical panel wave, most likely reflecting the shorter fieldwork period and lack of telephone fieldwork.

[3] More details of the BSA sampling approach can be found here: http://bsa.natcen.ac.uk/

Table 2:1 Survey response	
Response to the survey	
Issued	7,218
Deadwood	3
Achieved	4,863
Survey response rate	**67%**
Overall response	
BSA issued	44,356
BSA deadwood	4,159
BSA productive	18,361
Recruited to panel	11,301
BSA response rate	46%
Panel recruitment rate	62%
Panel deadwood	93
Overall survey response rate	**12%**

2.3 Sample profile

2.3.1 Socio-demographics

Table 2:2 Sample profile – socio-demographics

	BSA population estimate (weighted)[4]	**Panel survey estimate (weighted)**	**Panel survey sample (unweighted)**
Sex			
Male	49%	48%	43%
Female	51%	52%	57%
Age			
18-24	11%	11%	4%
25-34	17%	18%	13%
35-44	16%	17%	19%
45-54	18%	19%	21%
55-64	15%	15%	21%
65+	23%	20%	23%

[4] Estimates are based on combined BSA 2015 to 2019 datasets, each weighted to reflect the population at the time.

Region			
North East	4%	4%	4%
North West	11%	11%	11%
Yorkshire and The Humber	8%	8%	10%
East Midlands	7%	8%	10%
West Midlands	9%	8%	8%
East of England	10%	10%	11%
London	13%	13%	8%
South East	14%	14%	17%
South West	9%	9%	10%
Wales	5%	5%	4%
Scotland	9%	9%	7%
Social grade			
Managerial & Professional occupations	39%	40%	52%
Intermediate occupations	12%	14%	14%
Small employers & own account workers	9%	8%	8%
Lower supervisory & technical occupations	8%	8%	6%
Semi-routine & routine occupations	27%	27%	18%
Highest level of education			
Degree	26%	28%	36%
Higher education below degree	11%	12%	14%
A level or equivalent	18%	18%	18%
O level/CSE or equivalent	25%	25%	24%
Foreign or other	2%	2%	1%
No qualifications	17%	15%	7%
Household type			
Single person household	17%	16%	24%
Lone parent	3%	3%	5%
2 adults (no children)	36%	36%	38%
2 adults (with children)	21%	22%	21%
3+ adults (no children)	16%	15%	8%
3+ adults (with children)	7%	7%	3%
Economic activity			
Full time education	5%	4%	2%
Paid work	57%	59%	59%
Unemployed	5%	6%	4%
Retired	24%	21%	26%
Other	10%	9%	9%

Tenure			
Owned/being bought	63%	64%	73%
Rented (LA)	10%	8%	6%
Rented (HA/Trust/New Town)	7%	8%	6%
Rented (Other)	18%	18%	14%
Other	1%	1%	1%
Unweighted base	*18,361*	*4,863*	*4,863*

3 Non-response weights

Non-response for NatCen's probability panel surveys can occur at three stages: non-response at the survey used for recruitment (the British Social Attitudes survey), refusal to join the panel at the end of that interview and non-response in the survey of panel members itself. We compute a weight to account for non-response at each of these three stages. The final weight (Oct20_Weight) is the product of these three weights. We use this three-stage system because the variables underlying non-response could be different at each stage. With this system we also can maximise the use of all the information available from the British Social Attitudes Survey (BSA).

These are the three weights we have computed:

1. **BSA survey weight**: the panel members were recruited from BSA 2015, 2016, 2017, 2018 and 2019. Firstly, the BSA weights account for unequal chances of selection in the BSA sampling. Secondly, a non-response model is used to produce a non-response weight. This weight adjusts for non-response at the BSA survey using: region, type of dwelling, whether there were entry barriers to the selected address, the relative condition of the immediate local area, the relative condition of the address, the percentage of owner-occupied properties in quintiles and population density. Finally, the BSA weights make the sample of BSA respondents representative of the general British population in terms of gender, age and Government Office Region (GOR).[5]
2. **Panel weight**: this weight accounts for non-response at the panel recruitment stage where some people interviewed as part of the BSA survey chose not to join the panel. A logistic regression model has been used to derive the probability of response of each panel member; the panel weight is computed as the inverse of the probabilities of response. This weight adjusts the panel for non-response using the following variables: age and sex groups, GOR, BSA year, household type, household income, education level, internet access, ethnicity, tenure, social class group, economic activity, political party identification, and interest in politics.[6] The resulting panel weight has been multiplied by the BSA weights, so the panel is representative of the population.
3. **Survey weight**: this weight is to adjust the bias caused by non-response to this particular panel survey. A logistic regression model has been used to compute the probabilities of response of each participant. The panel survey weight is equal to the inverse of the probabilities of response. The initial set of predictors used to build the model was the same as for the panel weight; and at this wave the final set of variables used was also the same. Unlike the model used to calculate the panel weight, no interaction term between BSA survey year and internet access was used. As this wave of data collection was web-only, there were only a very small number of panellists (from each year of BSA) who took part in the survey but did not have access to the internet when they were interviewed for BSA. It was therefore deemed inappropriate to include the interaction term in the model.

   The final survey weight is the result of multiplying the survey weight by the compounded panel weight.

[5] More details on the BSA weight can be found at http://bsa.natcen.ac.uk/

[6] The characteristics that are likely to change with time for an individual and whose distribution differed between 2015 and 2019 BSA sample have been entered into the model in interaction with BSA year.

4 Questionnaire specification

4.1 Baseline ranking

{ASK ALL}
BLRankInt
"Our first set of questions will ask you to rank different foods on how healthy you think they are. We would like to know how long it takes for you to rank the different foods, so please click 'next' as soon as you have completed each task.

Please rank them from most healthy to least healthy."

DISPLAY

PROGRAMMER: Randomise order different foods appear.

{ASK ALL}
BLPizzaRnk [RANDOMISE BLPizzaRnk1… BLPizzaRnk3]
"Below are three pictures of different types of pizza. Please select how healthy each pizza is compared to the other 2 pizzas – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

[Pizza1BL image]
BLPizzaRnk1
1. Most healthy
2. In between
3. Least healthy

[Pizza2BL image]
BLPizzaRnk2
1. Most healthy
2. In between
3. Least healthy

[Pizza3BL image]
BLPizzaRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF BLPizzaRnk1 = BLPizzaRnk2 OR BLPizzaRnk1 = BLPizzaRnk3 OR BLPizzaRnk2 = BLPizzaRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
BLPizzaEnInf
"Do you feel you had enough information to rank the pizzas from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
BLHotChocRnk [RANDOMISE BLHotChocRnk1… BLHotChocRnk3]
"Below are three pictures of different types of hot chocolate. Please select how healthy each hot chocolate is compared to the other 2 hot chocolates – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

[HotChoc1BL image]
BLHotChocRnk1
1. Most healthy
2. In between
3. Least healthy

[HotChoc2BL image]
BLHotChocRnk2
1. Most healthy
2. In between
3. Least healthy

[HotChoc3BL image]
BLHotChocRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF BLHotChocRnk1 = BLHotChocRnk2 OR BLHotChocRnk1 = BLHotChocRnk3 OR BLHotChocRnk2 = BLHotChocRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
BLHotChocEnInf
"Do you feel you had enough information to rank the hot chocolates from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
BLCakeRnk [RANDOMISE BLCakeRnk1… BLCakeRnk3]
"Below are three pictures of different types of cake. Please select how healthy each cake is compared to the other 2 cakes – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

[Cake1BL image]
BLCakeRnk1
1. Most healthy
2. In between
3. Least healthy

[Cake2BL image]
BLCakeRnk2
1. Most healthy
2. In between
3. Least healthy

[Cake3BL image]
BLCakeRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF BLCakeRnk1 = BLCakeRnk2 OR BLCakeRnk1 = BLCakeRnk3 OR BLCakeRnk2 = BLCakeRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
BLCakeEnInf
"Do you feel you had enough information to rank the cakes from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
BLCrispsRnk [RANDOMISE BLCrispsRnk1... BLCrispsRnk3]
"Below are three pictures of different types of crisps. Please select how healthy type of crisps is compared to the other 2 types of crisps – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

[Crisps1BL image]
BLCrispsRnk1
1. Most healthy
2. In between
3. Least healthy

[Crisps2BL image]
BLCrispsRnk2
1. Most healthy
2. In between
3. Least healthy

[Crisps3BL image]
BLCrispsRnk3
1. Most healthy

2. In between
3. Least healthy

HARD CHECK: IF BLCrispsRnk1 = BLCrispsRnk2 OR BLCrispsRnk1 = BLCrispsRnk3 OR BLCrispsRnk2 = BLCrispsRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
BLCrispsEnInf
"Do you feel you had enough information to rank the crisps from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
BLYoghRnk [RANDOMISE BLYoghRnk1… BLYoghRnk3]
"Below are three pictures of different types of yoghurt. Please select how healthy each yoghurt is compared to the other 2 yoghurts – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

[Yogh1BL image]
BLYoghRnk1
1. Most healthy
2. In between
3. Least healthy

[Yogh2BL image]
BLYoghRnk2
1. Most healthy
2. In between
3. Least healthy

[Yogh3BL image]
BLYoghRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF BLYoghRnk1 = BLYoghRnk2 OR BLYoghRnk1 = BLYoghRnk3 OR BLYoghRnk2 = BLYoghRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
BLYoghEnInf
"Do you feel you had enough information to rank the yoghurts from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
BLBrkCerRnk [RANDOMISE BLBrkCerRnk1… BLBrkCerRnk3]
"Below are three pictures of different types of breakfast cereal. Please select how healthy each type of breakfast cereal is compared to the other 2 types of breakfast cereal – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

[BrkCer1BL image]
BLBrkCerRnk1
1. Most healthy
2. In between
3. Least healthy

[BrkCer2BL image]
BLBrkCerRnk2
1. Most healthy
2. In between
3. Least healthy

[BrkCer3BL image]
BLBrkCerRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF BLBrkCerRnk1 = BLBrkCerRnk2 OR BLBrkCerRnk1 = BLBrkCerRnk3 OR BLBrkCerRnk2 = BLBrkCerRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
BLBrkCerEnInf
"Do you feel you had enough information to rank the breakfast cereals from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
BLRankConf [FLIP SCALE]
"Thinking about all of the rankings you have just done…

How many do you think you ranked in the correct order – from most healthy to least healthy?"

1. All of them
2. Most of them
3. Some of them
4. None of them
5. Don't know

4.2 Experimental ranking

{ASK ALL}
ExpRankInt
"We will now present you with another set of pictures of the same food items. Remember, we would like to know how long it takes for you to rank the different foods, so please click 'next' as soon as you have completed each task.

Please rank them from most healthy to least healthy."

DISPLAY

PROGRAMMER: Different foods should appear in the same order as previously.

{ASK ALL}
ExpPizzaRnk [RANDOMISE ExpPizzaRnk1… ExpPizzaRnk3]
"Below are three pictures of different types of pizza. Please select how healthy each pizza is compared to the other 2 pizzas – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

{IF Oct20SampSplit1 = 1 [Pizza1BL image]}
{IF Oct20SampSplit1 = 2 [Pizza1MTL image]}
{IF Oct20SampSplit1 = 3 [Pizza1NS image]}
{IF Oct20SampSplit1 = 4 [Pizza1WL image]}
{IF Oct20SampSplit1 = 5 [Pizza1PCT image]}
ExpPizzaRnk1

1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [Pizza2BL image]}
{IF Oct20SampSplit1 = 2 [Pizza2MTL image]}
{IF Oct20SampSplit1 = 3 [Pizza2NS image]}
{IF Oct20SampSplit1 = 4 [Pizza2WL image]}
{IF Oct20SampSplit1 = 5 [Pizza2PCT image]}
ExpPizzaRnk2

1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [Pizza3BL image]}
{IF Oct20SampSplit1 = 2 [Pizza3MTL image]}
{IF Oct20SampSplit1 = 3 [Pizza3NS image]}
{IF Oct20SampSplit1 = 4 [Pizza3WL image]}
{IF Oct20SampSplit1 = 5 [Pizza3PCT image]}
[ExpPizzaRnk3

1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF ExpPizzaRnk1 = ExpPizzaRnk2 OR ExpPizzaRnk1 = ExpPizzaRnk3 OR ExpPizzaRnk2 = ExpPizzaRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
ExpPizzaEnInf
"Do you feel you had enough information to rank the pizzas from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
EXPHotChocRnk [RANDOMISE EXPHotChocRnk1… EXPHotChocRnk3]
"Below are three pictures of different types of hot chocolate. Please select how healthy each hot chocolate is compared to the other 2 hot chocolates – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

{IF Oct20SampSplit1 = 1 [HotChoc1BL image]}
{IF Oct20SampSplit1 = 2 [HotChoc1MTL image]}
{IF Oct20SampSplit1 = 3 [HotChoc1NS image]}
{IF Oct20SampSplit1 = 4 [HotChoc1WL image]}
{IF Oct20SampSplit1 = 5 [HotChoc1PCT image]}

EXPHotChocRnk1
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [HotChoc2BL image]}
{IF Oct20SampSplit1 = 2 [HotChoc2MTL image]}
{IF Oct20SampSplit1 = 3 [HotChoc2NS image]}
{IF Oct20SampSplit1 = 4 [HotChoc2WL image]}
{IF Oct20SampSplit1 = 5 [HotChoc2PCT image]}
EXPHotChocRnk2
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [HotChoc3BL image]}
{IF Oct20SampSplit1 = 2 [HotChoc3MTL image]}
{IF Oct20SampSplit1 = 3 [HotChoc3NS image]}
{IF Oct20SampSplit1 = 4 [HotChoc3WL image]}
{IF Oct20SampSplit1 = 5 [HotChoc3PCT image]}
EXPHotChocRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF EXPHotChocRnk1 = EXPHotChocRnk2 OR EXPHotChocRnk1 = EXPHotChocRnk3 OR EXPHotChocRnk2 = EXPHotChocRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
EXPHotChocEnInf
"Do you feel you had enough information to rank the hot chocolates from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
EXPCakeRnk [RANDOMISE EXPCakeRnk1… EXPCakeRnk3]
"Below are three pictures of different types of cake. Please select how healthy each cake is compared to the other 2 cakes – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

{IF Oct20SampSplit1 = 1 [Cake1BL image]}
{IF Oct20SampSplit1 = 2 [Cake1MTL image]}
{IF Oct20SampSplit1 = 3 [Cake1NS image]}
{IF Oct20SampSplit1 = 4 [Cake1WL image]}
{IF Oct20SampSplit1 = 5 [Cake1PCT image]}

EXPCakeRnk1
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [Cake2BL image]}
{IF Oct20SampSplit1 = 2 [Cake2MTL image]}
{IF Oct20SampSplit1 = 3 [Cake2NS image]}
{IF Oct20SampSplit1 = 4 [Cake2WL image]}
{IF Oct20SampSplit1 = 5 [Cake2PCT image]}
EXPCakeRnk2
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [Cake3BL image]}
{IF Oct20SampSplit1 = 2 [Cake3MTL image]}
{IF Oct20SampSplit1 = 3 [Cake3NS image]}
{IF Oct20SampSplit1 = 4 [Cake3WL image]}
{IF Oct20SampSplit1 = 5 [Cake3PCT image]}
EXPCakeRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF EXPCakeRnk1 = EXPCakeRnk2 OR EXPCakeRnk1 = EXPCakeRnk3 OR EXPCakeRnk2 = EXPCakeRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
EXPCakeEnInf
"Do you feel you had enough information to rank the cakes from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
EXPCrispsRnk [RANDOMISE EXPCrispsRnk1… EXPCrispsRnk3]
"Below are three pictures of different types of crisps. Please select how healthy each type of crisps is compared to the other 2 types of crisps – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

{IF Oct20SampSplit1 = 1 [Crisps1BL image]}
{IF Oct20SampSplit1 = 2 [Crisps1MTL image]}
{IF Oct20SampSplit1 = 3 [Crisps1NS image]}
{IF Oct20SampSplit1 = 4 [Crisps1WL image]}
{IF Oct20SampSplit1 = 5 [Crisps1PCT image]}
EXPCrispsRnk1
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [Crisps2BL image]}
{IF Oct20SampSplit1 = 2 [Crisps2MTL image]}
{IF Oct20SampSplit1 = 3 [Crisps2NS image]}
{IF Oct20SampSplit1 = 4 [Crisps2WL image]}
{IF Oct20SampSplit1 = 5 [Crisps2PCT image]}
EXPCrispsRnk2
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [Crisps3BL image]}
{IF Oct20SampSplit1 = 2 [Crisps3MTL image]}
{IF Oct20SampSplit1 = 3 [Crisps3NS image]}
{IF Oct20SampSplit1 = 4 [Crisps3WL image]}
{IF Oct20SampSplit1 = 5 [Crisps3PCT image]}
EXPCrispsRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF EXPCrispsRnk1 = EXPCrispsRnk2 OR EXPCrispsRnk1 = EXPCrispsRnk3 OR EXPCrispsRnk2 = EXPCrispsRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
EXPCrispsEnInf
"Do you feel you had enough information to rank the crisps from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
EXPYoghRnk [RANDOMISE EXPYoghRnk1… EXPYoghRnk3]
"Below are three pictures of different types of yoghurt. Please select how healthy each yoghurt is compared to the other 2 yoghurts – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

{IF Oct20SampSplit1 = 1 [Yogh1BL image]}
{IF Oct20SampSplit1 = 2 [Yogh1MTL image]}
{IF Oct20SampSplit1 = 3 [Yogh1NS image]}
{IF Oct20SampSplit1 = 4 [Yogh1WL image]}
{IF Oct20SampSplit1 = 5 [Yogh1PCT image]}
EXPYoghRnk1
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [Yogh2BL image]}
{IF Oct20SampSplit1 = 2 [Yogh2MTL image]}
{IF Oct20SampSplit1 = 3 [Yogh2NS image]}
{IF Oct20SampSplit1 = 4 [Yogh2WL image]}
{IF Oct20SampSplit1 = 5 [Yogh2PCT image]}
EXPYoghRnk2
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [Yogh3BL image]}
{IF Oct20SampSplit1 = 2 [Yogh3MTL image]}
{IF Oct20SampSplit1 = 3 [Yogh3NS image]}
{IF Oct20SampSplit1 = 4 [Yogh3WL image]}
{IF Oct20SampSplit1 = 5 [Yogh3PCT image]}
EXPYoghRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF EXPYoghRnk1 = EXPYoghRnk2 OR EXPYoghRnk1 = EXPYoghRnk3 OR EXPYoghRnk2 = EXPYoghRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}
EXPYoghEnInf
"Do you feel you had enough information to rank the yoghurts from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
EXPBrkCerRnk [RANDOMISE EXPBrkCerRnk1… EXPBrkCerRnk3]
"Below are three pictures of different types of breakfast cereal. Please select how healthy each type of breakfast cereal is compared to the other 2 types of breakfast cereal – most healthy, least healthy or in between.

Please select one answer for each image.

LAYOUT: On large screen, present images horizontally with answer boxes below each image. On small screen present images vertically with answer boxes below each image

{IF Oct20SampSplit1 = 1 [BrkCer1BL image]}
{IF Oct20SampSplit1 = 2 [BrkCer1MTL image]}
{IF Oct20SampSplit1 = 3 [BrkCer1NS image]}
{IF Oct20SampSplit1 = 4 [BrkCer1WL image]}
{IF Oct20SampSplit1 = 5 [BrkCer1PCT image]}
EXPBrkCerRnk1
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [BrkCer2BL image]}
{IF Oct20SampSplit1 = 2 [BrkCer2MTL image]}
{IF Oct20SampSplit1 = 3 [BrkCer2NS image]}
{IF Oct20SampSplit1 = 4 [BrkCer2WL image]}
{IF Oct20SampSplit1 = 5 [BrkCer2PCT image]}
EXPBrkCerRnk2
1. Most healthy
2. In between
3. Least healthy

{IF Oct20SampSplit1 = 1 [BrkCer3BL image]}
{IF Oct20SampSplit1 = 2 [BrkCer3MTL image]}
{IF Oct20SampSplit1 = 3 [BrkCer3NS image]}
{IF Oct20SampSplit1 = 4 [BrkCer3WL image]}
{IF Oct20SampSplit1 = 5 [BrkCer3PCT image]}
EXPBrkCerRnk3
1. Most healthy
2. In between
3. Least healthy

HARD CHECK: IF EXPBrkCerRnk1 = EXPBrkCerRnk2 OR EXPBrkCerRnk1 = EXPBrkCerRnk3 OR EXPBrkCerRnk2 = EXPBrkCerRnk3: "You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate."

{ASK ALL}

EXPBrkCerEnInf
"Do you feel you had enough information to rank the breakfast cereals from most healthy to least healthy?"

1. Yes
2. No

{ASK ALL}
ExpRankConf [FLIP SCALE]
"Thinking about the second set of rankings you have just done…

How many do you think you ranked in the correct order – from most health to least healthy?"

1. All of them
2. Most of them
3. Some of them
4. None of them
5. Don't know

4.3 Labels

{IF OctSampSplit1 = 2…5}
SeeLab
"Thinking about the second set of pictures of food we showed you, did you see any labels that looked like the following?"

{IF Oct20SampSplit1 = 2 [MTLLabel image]}
{IF Oct20SampSplit1 = 3 [NSLabel image]}
{IF Oct20SampSplit1 = 4 [WLLabel image]}
{IF Oct20SampSplit1 = 5 [PCTLabel image]}

1. Yes
2. No
3. Not sure

{IF SeeLab = 1}
UseLab [FLIP SCALE]
"And did you use those labels to help you rank how healthy the different foods are?"

{IF Oct20SampSplit1 = 2 [MTLLabel image]}
{IF Oct20SampSplit1 = 3 [NSLabel image]}
{IF Oct20SampSplit1 = 4 [WLLabel image]}
{IF Oct20SampSplit1 = 5 [PCTLabel image]}

1. I used the labels to help me rank all of the foods
2. I used the labels to help me rank some of the foods
3. I did not use the labels to help me rank the foods

{IF SeeLab = 1}
NotLab [FLIP SCALE]
"How easy or difficult did you find it to notice these labels?"

{IF Oct20SampSplit1 = 2 [MTLLabel image]}
{IF Oct20SampSplit1 = 3 [NSLabel image]}
{IF Oct20SampSplit1 = 4 [WLLabel image]}

{IF Oct20SampSplit1 = 5 [PCTLabel image]}

1. Very easy
2. Quite easy
3. Quite difficult
4. Very difficult

{IF SeeLab = 1}
UndLab [FLIP SCALE]
"And how easy or difficult did you find it to understand these labels?"

{IF Oct20SampSplit1 = 2 [MTLLabel image]}
{IF Oct20SampSplit1 = 3 [NSLabel image]}
{IF Oct20SampSplit1 = 4 [WLLabel image]}
{IF Oct20SampSplit1 = 5 [PCTLabel image]}

1. Very easy
2. Quite easy
3. Quite difficult
4. Very difficult

{IF OctSampSplit1 = 2…5}
AttLabUse
"Would you like these labels to be put on all food and drink packaging?"

{IF Oct20SampSplit1 = 2 [MTLLabel image]}
{IF Oct20SampSplit1 = 3 [NSLabel image]}
{IF Oct20SampSplit1 = 4 [WLLabel image]}
{IF Oct20SampSplit1 = 5 [PCTLabel image]}

1. Yes
2. No

{IF OctSampSplit1 = 2…5}
AttLabHelp[FLIP SCALE]
"How helpful do you think you would find these labels in choosing what to buy?"

{IF Oct20SampSplit1 = 2 [MTLLabel image]}
{IF Oct20SampSplit1 = 3 [NSLabel image]}
{IF Oct20SampSplit1 = 4 [WLLabel image]}
{IF Oct20SampSplit1 = 5 [PCTLabel image]}

1. Very helpful
2. Quite helpful
3. Not very helpful
4. Not at all helpful

{IF OctSampSplit1 = 2…5}
AttLabLong [FLIP SCALE]
"Which of the following best represents your views on how long it takes to use these labels?

{IF Oct20SampSplit1 = 2 [MTLLabel image]}
{IF Oct20SampSplit1 = 3 [NSLabel image]}
{IF Oct20SampSplit1 = 4 [WLLabel image]}
{IF Oct20SampSplit1 = 5 [PCTLabel image]}

1. These labels would take too long to use when buying food and drink
2. These labels would be quick enough to use when buying food and drink

4.4 Food habits

{ASK ALL}
FdHabInt
"Our next set of questions are about your current food shopping and eating habits"

DISPLAY

{ASK ALL}
FdShop
"Do you ever do any food shopping, either for yourself or someone else?"

1. Yes - I do some food shopping
2. No - someone else does food shopping for me

{ASK ALL}
RdNut [FLIP SCALE]
"How often do you read the nutritional information on the front of packaging when buying food or drink?"

1. Very often
2. Quite often
3. Occasionally
4. Rarely
5. Never

{IF RdNut <> 5}
RdNutInf [FLIP SCALE]
"And how often does the nutritional information on food packaging influence what you buy?"

1. Very often
2. Quite often
3. Occasionally
4. Rarely
5. Never

{ASK ALL}
HlthFdKnow [FLIP SCALE]
"How much knowledge would you say you have about healthy eating?"

1. A lot of knowledge
2. Some knowledge
3. A little knowledge
4. No knowledge

{ASK ALL}
HlthFdInt [FLIP SCALE]
"How interested would you say you are in healthy eating?"

1. Very interested
2. Quite interested
3. Not very interested

4. Not at all interested

{ASK ALL}
WeightLoss
"Are you currently trying to lose weight?"

1. Yes
2. No
3. Prefer not to say

{ASK ALL}
FdOft [MULTICODE: RANDOMISE ROWS 1…6]
"Thinking about the past 12 months, which of the following types of food or drink have you bought or consumed?

Please select all that apply"

GRID ROWS
1. Pizza
2. Hot chocolate
3. Cake
4. Crisps
5. Yoghurt
6. Breakfast cereal
7. None of these [EXCLUSIVE]

4.5Additional demographics

{ASK ALL}
HeightUnit
"We would like to know how to tall you are.

Would you prefer to tell us your height in metres and centimetres or feet and inches?

1. Metres and centimetres
2. Feet and inches
3. I'd prefer not to give my height

NO REF

{IF HeightUnit = 1}
HeightMet
"What is your height without shoes?

If you are unsure, please give your best estimate"

RANGE 0…2 metres

RANGE 0…99 centimetres

{IF HeightUnit = 2}
HeightImp
"What is your height without shoes?

If you are unsure, please give your best estimate"

RANGE 1…7 feet

RANGE 0…11 inches

{ASK ALL}
WeightUnit
"We would like to know how much you weigh.

Would you prefer to tell us your weight in kilograms or stones and pounds?

1. Kilograms
2. Stones and pounds
3. I'd prefer not to give my weight

{IF WeightUnit = 1}
WeightMet
"How much do you weigh without clothes and shoes?

If you are unsure, please give your best estimate"

RANGE 1…250 kilograms

{IF WeightUnit = 2}
WeightImp
"How much do you weigh without clothes and shoes?

If you are unsure, please give your best estimate"

RANGE 1…40 stones

RANGE 0…13 pounds

{IF FF_Sex = 1 AND FF_Age LT 50}
Preg
"Are you currently pregnant?"

1. Yes
2. No

{ASK ALL}
DisVis
"Do you have any physical or mental health conditions or illnesses lasting or expected to last for 12 months or more which affect your vision (for example blindness, or partial sight)?"

1. Yes
2. No

{ASK ALL}
DisLearn
"Do you have any physical or mental health conditions or illnesses lasting or expected to last for 12 months or more which affect your learning, understanding or concentration?"

1. Yes
2. No

{ASK ALL}
DisEat
"Do you have any physical or mental health conditions or illnesses lasting or expected to last for 12 months or more which affect what you eat?"

1. Yes
2. No

{ASK ALL}
DisEatSpec
"What is/are the physical or mental health condition(s) or illness(es) lasting or expected to last for 12 months or more which affect what you eat?"

OPEN

{ASK ALL}
HlthCond1 [MULTICODE: RANDOMISE ROWS 1…6]
"Has a doctor or other health professional ever told you that you have any of these conditions?"

Please select all that apply.

1. Coronary heart disease
2. Angina
3. Heart attack or myocardial infarction
4. Type 1 Diabetes
5. Type 2 Diabetes
6. High blood pressure/hypertension
7. None of these [EXCLUSIVE]

{ASK ALL}
EngLang
"Is English your first language?"

1. Yes
2. No

Nutri-Score Bandings

Aliments solides (points)	Boissons (points)	Couleur
Min à -1	Eau	Vert foncé
0 à 2	Min à 1	Vert clair
3 à 10	2 à 5	Jaune
11 à 18	6 à 9	Orange clair
19 à Max	10 à Max	Orange foncé

See Nutri-Score calculation tool Excel spreadsheet provided by Sante publique France https://www.santepubliquefrance.fr/media/files/02-determinants-de-sante/nutrition-et-activite-physique/nutri-score/tableur-calcul-nutri-score-en (accessed 14/02/23)

Multiple Traffic Light Criteria for 100g Food

Text	LOW[8]	MEDIUM	HIGH	
Colour code	Green	Amber	Red	
			>25% of RIs	>30% of RIs
Fat	≤ 3.0g/100g	> 3.0g to ≤ 17.5g/100g	> 17.5g/100g	> 21g/portion
Saturates	≤ 1.5g/100g	> 1.5g to ≤ 5.0g/100g	> 5.0g/100g	> 6.0g/portion
(Total) Sugars	≤ 5.0g/100g	> 5.0g to ≤ 22.5g /100g	> 22.5g/100g	> 27g/portion
Salt	≤ 0.3g/100g	> 0.3g to ≤ 1.5g/100g	>1.5g/100g	>1.8g/portion

Note: Portion size criteria apply to portions/servings greater than 100g. The NPM test dataset includes product data per 100g only and therefore the 'per portion' cut-offs were excluded from this analysis.

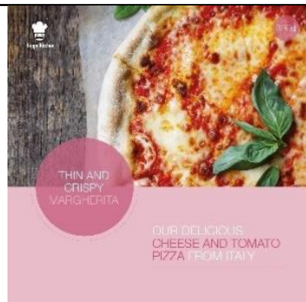
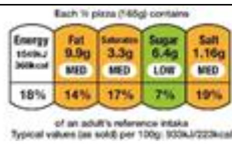
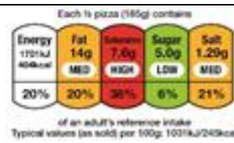
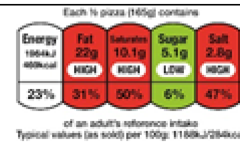
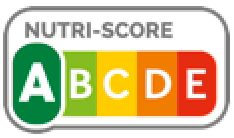
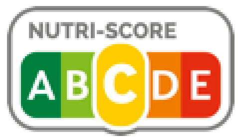
Multiple Traffic Light Criteria for 100ml Drink

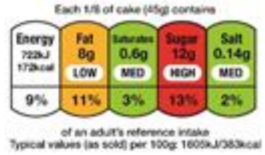
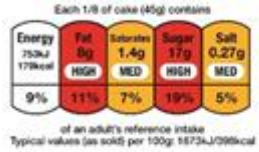
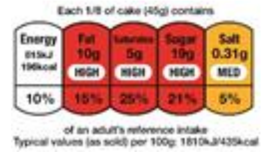
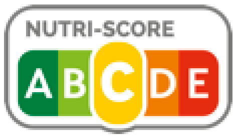
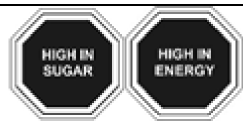
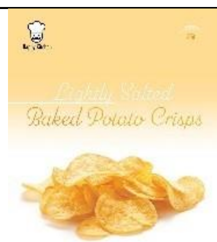
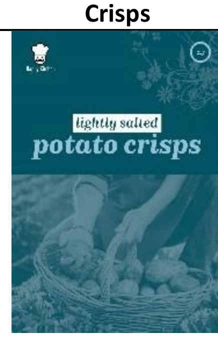
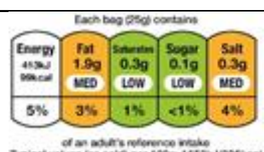
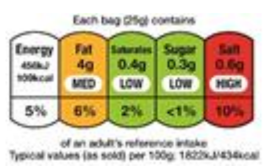
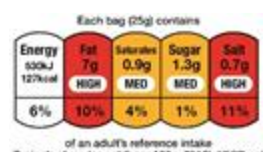
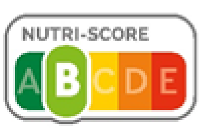
Text	LOW[9]	MEDIUM	HIGH	
Colour code	Green	Amber	Red	
			>12.5% of RIs	>15% of RIs
Fat	≤ 1.5g/100ml	> 1.5g to ≤ 8.75g/100ml	> 8.75g/100ml	>10.5g/portion
Saturates	≤ 0.75g/100ml	> 0.75g to ≤ 2.5g/100ml	> 2.5g/100ml	> 3g/portion
(Total) Sugars	≤ 2.5g/100ml	> 2.5g to ≤ 11.25g/100ml	> 11.25g/100ml	> 13.5g/portion
Salt	≤ 0.3g/100ml	>0.3g to ≤0.75g/100ml	> 0.75g/100ml	> 0.9g/portion

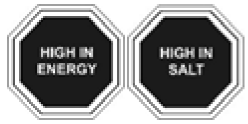
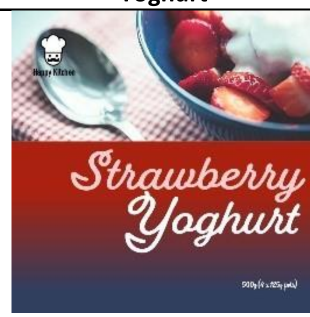
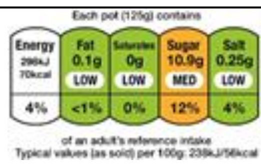
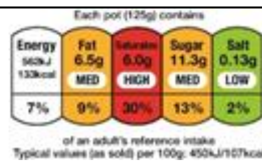
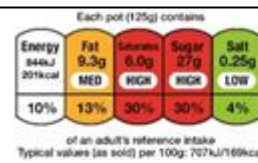
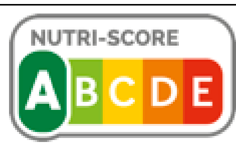
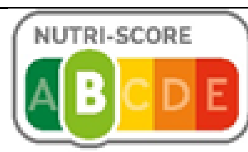
Table C1: Nutritional specification of products (including information necessary for Multiple Traffic Light labels)

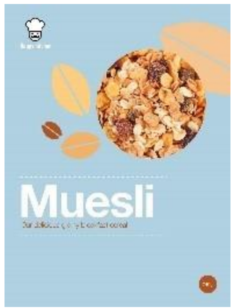
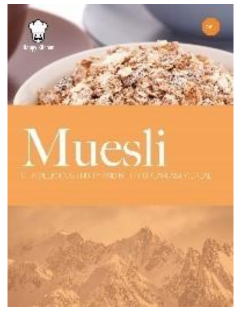
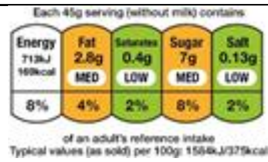
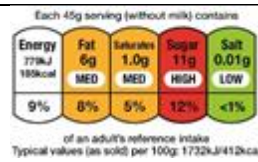
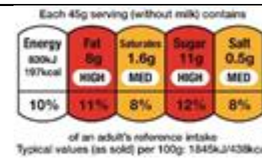
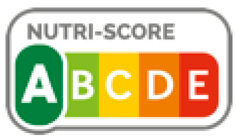
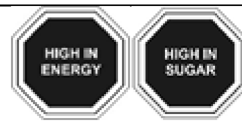
		Per package						**Per 100g**	
	Portion size	**kcal**	**kJ %RI**	**Fat %RI**	**Sat fat %RI**	**Sugar %RI**	**Salt %RI**	**kJ**	**kcal**
Pizza 1	165	368	1549 18	9.9 14	3.3 17	6.4 7	1.16 19	933	223
2		404	1701 20	14 20	7.6 38	5.0 6	1.29 21	1031	245
3		469	1964 23	22 31	10.1 50	5.1 6	2.8 47	1188	284
Cake 1	45	172	722 9	8 11	0.6 3	12 13	0.14 2	1605	383
2		179	753 9	8 11	1.4 7	17 19	0.27 5	1673	398
3		196	815 10	10 15	5 25	19 21	0.31 5	1810	435
Crisps 1	25	99	413 5	1.9 3	0.3 1	0.1 0	0.3 4	1653	395
2		109	456 5	4 6	0.4 2	0.3 <1%	0.6 10	1822	434
3		127	530 6	7 10	0.9 4	1.3 1	0.7 11	2119	507
Yoghurt 1	125	70	298 4	0.1 <1%	0 0	10.9 12	0.25 4	238	56
2		133	562 7	6.5 9	6.0 30	11.3 13	0.13 2	450	107
3		201	844 10	9.3 13	6.0 30	27 30	0.25 4	707	169
Cereal 1	45	169	713 8	2.8 4	0.4 2	7 8	0.13 2	1584	375
2		185	779 9	6 8	1.0 5	11 12	0.01 0	1732	412
3		197	830 10	8 11	1.6 8	11 12	0.5 8	1845	438
Hot Choc 1	230ml (30g powder + 200 ml water)	101	422 5	1.8 3	1.1 5	9.9 11	0.13 2	184	44
2		145	614 7	2.8 4	2.3 12	22.3 25	0.41 7	267	63
3		125	526 6	3.7 5	3.0 15	16.6 18	0.55 9	229	54

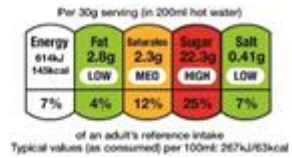
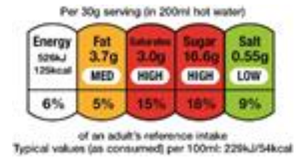
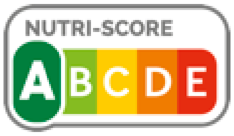
Table C2: Product images, applicable labels and design features

FOPL group	**Product**		
	Pizza		
No label (baseline and control)			
Multiple Traffic Light			
Nutri-Score			
Warning label			
Positive Choice tick			
Design features All: size=330g (standard pizza box)	<u>Description</u>: 'Thin and crispy margherita' and 'Our delicious cheese and tomato pizza from Italy' <u>Image</u>: Pizza with cheese, tomato sauce and fresh basil <u>Colour</u>: Light pink	<u>Description</u>: 'Margherita' and 'Our delicious cheese and tomato pizza from Italy' <u>Image</u>: Pizza with cheese (more than 1), tomato sauce and fresh basil (less than 1) <u>Colour</u>: Green	<u>Description</u>: 'Four cheese margherita' and 'Mouth-watering pizza from Italy' <u>Image</u>: Pizza with cheese (more than 2 and greasy) and very little tomato sauce <u>Colour</u>: Yellow
	Cake		
No label (baseline and control)			

Multiple Traffic Light			
Nutri-Score			
Warning label			
Positive Choice tick	N/A	N/A	N/A
Design features All: size=360g (whole cake)	<u>Description</u>: 'Fruit cake' and 'Our delicious fruit cake' <u>Image</u>: Plain loaf cake with fruit visible- no icing <u>Colour</u>: Beige	<u>Description</u>: 'Iced lemon cake' and 'Our delicious lemon cake' <u>Image</u>: Lemon cake loaf covered in icing and lemon zest <u>Colour</u>: Yellow	<u>Description</u>: 'Chocolate cake' and 'Our delicious chocolate cake' <u>Image</u>: Chocolate cake with chocolate icing and chunks <u>Colour</u>: Brown
		Crisps	
No label (baseline and control)			
Multiple Traffic Light			
Nutri-Score			

Warning label			
Positive Choice tick	N/A	N/A	N/A
Design features All: size=25g (single packet)	<u>Description</u>: 'Lightly salted baked potato crisps' – italicised font <u>Image</u>: Pile of crisps <u>Colour</u>: Light yellow and white	<u>Description</u>: 'Lightly salted potato crisps'- standard font <u>Image</u>: Potatoes in a basket <u>Colour</u>: Aqua – bluey green	<u>Description</u>: 'Sweet chilli potato crisps'- bold font <u>Image</u>: Large crisps <u>Colour</u>: Bright red
	Yoghurt		
No label (baseline and control)			
Multiple Traffic Light			
Nutri-Score			
Warning label			
Positive Choice tick			
Design features All: size=500g (4x125g pots)	<u>Description</u>: 'Strawberry yoghurt' <u>Image</u>: Yoghurt with some strawberry mixed in a yoghurt pot and strawberries in the background <u>Colour</u>: White	<u>Description</u>: 'Strawberry yoghurt' <u>Image</u>: Yoghurt with strawberries in a bowl <u>Colour</u>: Dull red	<u>Description</u>: 'Greek yoghurt with a thick strawberry layer' <u>Image</u>: Yoghurt with strawberries on top and strawberry compote layer in glasses <u>Colour</u>: Bright red
	Breakfast cereal		

No label (baseline and control)			
Multiple Traffic Light			
Nutri-Score			
Warning label			
Positive Choice tick			
Design features All: size=740g	Description: 'Muesli' and 'Our delicious grainy breakfast cereal' Image: Muesli with plenty of wholegrain flakes and animated oats/nuts Colour: Light blue	Description: 'Muesli' and 'Our delicious fruity and nutty breakfast cereal' Image: Muesli in a bowl with more fruit and what could be sugar frosting and mountains to indicate Swiss style muesli Colour: Dull orange	Description: 'Granola' and 'Our delicious fruity and nutty breakfast cereal' Image: Granola in a bowl Colour: Dark brown
	Instant hot chocolate powder (drink)		
No label (baseline and control)			

Multiple Traffic Light			
Nutri-Score			
Warning label			
Positive Choice tick			
Design features All: size = 320g	<u>Description</u>: 'Instant Hot Chocolate' – thin font <u>Image</u>: Plain hot chocolate in a mug <u>Colour</u>: Light purple	<u>Description</u>: 'Instant Hot Chocolate' – thicker font <u>Image</u>: Hot chocolate in a mug with a swirl of chocolate <u>Colour</u>: Dark purple	<u>Description </u>: 'Instant Hot Chocolate'- cursive font <u>Image</u>: Dark chocolate swirls <u>Colour</u>: Dark brown

Product Packaging Example for 'Least Healthy' Pizza with MTL, N-S and WL FOPLs

MTL	
N-S	
WL	

Full Survey

1.1 Baseline ranking

Our first set of questions will ask you to rank different foods on how healthy you think they are. We would like to know how long it takes for you to rank the different foods, so please click 'next' as soon as you have completed each task.

Please rank them from most healthy to least healthy

Baseline ranking task

Below are three pictures of different types of pizza. Please select how healthy each pizza is compared to the other 2 pizzas – most healthy, least healthy or in between.

Please select one answer for each image.

[Images of the 3 pizzas presented here with answer boxes below each image: 1 – Most healthy, 2 – In between, 3 – Least healthy]

- If participants ranked two products the same this message appeared: You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate.

Enough information question

Do you feel you had enough information to rank the pizzas from most healthy to least healthy?"

1. Yes
2. No

[Ranking task and enough information question repeated for the 5 other product categories]

Confidence question

Thinking about **all of** the rankings you have just done…

How many do you think you ranked in the correct order – from most healthy to least healthy?

1. All of them
2. Most of them
3. Some of them
4. None of them

1.2 Experimental ranking

We will now present you with another set of pictures of the same food items. Remember, we would like to know how long it takes for you to rank the different foods, so please click 'next' as soon as you have completed each task.

Please rank them from most healthy to least healthy.

Experimental ranking task

Below are three pictures of different types of pizza. Please select how healthy each pizza is compared to the other 2 pizzas – most healthy, least healthy or in between. Please select one answer for each image. **[Images of the 3 pizzas with allocated FOPL presented here with answer boxes below each image: 1 – Most healthy, 2 – In between, 3 – Least healthy]** - If participants had two products ranked the same this message appeared: You have ranked two foods the same. If you are not sure what order to put the foods in, please give your best estimate.
Enough information question Do you feel you had enough information to rank the pizzas from most healthy to least healthy?" 1. Yes 2. No **[Ranking task and enough information question repeated for the 5 other product categories]**
Confidence question Thinking about **the second set** of rankings you have just done… How many do you think you ranked in the correct order – from most healthy to least healthy? 1. All of them 2. Most of them 3. Some of them 4. None of them 5. Don't know
POST-RANKING QUESTIONS **1.3 Labels** (excluding control) Thinking about the second set of pictures of food we showed you, did you see any labels that looked like the following? **[shown image of their allocated FOPL condition]** 1. Yes 2. No 3. Not sure
And did you those labels to help you rank how healthy the different foods and drinks are? 1. I used the labels to help me rank **all** of the foods 2. I used the labels to help me rank **some** of the foods 3. I did not use the labels to help me rank the foods
And how easy or difficult did you find it to **understand** these labels? 1. Very easy 2. Quite easy

3. Quite difficult 4. Very difficult
Would you like these labels to be put on food and drink packaging in the UK? 1. Yes – all 2. Yes – some 3. No – none
How helpful do you think you would find these labels in choosing what to buy? 1. Very helpful 2. Quite helpful 3. Not very helpful 4. Not at all helpful
Which of the following best represents your views on how long it takes to use these labels? 1. These labels would take too long to use when buying food and drink 2. These labels would be quick enough to use when buying food and drink
1.4 Food habits Our next set of questions are about your current food shopping and eating habits Do you ever do any food shopping, either for yourself or someone else? 1. Yes – I do some food shopping 2. No – someone else does food shopping for me
How often do you read the nutritional information on the front of packaging when buying food or drink? 1. Very often 2. Quite often 3. Occasionally 4. Rarely 5. Never
And how often does the nutritional information on food packaging influence what you buy? 1. Very often 2. Quite often 3. Occasionally 4. Rarely 5. Never
How much **knowledge** would you say you have about healthy eating? 1. A lot of knowledge 2. Some knowledge 3. A little knowledge

4. No knowledge

How **interested** would you say you are in healthy eating?

1. Very interested
2. Quite interested
3. Not very interested
4. Not at all interested

Are you currently trying to lose weight?

1. Yes
2. No
3. Prefer not to say

Thinking about the past 12 months, which of the following types of food or drink have you bought or consumed?

Please select all that apply

1. Pizza
2. Hot chocolate
3. Cake
4. Crisps
5. Yoghurt
6. Breakfast cereal
7. None of these

1.5 Background questions

Our final set of questions is to help us with our analysis.

We would like to know how to tall you are.

Would you prefer to tell us your height in metres and centimetres or feet and inches?

1. Metres and centimetres
2. Feet and inches
3. I'd prefer not to give my height

What is your height without shoes?

If you are unsure, please give your best estimate

[Textbox response] (range 0…2 metres, 0…99 centimetres OR 0…40 stones, 0…13 pounds)

We would like to know how much you weigh.

Would you prefer to tell us your weight in kilograms or stones and pounds?

1. Kilograms
2. Stones and pounds
3. I'd prefer not to give my weight

How much do you weigh without clothes and shoes? If you are unsure, please give your best estimate **[Textbox response]** (range 0…250 kilograms OR 0…7 feet, 0…11 inches)
Are you currently pregnant? 1. Yes 2. No 3. Prefer not to say
Do you have any physical or mental health conditions or illnesses lasting or expected to last for 12 months or more which affect **your vision** (for example blindness, or partial sight)? 1. Yes 2. No
Do you have any physical or mental health conditions or illnesses lasting or expected to last for 12 months or more which affect **your learning, understanding or concentration**? 1. Yes 2. No
Do you have any physical or mental health conditions or illnesses lasting or expected to last for 12 months or more which affect **what you eat**? 1. Yes 2. No
What is/are the physical or mental health condition(s) or illness(es) lasting or expected to last for 12 months or more which affect what you eat? **[Textbox response]**
Has a doctor or other health professional ever told you that you have any of these conditions? Please select all that apply. 1. Coronary heart disease 2. Angina 3. Heart attack or myocardial infarction 4. Type 1 Diabetes 5. Type 2 Diabetes 6. High blood pressure/hypertension 7. None of these

Ethics Decision Letter

NatCen
Social Research that works for society

Application: P15640 Front of package labelling

Email: recadmin@natcen.ac.uk
Direct Line: 0207 549 7068

04/09/2020

Dear Curtis

RE: REC Application

Thank you for submitting your application for ethical review.

I am writing to let you know the decision of the REC with regards to your project.

Your study has been approved subject to the following:

- You will make sure that the person(s) responsible for checking the datasets are aware of safeguarding procedures and how to escalate any incidents.

Do not hesitate to contact myself or Nilufer Rahim should you have any questions.

Yours sincerely,

Kirsty Herrington
NatCen REC Administrator

NatCen Social Research is the trading name of the National Centre for Social Research.

Registered Office
35 Northampton Square
London EC1V 0AX
www.natcen.ac.uk

T. 020 7250 1866
F. 020 7250 1524
E. info@natcen.ac.uk
Follow us: @NatCen

A Company Limited by Guarantee
Registered in England No.4392418.
A Charity registered in England and Wales (1091768) and Scotland (SC038454)

UCL Data Protection Approval

From: "Crouch, Spenser" <s.crouch@ucl.ac.uk> on behalf of "data-protection@ucl.ac.uk" <data-protection@ucl.ac.uk>
Date: Wednesday, 9 September 2020 at 17:30
To: "Croker, Helen" <h.croker@ucl.ac.uk>
Subject: 20200909 Email confirm Z6364106 2020 09 31

Hi,

Thank you for your application to register with the Data Protection Office. The attached application form, indicates that you will be a data processor this this study.

Processors, like controllers, are required to implement appropriate security measures. What is appropriate is assessed in terms of a variety of factors including the sensitivity of the data, the risks to individuals associated with any security breach, and the nature of the processing. These measures might include pseudonymisation and encryption. Regular testing of the effectiveness of any security measures is also required where appropriate.

I am pleased to confirm that this project is now registered under, reference No Z6364106/2020/09/31 social research in line with UCL's Data Protection Policy.

You may quote this reference on your Ethics Application Form, or any other related forms.

When all essential documents are ready to archive, contact the UCL Records Office by email records.office@ucl.ac.uk to arrange ongoing secure storage of your research records unless you have made specific alternative arrangements with your department, or funder. Please note the UCL Records Office does not store student research data.

For data protection enquiries, please contact the data protection team at data-protection@ucl.ac.uk

For ethics enquiries, please contact the ethics team at ethics@ucl.ac.uk.

Regards,

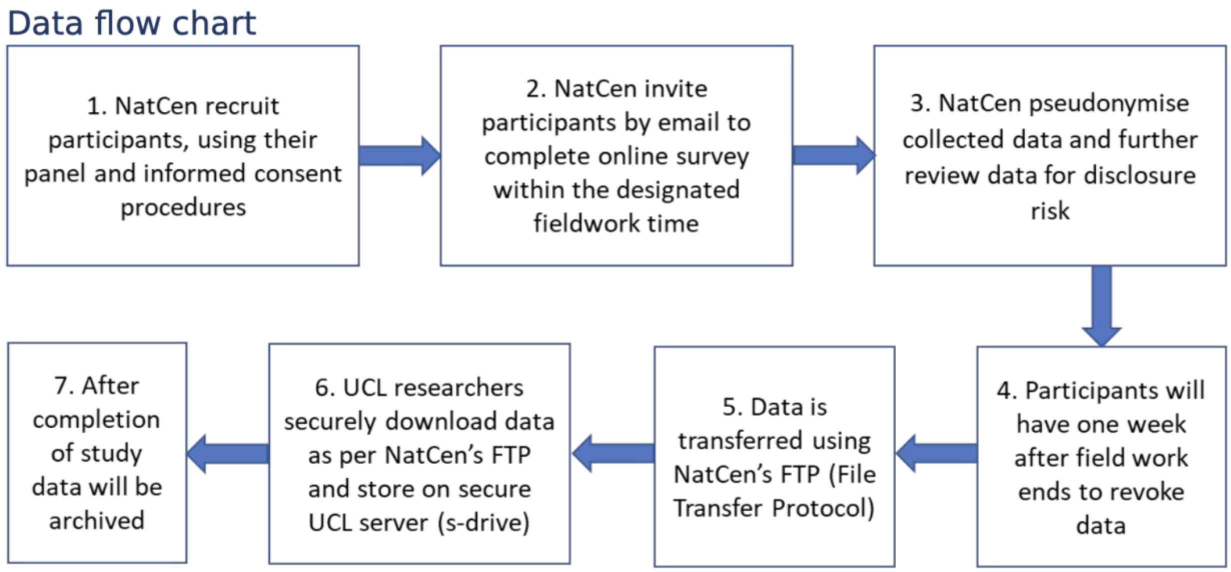
Spenser Crouch
Data Protection & Freedom of Information Administrator & Chief Web Editor
Legal Services, UCL | Gower Street | London | WC1E 6BT
Internal Address: 6th floor | Bidborough House | 38- 50 Bidborough Street | Kings Cross | London | WC1H 9BT
Email: s.crouch@ucl.ac.uk Data Protection: data-protection@ucl.ac.uk FOI: foi@ucl.ac.uk.
Telephone: 0203 108 8764 (internal 58764)

Please note my office working days are Tuesday and Thursday 7.30am – 3.30pm;
Home working days are Monday, Wednesday and Friday.

Please protect the Environment. Print only if necessary.

UCL

Data Flowchart



Appendix D: Primary Outcome Analysis (Chapter 5)

Figure D1: Proportion of participants who reported eating/buying each product category

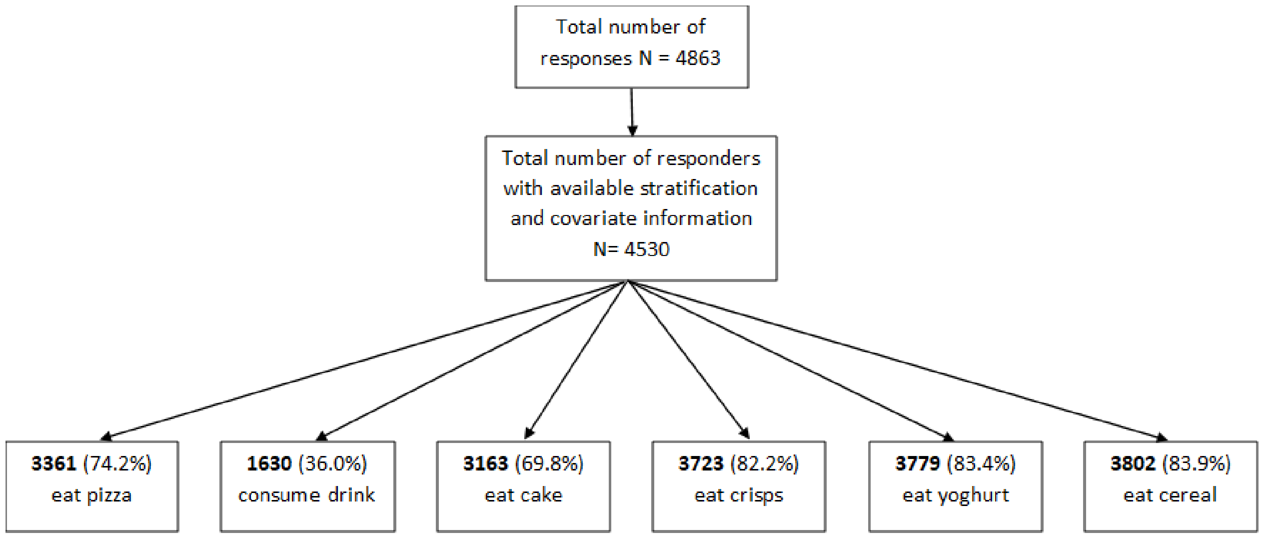


Table D1: Proportion of missing data per variable (*n* = 4,863)

Participant characteristics	***n* (%)**
Sex	0 (0)
Age	6 (0.1)
Ethnicity	58 (1.6)
Education	16 (0.3)
Children in household	21 (0.4)
Shopping responsibility	1 (0.0)
Current label use	0 (0.0)
Reported consuming or buying product in past 12 months*	0 (0.0)
Baseline ranking	
Pizza	55 (1.1)
Instant hot chocolate	98 (2.0)
Cake	53 (1.1)
Crisps	47 (1.0)
Yoghurt	71 (1.5)
Breakfast cereal	56 (1.2)
Follow-up ranking	
Pizza	48 (1.0)
Instant hot chocolate	55 (1.1)
Cake	40 (0.8)
Crisps	38 (0.8)
Yoghurt	49 (1.0)
Breakfast cereal	51 (1.0)
Enough information question	
Baseline	20 (0.4)
Follow-up	27 (0.6)
Reported seeing the label	2 (0.0)

Missing data includes the following responses: refused, don't know, not applicable.

Table D2: Mean global food score by experimental group and participant characteristics – without requirement of buying/eating all food products

	Experimental group					
	Control	**MTL**	**N-S**	**WL**	**PC**	**Overall**
Sex	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)
Female (*n*=2,586)	0.0 (0.8)	1.8 (1.3)	2.0 (1.3)	1.3 (1.4)	0.1 (0.9)	1.1 (1.4)
Male (*n*=1,944)	0.0 (0.8)	2.0 (1.4)	2.1 (1.4)	1.6 (1.5)	0.1 (0.8)	1.2 (1.5)
Age						
18-29 (*n*=279)	-0.1 (0.7)	1.9 (1.2)	2.3 (1.0)	1.5 (1.2)	0.6 (0.8)	1.3 (1.4)
30-39 (*n*=642)	0.0 (0.7)	1.8 (1.3)	2.3 (1.3)	1.5 (1.3)	0.1 (1.0)	1.1 (1.5)
40-49 (*n*=881)	0.0 (0.7)	1.9 (1.2)	2.0 (1.2)	1.3 (1.4)	0.2 (0.9)	1.1 (1.4)
50-59 (*n*=969)	0.0 (0.7)	1.8 (1.4)	2.1 (1.4)	1.3 (1.5)	0.1 (0.8)	1.1 (1.5)
60-69 (*n*=933)	0.0 (0.8)	2.0 (1.4)	1.9 (1.5)	1.5 (1.5)	0.1 (0.9)	1.1 (1.5)
70+ (*n*=823)	0.0 (1.0)	2.1 (1.5)	2.0 (1.5)	1.4 (1.5)	0.1 (0.8)	1.1 (1.6)

Ethnicity						
White (*n*=4,233)	0.0 (0.8)	2.0 (1.3)	2.1 (1.3)	1.4 (1.4)	0.1 (0.9)	1.1 (1.5)
Non-white (*n*=297)	0.0 (0.7)	1.6 (1.7)	2.0 (1.5)	1.1 (1.4)	0.0 (1.1)	1.0 (1.6)
Education level						
Degree or equivalent + (*n*=2,196)	0.0 (0.7)	2.0 (1.3)	2.1 (1.3)	1.5 (1.4)	0.2 (0.9)	1.2 (1.5)
A levels or vocational level 3 or equivalent (*n*=878)	0.1 (0.7)	2.1 (1.2)	2.0 (1.4)	1.4 (1.4)	0.2 (0.9)	1.2 (1.4)
Other qualifications below A levels or equivalent (*n*=784)	-0.1 (0.7)	1.9 (1.4)	2.0 (1.4)	1.5 (1.5)	0.0 (0.9)	1.1 (1.5)
Other qualification (*n*=271)	0.1 (1.2)	1.7 (1.3)	2.3 (1.4)	1.2 (1.5)	0.0 (0.8)	1.0 (1.5)
No qualification (*n*=401)	-0.1 (1.0)	1.4 (1.7)	1.7 (1.6)	1.0 (1.4)	0.1 (0.8)	0.8 (1.5)
Education dichotomised						
A Levels + (*n*=3,074)	0.0 (0.7)	2.0 (1.3)	2.1 (1.3)	1.5 (1.4)	0.2 (0.9)	1.2 (1.4)
Below A level (*n*=1,456)	0.0 (0.9)	1.7 (1.5)	1.9 (1.5)	1.3 (1.5)	0.0 (0.9)	1.0 (1.5)
Equivalised household income (£)						
More than 2000 (*n*=1,702)	0.0 (0.7)	2.1 (1.3)	2.1 (1.3)	1.5 (1.4)	0.2 (0.9)	1.2 (1.4)
1251- 2000 (*n*=1,148)	0.0 (0.8)	2.2 (1.2)	2.1 (1.4)	1.5 (1.4)	0.0 (0.8)	1.2 (1.5)
801-1250 (*n*=887)	-0.1 (0.9)	1.8 (1.3)	1.8 (1.4)	1.4 (1.4)	0.1 (0.9)	1.0 (1.5)
800 or less (*n*=631)	0.1 (0.7)	1.3 (1.5)	2.1 (1.5)	1.0 (1.5)	0.1 (0.9)	0.9 (1.5)
Income dichotomised						
More than 2000 (*n*=1,702)	0.0 (0.7)	2.1 (1.3)	2.1 (1.3)	1.6 (1.4)	0.2 (0.9)	1.2 (1.4)
2000 and below (*n*=2,666)	0.0 (0.8)	1.9 (1.4)	2.0 (1.4)	1.4 (1.5)	0.1 (0.9)	1.1 (1.5)
Current label use						
Often (*n*=2,491)	0.0 (0.7)	2.0 (1.3)	2.1 (1.2)	1.5 (1.4)	0.1 (0.9)	1.2 (1.4)
Not often (*n*=2,039)	0.0 (0.8)	1.9 (1.5)	2.0 (1.5)	1.3 (1.5)	0.1 (0.9)	1.0 (1.5)
Interest in healthy eating dichotomised						
Very (*n*=1,894)	0.0 (0.7)	1.9 (1.3)	2.1 (1.3)	1.4 (1.4)	0.2 (0.9)	1.2 (1.4)
Not very / quite and below (*n*=2,636)	0.0 (0.8)	1.9 (1.4)	2.0 (1.4)	1.4 (1.4)	0.1 (0.9)	1.1 (1.5)
Currently trying to lose weight* 165						
Yes (*n*=2,125)	0.0 (0.8)	1.9 (1.3)	2.1 (1.3)	1.4 (1.5)	0.2 (0.9)	1.1 (1.5)
No (*n*=2,240)	0.0 (0.7)	1.9 (1.4)	2.0 (1.4)	1.5 (1.4)	0.1 (1.9)	1.1 (1.5)

Children in household						
Yes (*n*=1,377)	0.0 (0.7)	1.9 (1.4)	2.1 (1.2)	1.3 (1.4)	0.2 (0.9)	1.1 (1.4)
No (*n*=3,153)	0.0 (0.8)	2.0 (1.4)	2.0 (1.4)	1.4 (1.4)	0.1 (0.8)	1.1 (1.5)
Medical conditions that affect diet						
Yes (*n*=740)	0.0 (0.9)	2.0 (1.4)	2.0 (1.3)	1.5 (1.4)	0.1 (0.9)	1.1 (1.5)
No (*n*=3,789)	0.0 (0.7)	1.9 (1.3)	2.1 (1.4)	1.4 (1.4)	0.1 (0.9)	1.1 (1.5)
Type 1 diabetes						
Yes (*n*=39)	0.0 (1.4)	2.4 (1.2)	2.2 (0.9)	1.4 (0.9)	-0.3 (1.0)	1.5 (1.4)
No (*n*=4,491)	0.0 (0.8)	1.9 (1.4)	2.1 (1.4)	1.4 (1.4)	0.1 (0.9)	1.1 (1.5)
Type 2 diabetes						
Yes (*n*=310)	0.1 (1.0)	2.0 (1.5)	1.9 (1.6)	1.3 (1.5)	0.1 (0.9)	1.0 (1.6)
No (*n*=4,220)	0.0 (0.7)	1.9 (1.3)	2.1 (1.3)	1.4 (1.4)	0.1 (0.9)	1.1 (1.5)
BMI category						
Underweight (*n*=21)	0.0 (0.4)	1.3 (1.8)	0.8 (2.4)	1.2 (1.8)	-0.1 (1.1)	0.7 (1.6)
Normal (*n*=757)	0.0 (0.7)	2.1 (1.3)	2.0 (1.3)	1.5 (1.3)	0.1 (0.9)	1.1 (1.5)
Overweight (*n*=572)	0.0 (0.8)	1.8 (1.3)	2.2 (1.4)	1.5 (1.5)	0.1 (0.8)	1.1 (1.5)
Obese (*n*=386)	0.0 (0.7)	1.9 (1.5)	2.1 (1.3)	1.2 (1.4)	0.1 (0.9)	1.1 (1.5)

Variables were dichotomised as: sex, female vs male; ethnicity, white (White British/White other) vs non-white (Mixed or multiple ethnic groups/ Asian or Asian British/ Black or Black British/ Other); current label use, often (very often/ quite often) vs not often (occasionally/ rarely/ never); interest in healthy eating, very (very interested) vs not very (quite interested/ not very interested/ not at all interested); currently trying to lose weight, yes vs no; children in household, yes vs no; medical conditions that affect diet, yes vs no; type 1 diabetes, yes vs no; type 2 diabetes, yes vs no.

Table D3: Individual characteristics of the analysis sample, by experimental group

	Experimental group					
	Control (*n* = 913)	**MTL (*n* = 907)**	**N-S (*n* = 924)**	**WL (*n* = 895)**	**PC (*n* = 891)**	**Overall (*n* = 4,530)**
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**
Sex						
Female	518 (57)	512 (57)	534 (58)	510 (57)	512 (58)	2,586 (57)
Male	395 (43)	395 (44)	390 (42)	385 (43)	379 (43)	1,944 (43)
Age						
18-29	61 (7)	65 (7)	56 (6)	56 (6)	41 (5)	279 (6)
30-39	139 (15)	101 (12)	141 (15)	132 (15)	129 (15)	642 (14)
40-49	169 (19)	181 (20)	171 (19)	180 (20)	180 (20)	881 (20)
50-59	199 (22)	204 (23)	187 (20)	185 (21)	194 (22)	969 (21)
60-69	179 (20)	191 (21)	197 (21)	181 (20)	185 (21)	933 (21)
70+	166 (18)	164 (18)	172 (19)	160 (18)	161 (18)	823 (18)
Ethnicity						
White British	799 (88)	789 (87)	809 (88)	792 (89)	785 (88)	3,974 (88)
White other	53 (6)	55 (6)	51 (6)	50 (6)	50 (6)	259 (6)
Mixed or multiple ethnic groups	9 (1)	13 (1)	15 (2)	6 (1)	11 (1)	54 (1)
Asian or Asian British	37 (4)	33 (4)	28 (3)	32 (4)	25 (3)	155 (3)
Black or Black British	12 (1)	14 (2)	17 (2)	13 (2)	16 (2)	72 (2)
Other	3 (0)	3 (0)	4 (0)	2 (0)	4 (1)	16 (0)
Education						
Degree or equivalent +	460 (50)	435 (48)	450 (49)	426 (48)	425 (48)	2,196 (49)
A levels or vocational level 3 or equivalent	174 (19)	174 (19)	179 (19)	174 (19)	177 (20)	878 (20)
Other qualifications below A levels or equivalent	163 (18)	155 (17)	167 (18)	158 (18)	141 (16)	784 (17)
Other qualification	46 (5)	70 (8)	44 (5)	49 (6)	62 (7)	271 (6)
No qualifications	70 (8)	73 (8)	84 (9)	88 (10)	86 (10)	401 (9)
Children in household						
Yes	280 (31)	262 (29)	269 (29)	286 (32)	280 (31)	1,377 (30)
No	633 (69)	645 (71)	655 (71)	609 (68)	611 (69)	3,153 (70)
Shopping responsibility						
Yes – some or all	868 (95)	866 (96)	895 (97)	859 (96)	852 (96)	4,340 (96)
No – someone else does	45 (5)	41 (5)	29 (3)	36 (4)	39 (4)	190 (4)

Current label use						
Very often	177 (19)	209 (23)	179 (19)	220 (25)	160 (18)	945 (21)
Quite often	297 (33)	345 (38)	305 (33)	308 (34)	291 (33)	1,546 (34)
Occasionally	269 (30)	234 (26)	267 (29)	266 (30)	282 (32)	1,318 (29)
Rarely	132 (15)	95 (11)	142 (15)	83 (9)	107 (12)	559 (12)
Never	38 (4)	24 (3)	31 (3)	18 (2)	51 (6)	162 (4)
Reported consuming or buying product in past 12 months						
Pizza	670 (73)	647 (71)	686 (74)	690 (77)	668 (75)	3,361 (74)
Instant hot chocolate	313 (34)	336 (37)	324 (35)	336 (38)	321 (36)	1,630 (36)
Cake	627 (69)	651 (72)	633 (69)	635 (71)	617 (69)	3,163 (70)
Crisps	752 (82)	739 (82)	772 (84)	753 (84)	707 (79)	3,723 (82)
Yoghurt	761 (83)	755 (83)	772 (84)	747 (84)	744 (84)	3,779 (83)
Cereal	779 (85)	757 (84)	755 (82)	756 (85)	755 (85)	3,802 (84)

MTL, Multiple Traffic Light; N-S, Nutri-Score; WL, Warning Label; PC, Positive Choice tick; RR: Relative Risk; CI: Confidence Interval.

Table D4: Summary of participants who correctly ranked products at baseline and follow-up, by FOPL group and product category

	Control *n* (%)		**MTL** *n* (%)		**NS** *n* (%)		**WL** *n* (%)		**PC** *n* (%)		**Overall** *n* (%)	
	B	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**
Pizza	465 (69)	456 (68)	473 (73)	582 (90)	508 (74)	640 (93)	476 (69)	563 (82)	490 (73)	491 (74)	2412 (72)	2732 (81)
Instant hot chocolate	198 (63)	196 (63)	212 (63)	286 (85)	183 (57)	306 (94)	221 (66)	280 (83)	194 (60)	221 (69)	1008 (62)	1289 (79)
Cake	422 (67)	417 (67)	428 (66)	593 (91)	445 (70)	590 (93)	399 (63)	547 (86)	422 (68)	411 (67)	2116 (67)	2558 (81)
Snack	453 (60)	459 (61)	468 (63)	675 (91)	478 (62)	723 (94)	504 (67)	629 (84)	454 (64)	464 (66)	2357 (63)	2950 (79)
Yoghurt	98 (13)	101 (13)	107 (14)	642 (85)	86 (11)	660 (86)	90 (12)	458 (61)	66 (9)	117 (16)	447 (12)	1978 (52)
Cereal	261 (34)	267 (34)	252 (33)	658 (87)	256 (34)	664 (88)	238 (32)	546 (72)	261 (35)	299 (40)	1268 (33)	2434 (64)

FOPL: front of pack label; B = baseline ranking; FU = follow-up ranking; MTL, Multiple Traffic Light; N-S, Nutri-Score; WL, Warning Label; PC, Positive Choice tick.

Table D5: Multiple regression analysis results—association between global food score and N-S vs MTL with education interaction (higher/lower), equivalised income interaction (higher/ lower), interest in healthy eating interaction (very interested/ not very)

	RR (95%CI)	***p-value***
Higher education + Nutri-Score interaction	-0.348 (-0.720, 0.023)	*0.066*
Higher equivalised income + Nutri-Score interaction	-0.261 (-0.619, 0.098)	*0.154*
Very interested in healthy eating + Nutri-Score interaction	0.093 (-0.267, 0.452)	*0.613*
	Across all RR (95%CI)	
Very interested vs not very interested	0.207 (0.056, 0.358)	*0.007*

Education level dichotomised as A-levels and above vs. below A-levels, higher or lower education, respectively; Equivalised income dichotomised as 2000 and above vs. below 2000, higher or lower income, respectively; Interest in healthy eating dichotomised as 'very interested' vs. 'quite interested', 'not very interested' or 'not at all interested'.

The multiple regression analysis results and interactions between participant characteristics (education, income and interest in healthy eating) are presented in Table D5. There were no significant interactions between FOPL group (N-S and MTL) and any of education level (higher vs lower), income (higher vs lower) and interest in healthy eating (very vs quite/not very/ not all) in terms of associations with global food score. Multiple regression analysis results show that across all FOPL groups global change score was significantly higher in participants who reported being very interested in healthy eating, by 0.2 points (p=0.007), compared to quite/not very/ not at all interested in healthy eating.

Appendix E: Secondary Outcomes Analysis (Chapter 6)

Table E1: Device usage for the experiment by age category

	Computer	Mobile	Both	Total
	n (%)	*n* (%)	*n* (%)	*n*
18-29	74 (27)	203 (73)	2 (1)	279
30-39	146 (23)	492 (77)	4 (1)	642
40-49	272 (31)	603 (68)	6 (1)	881
50-59	398 (41)	558 (58)	13 (1)	969
60-69	476 (51)	446 (48)	11 (1)	933
70+	498 (61)	314 (38)	11 (1)	823

Table E2: Summary of participants who correctly ranked the healthiest product at baseline and follow-up, by FOPL group and product category

	Control		MTL		N-S		WL		PC		Overall	
	B	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**	**B**	**FU**
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**
Pizza	537 (80)	531 (79)	547 (85)	604 (93)	575 (84)	656 (96)	557 (81)	592 (86)	565 (85)	574 (86)	2,781 (83)	2,957 (88)
Drinks	222 (71)	213 (68)	232 (69)	315 (94)	213 (66)	312 (96)	237 (71)	286 (85)	218 (68)	244 (76)	1,122 (69)	1,370 (84)
Cake	518 (83)	505 (81)	543 (83)	609 (94)	529 (84)	603 (95)	509 (80)	576 (91)	507 (82)	491 (80)	2,606 (82)	2,784 (88)
Crisps	546 (73)	549 (73)	556 (75)	695 (94)	569 (74)	739 (96)	565 (75)	668 (89)	530 (75)	542 (77)	2,766 (74)	3,193 (86)
Yoghurt	134 (18)	130 (17)	134 (18)	657 (87)	114 (15)	669 (87)	120 (16)	483 (65)	87 (12)	197 (27)	589 (16)	2,136 (57)
Cereal	384 (49)	370 (48)	354 (47)	678 (90)	354 (47)	684 (91)	339 (45)	577 (76)	349 (46)	409 (54)	1,780 (47)	2,718 (72)

MTL, Multiple Traffic Light; N-S, Nutri-Score; WL, Warning Label; PC, Positive Choice tick; B, baseline; FU, follow-up. Needed to consume the product to be included.

Table E3: Summary of baseline and follow-up ranking times (seconds), by product category, overall and FOPL group (with no requirement to consume product to be included in analysis)

	Control (*n*=913) Median (IQR)		MTL (*n*=907) Median (IQR)		N-S (*n*=924) Median (IQR)		WL (*n*=895) Median (IQR)		PC (*n*=891) Median (IQR)		Overall (*n*=4530) Median (IQR)	
	B	FU	B	FU	B	FU	B	FU	B	FU	B	FU
Pizza	34 (24-50)	17 (12-25)	35 (24-51)	20 (14-33)	34 (26-52)	16 (12-24)	34 (24-50)	22 (15-33)	35 (24-49)	17 (12-26)	35 (25-50)	18 (13-28)
Drink	33 (23-47)	15 (11-22)	34 (23-49)	20 (14-31)	33 (24-47)	16 (12-23)	32 (22-48)	19 (13-27)	32 (23-47)	16 (11-23)	33 (23-48)	17 (12-25)
Cake	25 (18-38)	14 (9-20)	27 (19-38)	19 (13-32)	27 (19-40)	16 (12-23)	27 (18-38)	22 (14-35)	25 (18-37)	15 (10-23)	26 (18-38)	17 (12-26)
Crisps	30 (22-42)	15 (11-22)	29 (22-41)	20 (14-32)	30 (22-42)	17 (12-24)	30 (22-43)	21 (15-29)	29 (21-41)	18 (12-25)	30 (22-42)	18 (13-26)
Yoghurt	36 (25-53)	16 (11-24)	36 (25-52)	21 (14-32)	36 (25-53)	16 (12-23)	36 (25-52)	21 (15-30)	35 (23-51)	17 (12-25)	36 (25-52)	18 (12-27)
Cereal	35 (24-52)	17 (12-25)	36 (25-54)	21 (14-33)	36 (25-52)	17 (16-32)	36 (25-52)	23 (16-32)	34 (24-50)	18 (12-27)	36 (24-52)	19 (13-28)

IQR: interquartile range; MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick
B, baseline; FU, follow-up. Participants needed to have complete covariate information.

Figure E1: Figure of median and interquartile ranges of ranking time (seconds) of baseline and follow-up rankings, by product category. Including outliers, excluding participants who did not have full co-variate information

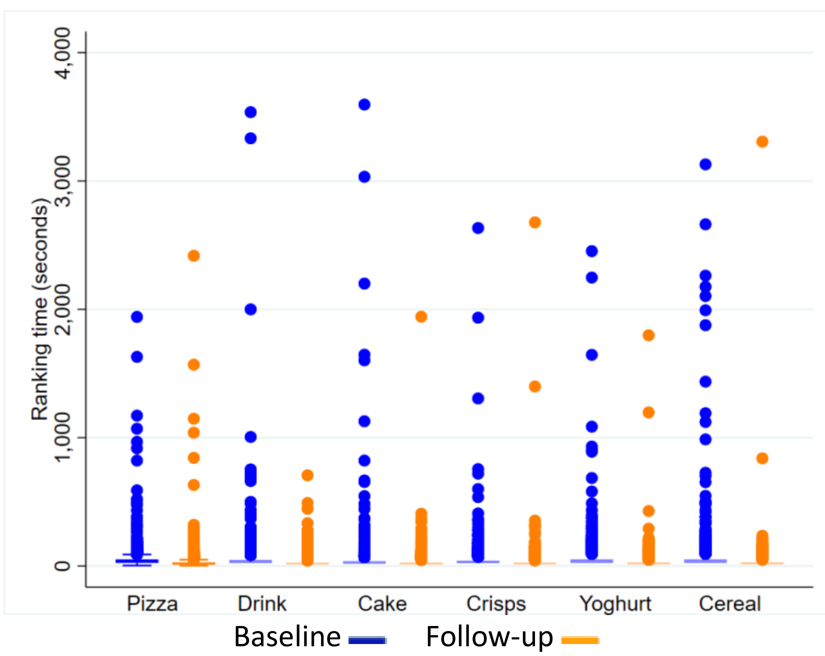


Table E4: Multiple regression analysis results- association between time taken to rank products at follow-up and FOPL group compared to control - without requirement for being correct (adjusted for baseline ranking time, device used, design factors and covariates)

	MTL v Control RM (95%CI)	**N-S v Control RM (95%CI)**	**WL v Control RM (95%CI)**	**PC v Control RM (95%CI)**	**N-S v MTL RM (95% CI)**
Pizza	1.21 (1.15, 1.28) **P < 0.001**	0.99 (0.94, 1.05) 0.779	1.22 (1.16, 1.29) **P < 0.001**	1.02 (0.96, 1.07) 0.591	0.82 (0.77, 0.87) **P < 0.001**
Drink	1.35 (1.25, 1.44) **P < 0.001**	1.08 (1.00, 1.16) 0.045	1.27 (1.18, 1.36) **P < 0.001**	1.05 (0.98, 1.13) 0.182	0.80 (0.75, 0.86) **P < 0.001**
Cake	1.40 (1.32, 1.48) **P < 0.001**	1.15 (1.09, 1.22) **P < 0.001**	1.54 (1.46, 1.63) **P < 0.001**	1.08 (1.02, 1.15) 0.008	0.82 (0.78, 0.87) **P < 0.001**
Crisps	1.39 (1.32, 1.45) **P < 0.001**	1.13 (1.08, 1.18) **P < 0.001**	1.28 (1.22, 1.34) **P < 0.001**	1.14 (1.08, 1.19) **P < 0.001**	0.81 (0.78, 0.85) **P < 0.001**
Yoghurt	1.26 (1.20, 1.33) **P < 0.001**	1.03 (0.97, 1.08) 0.321	1.23 (1.17, 1.29) **P < 0.001**	1.01 (0.96, 1.06) 0.788	0.81 (0.77, 0.86) **P < 0.001**
Cereal	1.25 (1.19, 1.31) **P < 0.001**	0.99 (0.95, 1.05) 0.846	1.32 (1.25, 1.38) **P < 0.001**	1.00 (0.95, 1.06) 0.867	0.80 (0.76, 0.84) **P < 0.001**

All analyses were adjusted for the five stratification factors (year of recruitment to panel, sex, age, government office region, household income) and the following pre-specified covariates: ethnicity, highest education level, household composition, food shopping responsibility, current FOPL use, baseline ranking time and device used. Participants needed to have complete covariate information and buy/eat cereal/pizza to be included. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning label; PC: Positive Choice tick.

Table E5: Median and interquartile ranges of ranking times (seconds) at follow-up, by product category, FOPL group and education level (higher/lower) – without requirement to be correct

Education	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Higher	18 (13-30)	16 (12-22)	21 (14-33)	16 (12-24)	16 (11-23)
Lower	22 (15-38)	18 (13-26)	23 (15-33)	18 (13-30)	19 (13-27)
Instant hot chocolate					
Higher	19 (13-29)	15 (11-21)	18 (13-25)	15 (11-22)	15 (10-21)
Lower	22 (15-34)	18 (13-27)	21 (14-29)	17 (12-30)	17 (12-24)
Cake					
Higher	18 (13-31)	15 (12-21)	23 (15-35)	15 (10-23)	16 (11-25)
Lower	21 (14-35)	17 (13-26)	21 (13-35)	16 (11-24)	16 (10-23)
Crisps					
Higher	19 (14-30)	16 (12-23)	20 (14-28)	17 (12-24)	15 (10-21)
Lower	23 (16-36)	19 (14-27)	22 (16-31)	19 (13-27)	17 (12-24)
Yoghurt					
Higher	19 (14-30)	15 (12-22)	20 (14-29)	16 (11-24)	15 (11-23)
Lower	24 (16-35)	18 (13-25)	23 (16-33)	19 (13-27)	18 (13-28)
Breakfast cereal					
Higher	19 (14-31)	16 (12-22)	22 (15-31)	17 (12-26)	17 (11-24)
Lower	24 (17-39)	18 (14-26)	24 (17-36)	19 (13-30)	19 (13-28)

Education level dichotomised as A levels and above vs. below A levels, higher or lower education respectively.

Table E6: Median and interquartile ranges of ranking times (seconds) at follow-up, by product category, FOPL group and income category (higher/lower) – without requirement to be correct

Income	FOPL group				
	MTL	**N-S**	**WL**	**PC**	**Control**
Pizza					
Higher	19 (13-31)	15 (11-21)	21 (15-30)	17 (12-24)	16 (11-23)
Lower	20 (14-33)	18 (13-25)	23 (15-34)	17 (12-26)	18 (12-25)
Instant hot chocolate					
Higher	19 (13-31)	14 (11-21)	18 (13-24)	15 (11-22)	15 (10-22)
Lower	20 (14-30)	17 (12-24)	20 (14-28)	16 (11-24)	16 (11-22)
Cake					
Higher	19 (13-31)	14 (12-20)	22 (14-34)	15 (11-24)	13 (9-20)
Lower	20 (13-33)	16 (12-24)	22 (15-35)	15 (10-23)	14 (10-21)
Crisps					
Higher	20 (14-30)	16 (12-22)	20 (14-28)	17 (12-24)	15 (10-21)
Lower	20 (15-32)	17 (13-25)	21 (15-29)	18 (12-25)	16 (11-23)
Yoghurt					
Higher	19 (14-30)	15 (11-21)	20 (14-29)	16 (12-26)	16 (10-24)
Lower	22 (14-32)	17 (12-25)	21 (15-31)	17 (11-25)	16 (11-24)
Breakfast cereal					
Higher	20 (14-31)	15 (12-21)	22 (15-31)	17 (12-26)	17 (11-26)
Lower	22 (15-34)	18 (13-25)	23 (16-34)	18 (13-28)	17 (12-25)

Equivalised income dichotomised as 2000 and above vs. below 2000, higher or lower income, respectively.

Appendix F: Latent Class Analysis (Chapter 7)

Table F1: Trend test for indicator variables by predicted class membership

Indicator variables	z	p-value
Current label use	59.3	P<0.001
Influenced by labels	62.9	P<0.001
Interested in healthy eating	36.2	P<0.001
Knowledge about healthy eating	28.0	P<0.001
Currently trying to lose weight	9.5	P<0.001
Responsible for shopping	-7.1	P<0.001
Children in the household	0.5	0.5974

Table F2: Mean global food score by predicted class membership and experimental group

	Total sample	'Highly engaged'	'Somewhat engaged'	'Less engaged'	F (*p*-value)
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	
Control	0.0 (0.8)	0.0 (0.7)	0.0 (0.8)	-0.1 (0.9)	0.44 (0.642)
MTL	1.9 (1.4)	2.0 (1.3)	1.8 (1.5)	1.8 (1.5)	1.00 (0.367)
N-S	2.0 (1.4)	2.1 (1.3)	2.0 (1.5)	1.9 (1.5)	1.61 (0.200)
WL	1.4 (1.4)	1.4 (1.4)	1.4 (1.4)	1.3 (1.6)	0.28 (0.756)
PC	0.1 (0.9)	0.2 (0.9)	0.0 (0.8)	0.1 (0.9)	1.94 (0.144)

MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning Label; PC: Positive Choice tick; SD = standard deviation.

Table F3: Median speed of ranking outcomes by predicted class membership

	Total sample	'Highly engaged' Class 1	'Somewhat engaged' Class 2	'Less engaged' Class 3
	Median (IQR)	**Median (IQR)**	**Median (IQR)**	**Median (IQR)**
Baseline speed (secs) (*n*=54)	205 (154-283)	230 (164-285)	160 (125-218)	215 (159-290)
Follow-up speed (secs) (*n*=1846)	117 (90-164)	120 (91-168)	112 (89-163)	115 (92-151)

Secs = seconds *For speed outcomes there was a requirement of being correct. SD = standard deviation; IQR = interquartile range.

Table F4: Label perception outcomes within FOPL schemes by predicted class and experimental group

	Total	**'Highly'**	**'Somewhat'**	**'Less'**	**χ^2 (*p*-value)**
Saw label (yes)	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	
MTL	740 (76)	457 (80)	191 (74)	92 (67)	13.5 (0.036)
N-S	851 (86)	455 (88)	244 (84)	152 (84)	4.7 (0.314)
WL	749 (77)	417 (79)	246 (78)	86 (70)	6.8 (0.150)
PC	530 (55)	290 (60)	153 (51)	87 (48)	19.3 (0.001)
Used label (all)					
MTL	658 (68)	418 (73)	166 (65)	74 (54)	25.9 (P<0.001)
N-S	730 (74)	400 (78)	209 (72)	121 (67)	17.2 (0.009)
WL	508 (53)	279 (53)	171 (54)	58 (47)	9.8 (0.134)
PC	96 (10)	60 (12)	27 (9)	9 (5)	16.7 (0.033)
Understanding labels (easy)					
MTL	710 (96)	447 (98)	179 (94)	84 (91)	11.6 (0.003)
N-S	753 (89)	406 (89)	222 (91)	125 (83)	6.6 (0.037)
WL	639 (85)	367 (88)	207 (84)	65 (76)	9.6 (0.008)
PC	289 (55)	162 (56)	85 (56)	42 (48)	1.6 (0.439)
Label helpfulness (helpful)					
MTL	939 (97)	567 (99)	253 (98)	119 (87)	59.7 (P<0.001)
N-S	888 (90)	465 (90)	272 (94)	151 (84)	12.3 (0.002)
WL	883 (91)	477 (90)	296 (94)	110 (89)	4.2 (0.121)
PC	705 (73)	324 (67)	251 (83)	130 (72)	24.0 (P<0.001)
Time to use label (quick enough)					
MTL	899 (93)	557 (97)	234 (91)	108 (79)	57.7 (P<0.001)
N-S	957 (97)	502 (97)	284 (98)	171 (95)	6.2 (0.398)
WL	881 (91)	489 (92)	284 (90)	108 (88)	6.9 (0.333)
PC	885 (92)	442 (92)	286 (95)	157 (87)	12.6 (0.013)
Labels on products (yes-all)					
MTL	963 (89)	537 (94)	229 (89)	97 (71)	75.5 (P<0.001)
N-S	748 (76)	404 (78)	227 (78)	117 (65)	18.7 (0.001)
WL	715 (74)	409 (77)	220 (70)	86 (70)	11.4 (0.023)
PC	484 (50)	243 (50)	162 (54)	79 (44)	23.4 (0.001)

*Control group not included. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning Label; PC: Positive Choice tick.

Table F5: Missing data across the indicator and sociodemographic variables in the sample that completed the ranking tasks (*n* = 4,863)

Indicator variables	***n***
Current label use	0
Influenced by labels	0
Interested in healthy eating	0
Knowledge about healthy eating	1/4863
Currently trying to lose weight	0
Responsible for shopping	1/4863
Children in the household	21/4863
Sociodemographic variables	
Sex	0
Age	6
Household income per month	199/4863*
Ethnicity	76/4863
Education	16/4863*
BMI	564/4863 (missing in complete = 512/4530 / n= 4018)

BMI = body mass index; * Missing responses include: refused/don't know/not applicable.

Sensitivity Analyses

The following tables (F6-F14) present the sensitivity analyses conducted in the sample with complete covariates variables (*n* = 4530).

Table F6: Sociodemographic characteristics (frequency or mean) by predicted class membership

	Total sample	'Highly engaged' Class 1	'Somewhat engaged' Class 2	'Less engaged' Class 3	
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	**χ^2 (*p*-value)**
Sex					
Female	2586 (57)	1504 (63)	754 (54)	328 (44)	
Male	1944 (43)	875 (37)	649 (46)	420 (56)	96.4 (<0.001)
Age					
18-29	279 (6)	112 (5)	95 (7)	72 (10)	
30-39	642 (14)	304 (13)	215 (15)	123 (16)	
40-49	881 (19)	470 (20)	293 (21)	118 (16)	
50-59	969 (21)	513 (22)	292 (21)	164 (22)	
60-69	933 (21)	531 (22)	258 (18)	144 (19)	
70+	823 (18)	447 (19)	249 (18)	127 (17)	46.6 (<0.001)
Household income per month (£)					
Less than 1,200	699 (16)	341 (15)	201 (15)	157 (22)	
1,201-2,200	1061 (24)	535 (23)	335 (25)	191 (26)	
2,201-3,700	1253 (29)	654 (29)	396 (29)	203 (28)	
3,701 or more	1355 (31)	764 (33)	417 (31)	174 (24)	35.4 (<0.001)
Ethnicity					
White	4233 (93)	2242 (94)	1292 (92)	699 (93)	
Non-white	297 (7)	137 (6)	111 (8)	49 (7)	6.7 (0.035)
Education					
A levels +	3304 (68)	1866 (73)	980 (66)	458 (57)	
Below A levels	1543 (32)	686 (27)	507 (34)	350 (43)	81.5 (<0.001)
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**F (*p*-value)**
BMI	26.7 (5.5)	26.3 (5.5)	27.1 (5.1)	26.9 (5.8)	7.44 (0.0006)

SD: standard deviation; BMI = body mass index.

Table F7: Multinomial logistic regression of sociodemographic variables on predicted class membership.

	'Somewhat engaged' vs 'Highly engaged'' Class 2 vs 1 (ref)			'Less engaged' vs 'Highly engaged'' Class 3 vs 1 (ref)			'Less engaged' vs 'Somewhat engaged'' Class 3 vs 2 (ref)		
	RRR	***p*-value**	**95% CI**	**RRR**	***p*-value**	**95% CI**	**RRR**	***p*-value**	**95% CI**
Age (18-29 ref)									
30-39	0.83	0.272	(0.60, 1.15)	0.63	0.012	(0.44, 0.91)	0.76	0.145	(0.52, 1.10)
40-49	0.74	0.051	(0.54, 1.00)	0.39	<0.001	(0.27, 0.56)	0.53	0.001	(0.37, 0.77)
50-59	0.67	0.011	(0.49, 0.91)	0.50	<0.001	(0.35, 0. 70)	0.74	0.104	(0.52, 1.06)
60-69	0.57	<0.001	(0.42, 0.78)	0.42	<0.001	(0.30, 0.60)	0.74	0.103	(0.51, 1.06)
70+	0.66	0.009	(0.48, 0.90)	0.44	<0.001	(0.31, 0.63)	0.67	0.038	(0.46, 0.98)
Household income per month (£) (less than 1,200 ref)									
1,201 - 2,200	1.06	0. 593	(0.85, 1.33)	0.78	0.047	(0.60, 1.00)	0.73	0.024	(0.56, 0.96)
2,201 - 3,700	1.03	0.806	(0.83, 1.27)	0.67	0.002	(0.53, 0.86)	0.66	0.002	(0.50, 0.860)
3,701 or more	0.93	0.475	(0.75, 1.14)	0.50	<0.001	(0.39, 0.64)	0.53	<0.001	(0.41, 0.70)
Ethnicity (non-white ref)									
White	0.71	0.010	(0.55, 0.92)	0.87	0.425	(0.62, 1.22)	1.23	0.253	(0.87, 1.74)
Education (below A-levels ref)									
A levels+	0.72	<0.001	(0.62, 0.83)	0.49	<0.001	(0.41, 0.58)	0.68	<0.001	(0.57, 0.82)
BMI (every 1-unit increase)	1.02	<0.001	(1.01, 1.04)	1.02	<0.001	(1.00, 1.04)	1.00	0.718	(0.98, 1.01)

RRR = relative risk ratio; CI = confidence interval. *The 'Highly engaged' class were chosen as the main reference group as it was the largest class.

Table F8: Mean ranking outcomes by predicted class membership

	Total sample	'Highly engaged' Class 1	'Somewhat engaged' Class 2	'Less engaged' Class 3	
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**F (*p*-value)**
Baseline ranking (0-6)	3.0 (1.2)	3.0 (1.2)	3.0 (1.2)	2.8 (1.2)	8.6 (P<0.001)
Follow-up ranking (0-6)	4.3 (1.7)	4.4 (1.7)	4.2 (1.7)	4.0 (1.8)	24.5 (P<0.001)
Global score (-5-+5)	1.1 (1.5)	1.2 (1.4)	1.0 (1.5)	1.0 (1.5)	11.3 (P<0.001)
	Median (IQR)	**Median (IQR)**	**Median (IQR)**	**Median (IQR)**	**χ^2 (*p*-value)**
Baseline speed (secs)	212 (160-290)	218 (165-296)	207 (157-290)	198 (149-266)	20.0 (P<0.001)
Follow-up speed (secs)	115 (87-161)	118 (88-166)	112 (87-159)	110 (85-149)	14.6 (0.001)

Secs = seconds *For speed outcomes there was no requirement of being correct. SD = standard deviation.

Table F9: Univariate multinomial logistic regression of ranking outcomes by predicted class membership.

	'Somewhat engaged' vs 'Highly' Class 2 vs 1 (ref)			'Less engaged' vs 'Highly' Class 3 vs 1 (ref)			Less engaged' vs 'Somewhat' Class 3 vs 2 (ref)		
	RRR	*p*-value	95% CI	RRR	*p*-value	95% CI	RRR	*p*-value	95% CI
Baseline ranking (correct)	0.98	0.508	(0.93, 1.04)	0.87	<0.001	(0.81, 0.93)	0.89	0.001	(0.82, 0.95)
Follow-up ranking (correct)	0.92	<0.001	(0.89, 0.96)	0.86	<0.001	(0.82, 0.90)	0.93	0.004	(0.88, 0.98)
Global food score	0.92	<0.001	(0.88, 0.96)	0.89	<0.001	(0.84, 0.94)	0.97	0.331	(0.91, 1.03)
Baseline speed	1.00	0.214	(1.00, 1.00)	1.00	0.080	(1.00, 1.00)	1.00	0.017	(1.00, 1.00)
Follow-up speed	1.00	0.056	(1.00, 1.00)	1.00	0.050	(1.00, 1.00)	1.00	0.598	(1.00, 1.00)

RRR = relative risk ratio; CI = confidence interval; BMI = body mass index. Class 2, the 'highly engaged' class is the reference group compared to 'somewhat' and 'less engaged' and 'somewhat' compared to 'less engaged'.

Table F10: Proportion of participant label perception outcomes, by predicted class membership

	FOPL groups combined*	'Highly engaged' Class 1	'Somewhat engaged' Class 2	'Less engaged' Class 3	
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	**χ^2 (*p*-value)**
Saw label					35.1 (<0.001)
Yes	2,692 (74)	1,512 (78)	792 (72)	388 (68)	
No/not sure/don't know	925 (26)	434 (22)	305 (28)	186 (32)	
Used label					46.9 (<0.001)
All	1,891 (52)	1,095 (56)	549 (50)	247 (43)	
Some	540 (15)	284 (15)	173 (16)	83 (14)	
Did not use	259 (7)	132 (7)	69 (6)	58 (10)	
Don't know/not applicable	927 (26)	435 (22)	306 (28)	186 (32)	
Understanding labels					24.7 (< 0.001)
Easy	2,260 (84)	1,305 (86)	661 (83)	294 (76)	
Difficult	431 (16)	207 (14)	131 (17)	93 (24)	
Label helpfulness					28.2 (<0.001)
Helpful	3,212 (89)	1,721 (88)	1,011 (92)	480 (84)	
Not helpful	404 (11)	224 (12)	86 (8)	94 (16)	
Time to use label⁺					33.5 (<0.001)
Quick enough	3,385 (94)	1,847 (95)	1,030 (94)	508 (89)	
Too long	225 (6)	95 (5)	64 (6)	66 (12)	
Labels on products in UK^					71.2 (<0.001)
Yes-all	2,649 (73)	1,501 (77)	789 (72)	359 (63)	
Yes-some	694 (19)	293 (15)	250 (23)	151 (26)	
No-none	271 (8)	150 (8)	57 (5)	64 (11)	

*Control group not included. ⁺ 14 do not know or refused responses; ^ 6 do not know responses. Helpful dichotomised as helpful (very helpful, quite helpful) vs not helpful (not very helpful, not at all helpful). Understanding dichotomised as easy (very or quite) vs difficult (quite or very).

Table F11: Frequency and proportion of ranking perceptions, enough information to rank the products and confidence in the ranking tasks, at baseline and follow-up, by predicted class membership

	Total sample ***n*** **= 4,530**		**'Highly engaged' Class 1** ***n*** **= 2,379**		**'Somewhat engaged' Class 2** ***n*** **= 1,403**		**'Less engaged' Class 3** ***n*** **= 748**	
	BL	**FU**	**BL**	**FU**	**BL**	**FU**	**BL**	**FU**
	n (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)
Reported 'yes' to having enough information to rank (*n* = 4,863)								
Pizza	698 (15)	1,871 (41)	330 (14)	990 (42)	224 (16)	576 (41)	144 (19)	305 (41)
Drink	371 (8)	1,792 (40)	167 (7)	955 (40)	117 (8)	547 (40)	87 (12)	290 (39)
Cake	740 (16)	1,965 (43)	336 (14)	1,034 (43)	242 (17)	604 (43)	162 (22)	327 (44)
Crisps	871 (19)	2,004 (44)	413 (17)	1,058 (44)	269 (19)	618 (44)	189 (25)	328 (44)
Yoghurt	488 (11)	1,806 (40)	214 (9)	960 (40)	160 (11)	549 (39)	114 (15)	297 (40)
Cereal	538 (12)	1,828 (40)	240 (10)	972 (41)	178 (13)	561 (40)	120 (16)	295 (40)
Confidence in ranking (how many of them do you think you ranked correctly?)								
All or most	645 (14)	2,043 (45)	354 (15)	1,142 (48)	186 (13)	597 (43)	105 (14)	304 (41)
Some	2,455 (54)	1,742 (38)	1,266 (53)	860 (36)	790 (56)	579 (41)	399 (53)	303 (41)
None	446 (10)	175 (4)	218 (9)	89 (4)	147 (10)	53 (4)	81 (11)	33 (4)
Don't know	984 (22)	570 (13)	541 (23)	288 (12)	280 (20)	174 (12)	163 (22)	108 (14)

BL = Baseline; FU = Follow-up.

Table F12: Mean global food score by predicted class membership and experimental group

	Total sample	**'Highly engaged' Class 1**	**'Somewhat engaged' Class 2**	**'Less engaged' Class 3**	**F (*p*-value)**
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	
Control	-0.0 (0.8)	0.0 (0.7)	-0.0 (0.8)	-0.0 (0.9)	0.56 (0.5708)
MTL	1.9 (1.4)	2.0 (1.3)	1.8 (1.5)	1.9 (1.5)	0.83 (0.4357)
N-S	2.1 (1.4)	2.1 (1.2)	2.0 (1.5)	1.9 (1.5)	2.10 (0.1231)
WL	1.4 (1.4)	1.4 (1.4)	1.4 (1.4)	1.4 (1.6)	0.09 (0.9173)
PC	0.1 (0.9)	0.2 (0.9)	0.0 (0.8)	0.2 (0.9)	2.11 (0.1213)

MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning Label; PC: Positive Choice tick; SD = standard deviation.

Table F13: The p-values of pairwise comparisons of mean global food score between classes, by experimental group

	'Less engaged' vs 'Highly engaged'	**'Highly engaged' vs 'Somewhat engaged'**	**'Somewhat engaged' vs 'Less engaged'**
Experimental group	*p*-value	*p*-value	*p*-value
Control	1.000	1.000	1.000
MTL	0.639	1.000	1.000
N-S	1.000	0.124	0.604
WL	1.000	1.000	1.000
PC	0.130	0.591	1.000
Total	P<0.001	0.001	1.000

MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning Label; PC: Positive Choice tick.

Table F14: Label perception outcomes within FOPL schemes by predicted class and experimental group

	Total	'Highly engaged'	'Somewhat engaged'	'Less engaged'	
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	**χ^2 (*p*-value)**
Saw label (yes)					
MTL	697 (77)	433 (80)	178 (74)	86 (68)	10.5(0.106)
N-S	807 (87)	429 (89)	236 (85)	142 (86)	4.3 (0.362)
WL	698 (78)	388 (80)	233 (79)	77 (69)	7.8 (0.101)
PC	490 (55)	262 (60)	145 (51)	83 (48)	16.4 (0.003)
Used label (all)					
MTL	627 (69)	400 (74)	156 (65)	71 (56)	19.8 (0.003)
N-S	697 (75)	379 (79)	204 (73)	114 (69)	17.2 (0.008)
WL	479 (54)	262 (54)	164 (55)	53 (48)	10.8 (0.093)
PC	88 (10)	54 (12)	25 (9)	9 (5)	15.2 (0.056)
Understanding labels (easy)					
MTL	669 (96)	423 (98)	166 (93)	80 (93)	8.7 (0.013)
N-S	717 (89)	385 (90)	215 (91)	117 (83)	6.5 (0.039)
WL	601 (86)	346 (89)	198 (85)	57 (74)	12.7 (0.002)
PC	273 (56)	151 (58)	82 (57)	40 (48)	2.3 (0.311)
Label helpfulness (helpful)					
MTL	885 (98)	537 (99)	236 (99)	112 (89)	46.7 (P<0.001)
N-S	845 (91)	442 (92)	261 (94)	142 (86)	8.4 (0.015)
WL	822 (92)	443 (91)	278 (94)	101 (91)	2.5 (0.280)
PC	660 (74)	299 (69)	236 (83)	125 (73)	18.3 (P<0.001)
Time to use label (quick enough)					
MTL	846 (93)	527 (97)	217 (91)	102 (81)	49.2 (P<0.001)
N-S	902 (98)	472 (98)	272 (98)	158 (96)	5.8 (0.212)
WL	816 (91)	448 (92)	270 (91)	98 (88)	3.4 (0.754)
PC	821 (92)	400 (92)	271 (95)	150 (87)	12.0 (0.017)
Labels on products (yes-all)					
MTL	813 (90)	510 (94)	213 (89)	90 (71)	68.9 (P<0.001)
N-S	718 (78)	387 (80)	219 (79)	112 (68)	14.5 (0.006)
WL	667 (75)	380 (78)	206 (70)	81 (73)	11.6 (0.021)
PC	451 (51)	224 (51)	151 (53)	76 (44)	19.9 (0.003)

*Control group not included. MTL: Multiple Traffic Lights; N-S: Nutri-Score; WL: Warning Label; PC: Positive Choice tick.

Appendix G: Patient and Public Involvement and Engagement Session (Chapter 8)

Mentimeter Survey

Baseline ranking task

Baseline ranking (without label): "*Please rank the [cakes/pizzas/cereals] in order of healthiness – most healthy, in between, least healthy*." Under each product was a letter – A, B, or C and participants would select a product via the corresponding letter from the drop-down box for most healthy (1st) to least healthy (3rd), see Figure A1.

Enough information (baseline ranking): "*Did you feel you had enough information to rank the [cakes/pizzas/cereals]?*", with response options 'yes' or 'no'

Follow-up ranking task

Follow-up ranking (with label): "*We are now going to show you more pictures of the same food items; Please rank the [cakes/pizzas/cereals] in order of healthiness – most healthy, in between, least healthy*." Under each product was a letter – A, B, or C and participants would select a product via the corresponding letter from the drop-down box for most healthy (1st) to least healthy (3rd).

Enough information (follow-up ranking): "*Did you feel you had enough information to rank the [cakes/pizzas/cereals]?*", with response options 'yes' or 'no'

Survey

Label perceptions

- *"Which label did you like most?"* with response options 'MTL', 'N-S', or 'PC'
- *"Which label did you find easiest to use?"* with response options 'MTL', 'N-S', or 'PC'
- *"Which label did you find quickest to use?"* with response options 'MTL', 'N-S', or 'PC'
- *"Would you like labels to be put on food and drink packaging in the UK?"* with response options 'Yes – all products', 'Yes – some products', 'No – none'

Food shopping and eating habits

- *"Do you ever do any food shopping, either for yourself or someone else?"* with response options 'Yes – I do some food shopping' or 'No – someone else does food shopping for me'
- *"How often do you read the labels on the front when buying food or drink?"* with response options 'Very often', 'Quite often', 'Occasionally', 'Rarely', 'Never'
- *"How much knowledge would you say you have about healthy eating?"* with response options 'A lot of knowledge', 'Some knowledge', 'A little knowledge', 'No knowledge'
- "How interested would you say you are in healthy eating?" with response options 'Very interested', 'Quite interested', 'Not very interested', 'Not at all interested'

Figure G1: Ranking task example from Mentimeter

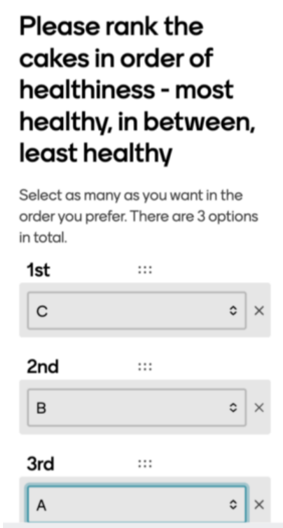


Figure G2: Poster presented to DHSC and submitted to the UCL Great Ormond Street Institute of Child Health Postgraduate Research Showcase