I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12675

Facilitating Healthy Dietary Habits: An Experiment with a Low Income Population

Michèle Belot
Jonathan James
Jonathan Spiteri

OCTOBER 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12675

Facilitating Healthy Dietary Habits: An Experiment with a Low Income Population

Michèle Belot
European University Institute and IZA

Jonathan James
University of Bath

Jonathan Spiteri
University of Malta

OCTOBER 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Facilitating Healthy Dietary Habits: An Experiment with a Low Income Population*

This paper tests an intervention aimed at facilitating (cognitively) the adoption of healthy dietary habits. We provide easy-to-understand information about the risks of developing diabetes or heart diseases and give easy-to-follow dietary recommendations to minimize these risks. We implement two variations, one consisting of generic information, the other consisting of information tailored to the individual, the latter resembling newly developed on-line health assessment tools. On top of the information treatment, we implement a second experimental variation nudging people into spending more time thinking about their dietary choices. We find evidence that the information intervention leads to healthier choices in the short run, but mostly in the generic treatment. Surprisingly, we find that people are on average pessimistic about their health, and therefore receive good news on average when the information is tailored to them. We find no evidence that increasing the time available to make choices leads to healthier choices, and find no evidence of long-term changes in habits. These results do not support a bounded rationality explanation for poor dietary choices.

JEL Classification: I12, I18, H51, D83

Keywords: health risks, dietary habits, bounded rationality, heuristics, information, time availability, laboratory experiments

Corresponding author:
Michele Belot
School of Economics
University of Edinburgh
30 Buccleuch Place
Edinburgh EH8 9JT
United Kingdom
E-mail: michele.belot@ed.ac.uk

* This project is funded by the European Union's Seventh Framework Programme, under grant agreement no. 607310. A pre-analysis plan was submitted and published on the American Economic Association's RCT Registry in June 12th 2016 (Reference: AEARCTR-0001189. https://doi.org/10.1257/ rct.1189-1.0). The deviations from the pre-analysis plan are reported in the Appendix. We are grateful to seminar participants at Newcastle, Bath, and European Society of Population Economics (2017), Royal Economic Society (2018), and Health Economics Study Group (2018) conferences for their comments and suggestions. This paper is accompanied by Supplementary Material in the Online Appendix available at: http://www.mwpweb.eu/1/132/resources/document_1141_1.pdf

I. Introduction

A poor diet is now the leading contributor to early deaths around the world (Forouzanfar et al. (2015)). Poor dietary choices have been linked to various risk factors and diseases, such as high blood pressure, diabetes and obesity.

Over the last decade behavioral economists have developed interventions targeting unhealthy habits. The prevalent approach has followed the Becker and Murphy (1988) view of habits, where habits are modelled as the result of a process of consumption shaping future preferences and future consumption. However, as it is, there is so far little evidence for such a process being at work in the context of diet (Belot et al. (2018)). In contrast, psychologists view habits as the result of a cognitive process (see Verplanken and Aarts (1999), Wood and Rünger (2016), Wood and Neal (2016)). The idea is that optimizing is cognitively costly, and developing simple heuristics (habits) saves on cognitive costs. Such heuristics should a priori be optimal, but they may not be if the environment changes and individuals do not re-optimize. In a bounded rationality world, re-optimization only takes place when changes in the environment are sufficiently large for people to notice them and find it worthwhile to re-optimize (see Wood and Neal (2016) for a recent discussion). In the context of health-related behaviors, an example could be an information shock, such as being diagnosed with a disease. One worry is that at that point changes may not be as effective, and as a consequence, it may be useful to think of interventions that alert individuals early on about the necessity to change their behavior. It is also plausible that heuristics are suboptimal because adopting a healthy diet is not straightforward. In contrast to other health-related behaviors such as smoking or exercising, adopting a healthy diet is a priori more complex and requires processing a fair amount of information. Finally, dietary heuristics could also be suboptimal because people are systematically misinformed, e.g. are too optimistic about their health or unaware of the link between their lifestyle and their health. Previous research has found that individuals systematically overestimate the risk of developing lung cancer as a result of smoking (Viscusi (1991) Viscusi (1990)), alcohol addiction (Lundborg and Lindgren, 2002) and the risks of having a heart disease or a stroke as a result of being obese (Winter and Wuppermann (2014)).

In this paper, we conduct a laboratory experiment testing an intervention designed around the cognitive approach to habits. The intervention aims at facilitating re-optimization and the adoption of healthy dietary habits. We provide easy-to-understand information about health risks and giving simple recommendations. We implement two variations: One group receives generic information on the average risks of contracting heart disease or diabetes, as well as easy-to-follow recommendations on lifestyle changes that can reduce the risks of developing either of these diseases. The information is provided in an easy-to-understand manner (e.g. "eat 5 portions of fruit and vegetables a day"), but is not tailored to the individual. The second group of participants receives personalized health information via a specialized computer-based tool. The tool is an adapted version of a publicly available assessment tool called 'Your Disease Risk' (YDR) (Baer et al. (2013)), which provides individual information about risks of developing diabetes and heart disease, and then provides tailored recommendations to reduce the risks. A third group (control), receives neutral information that has no link to health or diet (an article on architecture). We measure participants' dietary choices through an incentivized shopping task on a digital platform. Participants receive GBP 30 to spend and can choose across over 100 food and drink items, varying in their nutrition profiles.

On top of this intervention, we test the bounded rationality hypothesis more directly by encouraging people to devote more time to their dietary choices. Participants (across all information groups) are split into two further groups - a 'long time' group and a 'short time' group. The long time group has 10 minutes to choose their basket of food and drinks, while the short time group has 3 minutes. Note that the only variation we introduce here is in the time participants have to ponder their purchases. We deliberately do not give them access to more or less information about products across treatments. The idea here is simply to induce participants to spend more time "thinking" about their choices, holding information constant.

We recruited participants with a household income below the UK median. We focus on this group for two reasons: First, disadvantaged socioeconomic groups appear more vulnerable to the obesity epidemic. Levine (2011) finds that in the U.S. the prevalence

of obesity among adults is 145% higher in counties with poverty rates over 35%. Second, recent evidence by Mani et al. (2013) and Laraia et al. (2015) argue that the poor may be more at risk of poor decision-making. The hypothesis is that poverty takes up a considerable proportion of cognitive resources, and therefore causally affects decision-making. If this view is correct, interventions that minimize the cost of re-optimizing may be particularly appropriate for this group.

Using the nutritional information from the chosen basket of food and drink items, we evaluate the impact of health information and time availability on food choices, controlling for a number of factors like current state of hunger, prior health knowledge, prior health status (and knowledge of such status), current dietary habits, socio-economic factors and demographic indicators. We also conduct a follow-up session 3 months later in order to measure the long-term impact of information provision on people's food choices.

The results provide evidence that participants in the information treatments select healthier food baskets that contained a lower proportion of unhealthy items[1]. The results are however stronger in the generic health information treatment where selected food baskets contained around 17% less total fat, 20% less saturated fat and 15% less salt relative to the no information group.

The fact that the effects are stronger in the generic treatment may appear surprising. To understand why this may be the case, we examine how the treatments affected participants' beliefs about their health. We find that participants in the tailored health information group received good news on average, that is, they were told that their relative risk of heart disease and diabetes is lower than they thought. As a consequence, beliefs about their own health status became more optimistic after receiving the tailored information. Notably, the good news did not translate into statistically worse dietary choices, however. If anything, we find evidence that participants chose food items with lower calories, irrespective of whether they received good or bad news. Because of the randomization, we can assume that participants in the generic information were most likely too pessimistic as well, but they did not receive tailored information and they did

[1]To classify items in the tool as healthy or unhealthy we use the nutrient profiling technique developed by the UK's Food Standards Agency (FSA). This is set out in more detail in section III.F.

not update their beliefs about their own health. Most participants report that the information was not novel, thus, it seems unlikely that the channel driving behavioral change is information. Rather it seems that the intervention affected behavior through salience.

Our second experimental variation - variation in the allotted time to shop for a food basket - did not affect dietary choices significantly. End surveys suggest that the variation worked in the sense that people spent more time on their choices: those in the longer time treatment were more likely to report they had enough time to choose their purchases. But this extra time did not lead to different choices. In fact, combining with the results on the effects of information, the evidence suggests that people do not need much thinking time to adjust their choices. They can easily follow simple healthy dietary recommendations.

A follow-up session held three months after the initial experiment shows no strong evidence of long term changes in dietary habits. We only find a negative effect of the tailored health information treatment on the number of calories in foods chosen. Overall, our results confirm that salience and attention may play a key role in behavior that is habitual in nature, but also suggest that people are in fact not too optimistic about their health. This may explain why recent efforts by public health agencies to offer personalized health information have had limited success so far. For example, the NHS's Health Check programme, aimed at providing a free health assessment for people aged 40 to 74, appears to have had limited effectiveness despite its £165 million annual cost (Chang et al. (2016)).

The remainder of the paper is structured as follows. Section 2 summarizes the related literature. Section 3 describes the design of our experiment. Section 4 presents the short run results and Section 5 focuses on the longer run effects. Section 6 concludes.

II. Related Literature

To our knowledge, there has been little work on interventions targeting re-optimization of dietary choices. A recent study by Carrera et al. (2016) looks at whether the provision tailored health information in relation to cholesterol levels has any impact on food choices among employees at a hospital in the US. Hospital workers were incentivised to undertake a biometric health screening test to measure various health characteristics, including

cholesterol, glucose levels, blood pressure and BMI (although the authors focus solely on cholesterol levels in their study). The authors combine these readings with data on weekly food purchases from the hospital cafeteria and cafes both before and after the actual health assessment. The results show a statistically-significant decline in total spending on food purchases among those who were diagnosed as 'high risk', particularly among those who were previously unaware of their cholesterol status, as well as a significant increase in the proportion of healthy items purchased among high-risk participants over the age of 55, and women.

Another related study is by Oster (2015) who looks at the impact of diabetes diagnosis on food purchases using household scanner data. Individual diagnosis of diabetes is inferred from purchases of glucose testing products (like glucose monitors) that show up on the household scanner data over time, since such items are required in order to manage their disease and track their blood sugar levels. The results show that, post-diagnosis, households purchase slightly fewer calories, and that in the first month, healthy food purchases increase while unhealthy foods decrease, although only this latter effect persists over time. This is in line with a growing literature on the impact of disease diagnosis on eating patterns, like for example Zhao et al. (2013) who, using panel data from China, find that people reduce their intake of fat in the 12 months immediately after a hypertension diagnosis, although this effect is strongest among the richest third of the population.

These papers focus on the effects of individually tailored information. In contrast to these studies, the interventions we propose aim at facilitating re-optimization by making concrete easy-to-follow recommendations, in addition to providing information about health. We also have a generic information treatment, which does not require access to individual information and may therefore be easy to implement.[2] Another important dif-

[2]The type of tailored information we provide is also different. In Carrera et al. (2016) participants were told their cholesterol levels (among other biometric measures), which is one of the risk factors that can lead to cardiovascular disease. At the other end of the scale, in Oster (2015) people have received a diagnosis of diabetes. Our tailored information lies somewhere in between, since although the Your Disease Risk (YDR) tool is not intended to diagnose either heart disease or diabetes (in fact people with pre-existing diseases were excluded from this study), it combines several risk factors (including self-reported cholesterol, among other things) to calculate the relative risk of developing heart disease and diabetes over the next 10 years relative to the average person of the same age and gender living in Scotland. In this respect, our paper is more in line with the likes of Dupas (2011), who find that providing teenagers in Kenya with the relative risk of developing HIV according to their partner's age significantly

ference is the nature of our sample. We focus on a low-income population, whereas these studies do not have a specific target group. It could be that re-optimization is a more acute issue in a poor population, as hinted by the recent work by Mani et al. (2013). Finally, our data comes from a lab experiment whereas the other two studies use data from the field. The advantage here is that we have detailed information on the individuals, their health history, as well as their beliefs about their health and anthropometric measurements. The caveat is that the choices we study are made in a laboratory setting. We discuss external validity in Section III.H.

Also related to our work, a number of recent studies look at the effects of providing general health and nutritional information to consumers on food choices. For example, Wisdom et al. (2010) find that providing calorie content information on menus at Subway restaurants reduced calorific intake by approximately 7%. Similarly, Bollinger et al. (2011) look at calorie posting at another chain restaurant, this time Starbucks, with results showing that although average calories per transaction fell by around 6%, this was solely driven by changes in food choices, with zero impact on drinks. Our paper differs in that we study the impact of more general, easy-digestible health information that incorporates facts regarding heart disease and diabetes as well as dietary and lifestyle recommendations to reduce the risk of illness.

This paper also relates to the literature on uncertainty and updating of beliefs regarding people's health and associated behaviors. The uncertainty and lack of knowledge regarding health or disease incidence is well-documented (e.g. Crossley and Kennedy, 2002 and Barrett-Connor et al., 2011), as is the general lack of awareness regarding lifestyle risk factors (e.g. Sanderson et al., 2009). In both cases, standard economic theory would suggest that people will, when faced with new health information, update their beliefs and will make healthier dietary choices, as suggested by the evidence presented in studies such as Carrera et al. (2016) and Zhao et al. (2013). However, there is also evidence to suggest that when it comes to certain health-related behaviors, people actually *overestimate* the risks involved in terms of falling ill or developing a disease. For example, Viscusi and

reduces the risk of unprotected sex.

Hakes (2008) report that adults on average overestimate the lung cancer risks of smoking, as well as the mortality risks and life expectancy losses. The authors find that higher risk beliefs reduce the likelihood of starting to smoke, and increase the probability of smoking cessation among smokers. It follows that new information regarding the true risks of smoking would lead to an increase in smoking and reduced efforts to quit. In our paper, we find that in the tailored information group, people's self-reported beliefs regarding their own health improve on average after receiving the personalized health information, since the overwhelming majority of participants receive good news regarding their risk of developing heart disease and diabetes. This in turn does not lead to healthier food choices.

Our paper also contributes to the broader topic of information processing and bounded rationality. Although the literature on bounded rationality is voluminous (see Conlisk, 1996 for a comprehensive survey), our understanding of how bounded rationality may affect food choices is limited. In a computerized experiment, Scheibehenne et al. (2007) find that a simple heuristic whereby participants focused on one product attribute (e.g. convenience) could be used to explain people's food choices during the experiment. Another related experiment is by Reutskaja et al. (2011) who use eye-tracking technology to analyze the computational processes that people undertake when selecting among various snack items, with a time limitation of 3 seconds to make each choice, in order to mimic a real-world supermarket situation. The authors find that in general when making choices people search for a random amount of time, depending on the value of the items under selection, and then pick the best option that they have seen, at odds with optimal search models. Our paper fits in with this literature by introducing a time availability treatment whereby some participants only have 3 minutes to select their £30 food and drink basket as opposed to 10 minutes. The key difference in our case is that rather than looking at 'optimal' choices, our aim is to see whether restricted time availability has any effect on the healthiness of the food choices.

III. Experimental Design

The first and main part of the study was conducted from Monday 13th June to Friday 17th June 2016, and was held in the Behavioural Laboratory at the University in Edinburgh (BLUE). Each day there were four time slots: 9.30am-11am, 11.30am - 1pm, 2.30pm - 4pm, 5.30pm - 7pm. We conducted the experiment in 20 sessions, with up to 18 individuals per session. The 20 sessions were spread over 5 consecutive days. We assigned treatments to sessions and pre-assigned sessions to specific slots in order to guarantee a balance of treatments across times of the day and days of the week. Participants were offered £50 compensation for taking part in the experiment. Participants were able to indicate their preferred time slots, but were not informed in advance of the treatments associated with each time slot. Participants received an information leaflet in advance at their home address and were asked to sign a consent form on the day of their first visit.

III.A. Sample and Recruitment Procedure

In total, we recruited a sample of 318 participants, with a focus on low-income individuals (with an annual income below £26,500) living in the proximity of the University of Edinburgh's main campus. More specifically, participants had to satisfy the following eligibility criteria:

- Must be over 18 years of age;
- Must live in Edinburgh;
- Must be fluent in English;
- Must have an annual household income below £26,500;
- Must not be currently undertaking any regular medical treatment;
- Must not be pregnant.

We recruited participants via three main channels, namely the distribution of information leaflets by post to home addresses in the more deprived neighborhoods in the vicinity

of the University, which was done through a local marketing intelligence firm; online advertisements on one of the leading classified advertising websites in the UK (Gumtree); as well as promotional emails sent out to non-academic and non-student members of the BLUE mailing list.[3]

III.B. Procedure

Upon arriving at the BLUE lab, all participants were asked to fill out an initial questionnaire, which included questions related to demographics, socio-economic background, education, employment status, as well as various questions related to their prior knowledge regarding health, nutrition and their own health status. Participants were also asked to complete a short food frequency questionnaire[4] (based on the National Cancer Institute's Dietary Screener Questionnaire) in order to obtain a measure of their typical eating habits.

Following this initial stage, we then moved on to the actual interventions.

Participants were assigned to one of six (6) groups upon registration (based on their registration slot, which was pre-assigned to a specific treatment[5]):

1. No Information/Short Time
2. No Information/Long Time
3. Generic Information/Short Time
4. Generic Information/Long Time
5. Tailored Information/Short Time
6. Tailored Information/Long Time

[3]A dedicated website and registration page was created in order to handle registrations, and prospective participants were given a contact number and email address in case of any queries. Ethical approval was granted on May 10th, 2016 by the School of Economics' Ethics Committee. A pre-analysis plan was submitted and published on the American Economic Association's RCT Registry in May 2016 (reference: AEARCTR-0001189). The analysis in this paper, except where otherwise stated, follows what was originally set out in this plan.

[4]The food frequency questionnaire is shown in full in the appendix (Figure A.1).

[5]The allocation of treatments to session days and times can be found in Table A.1 in the appendix.

III.C. Information Intervention

The first intervention in our study is related to the provision of different types of health information to our participants. In particular, we are interested in understanding the impact of providing easy-to-digest information, and the distinction between tailored versus generic health information. We had a no information (control) group, which was asked to read a non-health related article on architecture, taken from Wikipedia. This article was chosen following a pre-experimental session held on Thursday June 2nd 2016 at 4pm, where it was deemed by participants to be both unrelated to health and emotionally-neutral, which is particularly important given the well-documented impact of emotional state on food choices (Gibson (2006)).

The tailored information treatment group was provided with personalized health information via an adapted version of a computer-based health assessment tool called 'Your Disease Risk' (YDR) (Baer et al. (2013)). The algorithms developed for this tool are used to predict the chances of developing a particular disease over the next 10 years, relative to the average person of the same age and gender. We adapted the algorithm for the Scottish population using data from the Scottish Health Survey, 2008-2012. These calculations are done on the basis of a series of questions that respondents are asked to fill in, related to their medical history, their parents' medical history, dietary habits (e.g. consumption of fruit and vegetables per day) and lifestyle choices (e.g. smoking, exercise, average daily alcohol consumption). Once these questions are answered, the YDR tool provides a scale showing the respondent's relative risk of developing a particular disease, which ranges from 'Very Much Below Average' to 'Very Much Above Average'. The system also provides tailored recommendations to respondents which would help them to lower their risk (e.g. 'Eat more unsaturated fats', 'Stop smoking'.). For the purpose of this study, we focus solely on two diseases, namely heart disease and diabetes. Note that the questions required as inputs for the YDR tool are already included in the initial questionnaire described earlier, and hence were answered by all participants. However, the Tailored Information group are the only participants to receive the YDR risk and recommendations (other groups were not aware of this treatment).

The generic information treatment group were given a two-page document with information published by the NHS and Harvard Medical School. In essence, this document first provided information regarding the average risk of developing heart disease and diabetes in Scotland (for the entire adult population), as well as some details regarding each disease. The wording is exactly identical to that used in the tailored information treatment, with the only difference being that in this case the risk is general for the entire population rather than personalized for each individual based on their answers. The document also provided the full list of recommended actions that would lower the risk of developing each disease, as specified in the YDR tool. Again, the wording is exactly identical to that used in the tailored info treatment, with the only difference being that in this case the full list of potential recommendations is provided rather than those specifically pertaining to each individual. Thus, the design is set up to ensure maximum comparability across treatments as this allows us to assess the marginal impact of tailored health information relative to generic information. A copy of the generic information document related to heart disease is provided in the Appendix, along with a sample results page from the tailored information tool (Figure A.2A and A.2B).

III.D. Time Intervention

The second intervention in our study is related to the time available for each participant to make their food choices. The idea behind this treatment is that people have limited cognitive resources that must be allocated across various competing concerns, meaning that matters related to nutrition and health may not be high on an individual's list of priorities. The lack of health prioritisation may be more acute for people from a low income background since various studies (e.g. Mani et al, 2013) have shown that poverty-related concerns occupy a large portion of a low-income person's mental resources, leaving less room for focusing on other matters. Therefore in this study we are explicitly trying to test this idea by introducing exogenous variation in the amount of time available for participants to make their dietary choices. Note that we chose this intervention rather than an intervention aiming at depleting cognitive resources, because cognitive resource

depletion may affect behavior through self-control, which is not the mechanism we are interested in here. We want to test whether nudging people to think harder about their choices has an effect on their decisions, holding self-control constant.

At the start of this intervention, all participants were allocated a budget of £30 to spend on food and drink from a specially-designed choice tool that appears similar to an online supermarket developed specifically for this study[6]. The choice tool contains a total of 120 food and drink items. We chose the 10 most popular items across the following grocery categories: fruit and vegetables, meat, fish, confectionery, chilled meals and drinks. We ended up with a mix of 66 "healthy" and 54 "unhealthy" items.[7] Apart from capturing the participants' food choices in terms of which items were actually selected, the system has been designed to calculate the nutritional value of each basket along several key nutrients, namely calories, carbohydrates, total fats, saturated fats and sugar content. These will be used to construct our main outcome variables. All prices used in the supermarket tool reflect current market prices at the leading high street supermarkets in the UK, in order to make the food selection task more realistic. Participants were allowed to spend their budget on any of the items listed in the supermarket tool, just as long as they did not exceed the £30 limit.

The experimental variation is related to the *time* available to select food and drink items. The Long Time group were given 10 minutes to make their choices (and were required to stay for the entire duration), while the Short Time group were given just 3 minutes. Both time periods were pre-tested in BLUE before the start of the experiment. At the end of each session, 1 subject per session was picked at random and his/her food basket was delivered to his/her home address two weeks after participation. This waiting period was chosen to ensure that their choices would not in some way be influenced by the current stock of food that participants had at home at the time of the experiment.

[6]For a full presentation and evaluation of the tool see Spiteri et al. (2019).

[7]A complete list of all the food and drink items included in this food choice tool is provided in Table A.2 in the appendix.

III.E. Post-Treatment and Follow-up

At the end of the session, all participants were asked to fill in a short questionnaire, which is primarily designed to answer three questions:

- Whether the participants updated their beliefs regarding their own health status following the information treatment;
- Whether the participants believe that the information provided was credible/trustworthy or not;
- Whether the choice tool was easy to use and comparable to their typical supermarket shopping experience.

In order to gauge the long-term impact of the health information intervention, we also ran a follow-up session 3 months later from Monday 12th September to Friday 16th September 2016. Participants were asked to complete a short questionnaire aimed at eliciting their beliefs regarding their health status and whether they had undertaken any dietary changes (particularly for those who received tailored health information). They were also asked to complete a food frequency questionnaire, a 24-hour dietary recall (using the INTAKE24 software developed specifically for the UK[8]), and were once again allocated a £30 budget to spend using our food choice tool. In this instance there were no time restrictions on their food choices - all participants had a maximum of 10 minutes to make their choices.

III.F. Outcome variables and Hypotheses

As mentioned above, we use the data gathered from the food choice stage to construct our outcome variables, which include the following:

- The proportion of unhealthy items, where an item is classified as 'unhealthy' is based on the UK Food Standards Agency's nutrient profiling technique. Points for each item are allocated on the basis of the nutrient content of 100g of a food or drink.

[8]Screenshots of the programme can be found in appendix (Figure A.3.)

Points are awarded for energy, saturated fat, total sugar and sodium (A-nutrients), and for fruit, vegetables and nut content, fibre and protein (C-nutrients). The points from C-nutrients are then subtracted from the score for A-nutrients to calculate a final score. The unhealthy items are then classified as foods with 4 or more points and drinks with 1 or more points.[9]

- The nutrient content of each participant's food basket, where the nutrients under consideration are calories, total fat, saturated fat, sugar, salt, fibre, and protein (we estimate a separate regression for each nutrient)[10].;
- The proportion of the budget (£30) spent on fruit and vegetables;

We are interested in testing the following hypotheses:

1. Participants who receive generic health information will on average select healthier food/drink items relative to those who receive no health information.
2. Participants who receive tailored health information will on average select healthier food/drink items relative to those who receive no health information.
3. Participants who have more time available to make their food choices will select a healthier food basket than those with less time availability.

The first two hypotheses will hold if participants lack information and, perhaps additionally, overestimate how healthy they are. For the second hypothesis, the direction of the change should depend on the type of health information received (good or bad news): those with below average risks could 'reward' themselves with unhealthier choices, while those with above average risks could pick healthier choices.

The third hypothesis ties in with the literature on the impact of time constraints on decision-making (Svenson and Edland, 1987; Van Herpen and Van Trijp, 2011). However, it is important to note that this prediction is highly dependent on the type of decision-making rule used by individuals when making choices under time pressure. Findings from

[9]For full details of how the points are calculated see `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216094/dh_123492.pdf`

[10]We present total nutrient content in the main text and nutrients per 100g in the appendix.

the sizable literature on behavioral psychology (e.g. Beach and Mitchell, 1978 and Ford et al., 1989) suggest that time limitations have a significant impact on people's decision-making processes, leading to the minimization of cognitive effort as people increasingly rely on a variety of heuristics or shortcuts rather than carefully-considered judgments. These heuristics can be both intrinsic to the individual and related to his/her own biases or extrinsic and part of the decision-making context. In our case, the somewhat artificial laboratory setting, where participants are fully aware that they are part of a study, may lead to healthier choices even under time constraints. Similar to the previous discussion, the way in which intrinsic heuristics may influence people's food choices will to some extent depend on the individual's existing diet/health, with people typically relying on what they are already familiar with in order to make rapid decisions (Dijker and Koomen (1996)).

For the long term analysis, we will also look at dietary or lifestyle changes that the participants report having undertaken in the 3 months after the initial intervention. The follow-up questionnaire contains two questions related to this matter. The first question states: "Looking back over the last 3 months, have you made any changes to your diet or lifestyle habits?", which elicits a simple 'Yes' or 'No'. The second question follows-up on this point: "If yes, please indicate the change in your diet or lifestyle that you have undertaken from the list below", which then lists the following options: "Stop smoking, Do more exercise, Eat more fruit and vegetables, Eat less junk food and processed foods, Eat less sugar, Eat less red meat, Take vitamins and other supplements, Drink less alcohol". Therefore, we can use this data to obtain two outcome variables:

1) A simple dummy variable indicating whether any changes have been made; 2) A series of eight dummy variables for each response.

In addition, we will also look at the changes in BMI and waist circumference.

The hypothesis we wish to test here is that those who received generic or tailored information will have report more lifestyle changes than the control group, and as a consequence may have a lower BMI and waist circumference.

III.G. Power calculations

We ran power calculations on our data, based on the sample of 309 subjects across all interventions, both in terms of the information treatments (90 control group; 111 generic information group; 108 tailored information group) and time treatments (153 short group; 156 long group). When it comes to the information treatments, our power calculations show that our sample size is sufficient to detect aggregate nutrient content differences of 15% across treatments with over 80% power for total calories, sugar, fibre and protein, 76% power for differences in salt content, 67% power for differences in fat content, and 48% power for differences in saturated fat content. In the time availability intervention the sample size is enough to detect effect sizes of 15% in nutrient content with at least 80% power, depending on the nutrient under consideration.

III.H. External Validity

The design has been chosen to mimic a familiar shopping environment for food choices. There are however potential limitations to the external validity of the study, which we discuss briefly here.

First, one may be concerned that the procedure (lab experiment) may lead to a Hawthorne effect. In contrast to studies conducted in the field, where participants are unaware they are part of an experiment, here participants will easily understand that the study is about health and food choices. They may therefore make healthier choices than they otherwise would. Perhaps even more relevant for our research question, the fact that participants are put in a novel environment and are presumably paying more attention to the tasks than they would in the field, it could be that the experiment triggers re-optimization for everyone, including the control group. This is plausible, and of course means that the treatment effects would be underestimated. Another related concern is that our participants may perceive the £30 budget as a 'windfall' gain, and as a consequence spend it differently than they normally would. To gauge whether these concerns are legitimate, we compare the average proportion of the total budget spent on different food categories in our experiment to the mean household choices as recorded in

ONS Family Spending data. The results for each category are shown in the Appendix in Table A.3. Because of the experiment we would of course expect that choices may be healthier than the average. Also, as we will show below, our cohort is on average relatively healthier than the Scottish population. Nevertheless, we find that spending across each category is relatively comparable across both our experiment and the population average. Crucially, our participants spent almost an identical percentage of their budget on meat and fish relative to the household mean (34.1%), which in our food choice tool was the most expensive category with a average price per item of 3.06. We do however observe that participants in our experiment spent a higher proportion of their budget on fruit and vegetables (34.6%), which could be due to the composition of participants but also to a possible Hawthorne effect.

Second, our experimental sample is not representative of the (low income) Scottish population and may be positively selected in terms of health. We will show below that this is indeed the case. However, there is concern and evidence that those who participate and take up the NHS health check are the so-called "worried well" - those who are already healthy and proactive and are looking for reassurance, (Riley et al. (2016) Gøtzsche et al. (2014)). Therefore, the sample could still be representative of potential users of this type of tools.

Third, the study also focuses on planned consumption rather than immediate consumption. It may be that different heuristics apply in the case of immediate consumption. Nevertheless, about 5 percent of purchases in Great Britain now takes place on-line and therefore involves planned consumption[11] and the online supermarket shopping channel is the fastest growing food purchase channel. There is evidence that planned food choices tend to be healthier than immediate food choices (Milkman et al. (2010)), thus we will not claim the conclusions we find here apply to immediate consumption as well.

[11]Statistics reported by Kantar WorldPanel for 2017

V. Results

V.A. Summary Statistics

We begin by presenting summary statistics of participants across the different treatment groups. Table 1 looks at participant demographic characteristics, education, and income. We also examine features of the experiment such as the timing of the experiment, which page the participants saw first from the food choice tool, and whether they were hungry at the time of the experiment. Finally, we present summary statistics on measures of pre-experiment dietary habits and health behaviours.

The first column presents the summary statistics for all the 309 participants who took part in the experiment. The sample was 39% male. Most were single (70%) with an average age of 36. The average participant was just overweight with a BMI of 25. The majority of the sample had a University degree, and 56% were employed. Nearly all (93%) had a household income lower than £25,000 which is in line with the recruitment criteria.

In Table A.4 in the Supplementary Appendix we present a comparison of our sample to that of the Scottish population from the Scottish health survey (SHS). Our sample, despite being drawn from the Scottish population as opposed to the student cohort, is healthier than the average person living in Scotland. We can examine this by comparing various health and dietary measures in our data with the average recorded in the Scottish Health Surveys from 2008 to 2015. This comparison is seen in Table A.4. We make two comparisons. First, with all those in the Scottish Health Survey, and second with those with a household income below £26,000 to match as closely as possible to our experimental sample income criteria. Our sample has a similar gender composition compared to that from the SHS, however, it is younger, less white and more likely to be married. Our sample exhibits several characteristics that are consistent with a healthier lifestyle relative to the general Scottish population, including a lower proportion of overweight or obese people (BMI $\geq$ 25), a lower incidence of family heart problems, a higher proportion of people who eat five or more portions of fruit and vegetables a day. There is similarity in taking of vitamins between our sample and the two samples of the SHS. The experimental sample seems to be similar to the overall Scottish population along the lines of current

smoking prevalence but the prevalence is lower compared to the low income sample. One health behaviour where our sample displays less healthier behaviours is that of alcohol consumption. Our sample have a higher rate of daily alcohol consumption than both SHS samples.

Returning to Table 1 The distribution of front pages of the food choice tool is broadly even between the various categories, as is the timing of the experiment. Most (around 80%) of those taking part were not hungry at the time of the experiment. Respondents were also asked about their diet and health behaviours. The majority of the participants followed a diet without dietary restrictions (78%), 48% eat fish at least twice a week, however 52% did not regularly eat 5 or more fruits and vegetables per day. Around 20% currently smoke and just under 30% have quit smoking.

Columns 2 to 4 in table 1 present the means of variables by treatment group for the information treatment and columns 6 and 7 present the means for the time treatment. We test for balance in the information treatment by regressing the characteristics on the two treatment indicators, column 5 presents the p-value of the test of their joint significance hence testing the equality of the three groups, and column 6 presents the p-value of the t-test of the 'long time' dummy.

Over some measures, there are a few significant differences between participants in the treatment and control groups. We find that there is a higher proportion of men in the tailored information treatment and a lower proportion in the generic information group relative to the control, the two treatment groups are also slightly younger than the control group. We also find that the tailored information group was on average more likely to report that they were hungry at the start of the experiment. We will address this and the other imbalances in more detail below. There are no differences across the control group and treatment groups in qualifications, employment, and income, nor for the view of the front page of the tool, or along the diet and health behaviors prior to the experiment. For the time treatment, we find a significantly higher proportion who are employed in the long time treatment. For the other categories (besides the timing of the experiment), we do not find other significant differences. For both treatment arms we do find systematic

differences between the treatment and control groups along the timing of the experiment. It is worth recalling that we in essence have 6 treatment arms, and that treatments were assigned at the session level. Hence having one or two more sessions in the morning or afternoon with a particular treatment can lead to imbalance at the individual level. It was the case that all treatments took place at each time slot, but because of the session sizes, we have either 3 or 4 sessions for each of the main information treatments (tailored, generic, no info) in the morning or in the afternoon. Due to the differences across groups we will report how the treatment effects change when controlling for the variables for which there is imbalance. These are shown in Figure 1 and are discussed in more detail later. In brief, our results are robust to the inclusion of these variables.

Next we present the results from the pre-intervention survey where we elicited the beliefs of the participants regarding their risk of developing heart disease and diabetes over the next ten years. There is reason to believe that participants may systematically overestimate or underestimate their risk of developing certain diseases. Previous research has found that individuals systematically overestimate the risk of developing lung cancer as a result of smoking (Viscusi (1991) Viscusi (1990)), alcohol addiction (Lundborg and Lindgren, 2002) and the risks of developing obesity related diseases as a result of being obese (Winter and Wuppermann (2014)). Table 2 shows that participants systematically overestimate their risk of developing heart disease. The table shows the joint distribution of the participants'beliefs about their pre-treatment risk of developing heart disease (panel A) and diabetes (panel B) and their risks as assessed by the Your Disease Risk tool. The percentage who hold correct beliefs are shown along the diagonal in bold. Only 18% and 24% correctly identify their correct risk for heart disease and diabetes respectively. We now turn to the main results of the experiment.

V.B. Information Treatment

The most basic specification in this case is a linear model where we regress the nutritional content of each participant's food choices on each of the basic treatment dummies:

$$Y_i = \alpha + \sum_{k=1}^{2} \beta_k I_k + \mathbf{X}_i' \delta + \epsilon_i \quad (1)$$

where Y_i is one of the four outcome measures for participant i, β_k is the coefficient of interest related to each of the $k = 2$ treatments [12], I_k is a dummy variable for each of the information treatments, and ϵ is an idiosyncratic error term. We also add a vector of control variables $\mathbf{X}_i$ including age, being male, a set of indicators for the time of the experimental session and dummy indicating whether the participant was hungry or not.

In the first instance we present robust standard errors. These are on average more conservative than standard errors that are clustered at the session level. Although the participants could not see or interact with each other there could be a concern that the there is correlation at the session level for a variety of reasons. Given the small number of clusters, however, we therefore perform the wild bootstrapping procedure, clustering at the session level (Cameron et al. (2008)). In addition, we also perform three further corrections to the p-values. First, we perform a randomization inference procedure set out in Young (2019). This involves a test of a sharp null (all participants of a particular treatment have a zero treatment effect rather than an average treatment effect of zero). Second, we take into account of the multiple comparisons problem by using the by using the False Discovery Rate, FDR (the share of significant estimates that are expected to be false positives) of Anderson (2008) as set out in Anderson (2008). We also correct for multiple comparison issue using the more conservative adjustment for the family wise error rate (FWER) as proposed by Romano and Wolf (2005a,b) - this shows us the chance that at least one of our outcomes within the family of outcomes is significant when the null hypothesis of no effect is true. We do not include the proportion spent on unhealthy items in our multiple hypothesis correction as that outcome is a composite measure of the nutrients.

The results in Table 3 show that participants in the generic information treatment

[12] Where $k = 1$ denotes the generic health information treatment, while $k = 2$ denotes tailored health information.

made dietary choices that were on average healthier. They chose a basket that had 4.5 percentage points (22.8%) fewer unhealthy items relative to the control group. As previously mentioned in section III.F, unhealthy items are classified using the UK Food Standards Agency's nutrient profiling system that allocates points for each item on the basis of the nutrient content of 100g of a food or drink. Points are awarded for energy, saturated fat, total sugar and sodium (A-nutrients), and for fruit, vegetables and nut content, fibre and protein (C-nutrients). The points from C-nutrients are then subtracted from the score for A-nutrients to calculate a final score. The unhealthy items are then classified as foods with 4 or more points and drinks with 1 or more points. We next turn to these other elements that determine whether something is classified as unhealthy or not to understand which nutrients were driving the improvement in the healthiness of the basket.

Baskets in the generic treatment were 43g (17%) lower in total fat, 19g (24%) in total saturated fat, and 3.2g (15%) lower in salt. The estimates for fat and salt remain statistically significant at the 5% level when the wild bootstrap and the randomization inference procedures are carried out. Once we correct for multiple hypothesis testing using the FDR approach the coefficients on fat, saturated fat and salt are not longer significant at the 5% level but remain so at the 10% level. However, when the we correct using the more conservative approach correcting for the FWER (Romano and Wolf (2005a,b)) our estimates no longer statistically significant at the 10% level (the p-values are 0.126, 0.165 and 0.127 respectively on fat, saturated fat and salt). Overall, it appears that the basket became healthier due to the reduction in A-nutrients and not an increase in the C-nutrients.

By contrast, we find weaker effects of the tailored information treatment. Overall, the baskets in the tailored treatment were on average healthier by around 3% however the estimate is very imprecise. Most of the coefficients for the various nutrients are negative

and none are statistically significant at conventional levels[13][14].

Since there are small differences in covariates across treatment and control, we examine the impact that the inclusion of controls has on the estimates. Figure 1 shows the impact of the information treatments for a baseline without controlling variables (1), and six different sets of controls: (2) includes demographic and controls for body size, (3) includes controls for qualifications and income, (4) includes controls for the experimental conditions of time of day and the front page of the tool, (5)controls for measures of pre-experiment diet and health behaviors, and finally (6) includes all control variables. The treatment effects are robust and relatively unaffected by the inclusion or exclusion of controls. The magnitude of both treatment effects is greater (in absolute terms) when time of the experiment and front page of the experiment are controlled for, although these differences are not statistically significant. All treatment effects are of a very similar magnitude and point to the same conclusion: There is an impact of providing generic information on the healthiness of the basket and in particular on the amount of fat, saturated fat and salt in the foods chosen but there is little impact of the tailored information treatment.

Of course the average treatment effect in the tailored information treatment could mask important heterogeneity, depending on whether people received good or bad news.

Good News and Bad News[15] — We next examine the nature of the tailored information that was provided by the YDR tool. In our initial questionnaire we ask the participants to indicate what they think their relative risk of developing a particular disease is along the same scale as the YDR tool – ranging from 'Very Much Below Average' to 'Very Much Above Average'. We then know whether the information provided by the YDR tool gave the respondent good news or not. The information provided by the tool falls into four main categories. First, an individual can receive good news, this is when the information

[13]In the appendix we also present (Table A.5) the analysis where the dependent variables are the mean amount of the nutrient per 100g. The point estimates in this instance are all in the same direction, however, the estimates of the impact of the generic treatment on the consumption of fat and saturated fat are no long statistically significant in contrast the coefficient on the generic treatment on salt remains statistically significant.

[14]We also interact the tailored treatment with the specific recommendation given by the YDR tool. The estimates of this analysis are presented in the appendix (Table A.6). We do not find a relationship between the tool's specific recommendations and basket choices.

[15]Note that this is additional analysis not specified in the pre-analysis plan in order to further understand our main findings.

is better than the participants' expectations. As we have provided information regarding two diseases, we define good news as getting good news about one disease and at least expected news about the other. The second category is mixed news, this is where the individual receives good news (or expected news) for one disease and bad news for the other. The third group is where in both cases the respondent YDR tools reports the risk that they expected to receive. The final category is bad news where in each case the information provided from the tool suggested a higher relative risk than the responded expected. For the majority of our participants who were in the tailored information group, 71.3%, received good news 14.8% were given mixed news, 6.5% received the news they expected and 7.4% got bad news.

We again estimate the impact of the information treatments, this time breaking up the tailored information into main two categories – receiving good news or not, where not receiving good news is made up of the other three categories described above. The results of this exercise are presented in table 4. We find some heterogeneity in the response to the tailored health information, but the difference between the two tailored information is not statistically different. The estimates for both the tailored information groups are imprecisely estimated.

Post Treatment Beliefs — We are also interested in looking at variation in dietary choices according to people's beliefs *after* treatment. This is because post-treatment beliefs would provide a good indicator of whether participants have understood or indeed believe the health information provided and updated their beliefs regarding their own health, which in turn is more likely to influence their dietary choices.[16] In the post-experiment questionnaire we ask: "Having read the information provided in this study, indicate the extent to which you believe that you are leading a healthy life?" which is then followed by a scale from 0 to 100 ranging from 'Very Unhealthy' to 'Very Healthy'. This is directly comparable to the question asked in the initial questionnaire pre-treatment, therefore enabling us to evaluate whether participants have changed their beliefs regarding their

[16]In the post-experiment questionnaire, an average of 98% of respondents across the two treatment groups stated that they found the information to be 'credible or trustworthy', and a further 92% found the information to be 'useful'.

own health following treatment. On average, we find that people across all 3 information treatments initially rate themselves at 67, i.e. moderately good health (there is no statistically-significant difference across the 3 treatments in the initial survey period). As described above we also ask respondents, in the initial questionnaire, what they believe their chances are of developing heart disease and diabetes. We ask these questions again after the treatment. In addition, we ask whether the information that was provided was new to them or not.

Table 5 shows the impact of the information treatments on the changes in lifestyle beliefs, changes in the beliefs of disease risks and whether the information that was provided was new. We find that the tailored health information treatment leads people to positively update their beliefs regarding their lifestyle. They report a higher score of living a healthy life after they received the tailored information. There is an increase by around 4.9 percentage points (around 7%) in the extent to which participants in the tailored treatment believe they are living a healthy lifestyle. Given, as described above, that most people received good news from the tailored information treatment then this is what we would expect to see. The difference between the generic and tailored treatments is significantly different along the lifestyle measure.

In the next two columns we examine the change in beliefs of developing heart disease (column 2) and diabetes (column 3). The tailored information results in participants reporting a lower risk for both diseases with a coefficient of -1.1 (heart disease) and -0.5 (diabetes). To put this into context, in the control group prior to being given information the mean of these variables are (3.17 and 3.19 respectively). The reductions we find in the changes in perceived risk are therefore large. This is in line with the information provided and the results on changes in lifestyle and provides context for the lack of statistically significant impact of tailored health information on food choices. Furthermore, figure 2 also shows that on average, people in this group received positive news regarding their health, leading to an upward revision in their beliefs, which in turn may not have induced them to select healthier items from the food choice tool.The final column in Table 5 shows that the tailored information participants were more likely to report the information

provided was new. This is a large but imprecisely estimated effect.

The perceptions of health risks in the generic group were largely unaffected by the treatment. This result, along with participants revising down their lifestyle rating, points to generic information not influencing food choices by virtue of novelty, but rather due to salience, by reminding participants of the risk of heart disease and diabetes and the dietary changes required to lower this risk, as mentioned in the previous section.

To further evaluate the issue of salience and novelty[17]. We next split the treatment groups into two further categories for each treatment type. Each treatment is split according to whether the participant reported that the news they received was new or not. Within the Generic information treatment 90 participants (81%) reported that the information was not new, with the remaining 19% reporting that it was. Within the Tailored information treatment 80 people (26%) reported that the information was new compared to 28 (74.1%) reporting that it was not. Overall, the basket of those who thought the information was not new was significantly healthier by 5 percentage points relative to the control, for those in the generic treatment who found the information was new their basket was around three points healthier than the control albeit imprecisely estimated - furthermore, the difference between the old news and new news groups was not statistically significantly different. The impact of generic information does lead to a basket lower in salt (and higher in spending on fruits and vegetables) that is significantly different for those who did not find the generic information to be new. The generic information appears to have an impact if it acts as a reminder rather than providing surprising information.

We do not find a significant impact on the healthiness of the food basket in the tailored information treatment for either the new or old news groups. Overall, there is a mixed picture coming from the tailored group - with those who did not think the information was new choosing less protein but spend more on fruit and vegetables compared to those who found the new information novel. However, the numbers in these groups are small so we should take this analysis with caution.

[17]Note that this is additional analysis not specified in the pre-analysis plan in order to further understand our main findings.

V.C. Time Availability Treatment

We now evaluate the second experimental variation, related to the amount of time available for participants to select their food and drink items from the food choice tool. The basic specification is a linear model which can be described as follows:

$$Y_i = \rho A_i + \mathbf{X}_i'\alpha + \upsilon_i \tag{2}$$

where Y_i is the nutritional content of participant i's chosen food basket, A_i is a dummy variable denoting whether the participant was part of the long-time treatment group, ρ is our parameter of interest and υ_i is an idiosyncratic error term.

We do not have precise records on the time spent on the task, but we included a question in the end survey asking whether participants felt they had enough time to shop. For those in the Long Time (10 minutes) treatment 80% reported that the time they had was just right, and 20% reported they had too much time to spend the £30. None reported that they had too little time. In contrast, 22% of those in the Short Time (3 minutes) treatment reported that they had too little time with 3.9% still reporting they had too much time - the remaining 73.9% reported that the time was just right.

The results in Table 7 show that there is no statistically-significant difference between the two time treatments in terms of the nutritional content of the food choice baskets selected by participants in either group. We do observe coefficients on our high time dummy that go in the direction of healthier purchases (except for sugar), but all of these estimates are imprecise. For comparison, the point estimates for all the outcomes, except sugar which is almost three times as large, are a fraction of the estimates of the generic information. These results are not consistent with the hypothesis that people choose unhealthy foods because of time constraints.

It is possible that participation in the study made health concerns salient and lead to healthier choices than in real life. As described in the experimental design, the study involved a series of questionnaires related to health, lifestyle and nutrition, quite apart from the health information treatment(s). Thus, matters related to health and nutrition

were already quite salient in the minds of participants when reaching the food choice stage. Those in the long time treatment were able to ponder their choices, while participants in the short time treatment may have had to rely on simple heuristics when selecting food items. It is possible that the context of the experiment triggered a "healthy" heuristic. This is a key consideration since our theoretical predictions depend on the type of decision-rule people use when making their dietary choices. In both tailored and generic treatments participants received recommendations that were relatively easy to apply immediately. It appears that, when possessing such information, being nudged to spend more time on decisions has little effect[18]

V.D. The Long-Term Impact of Information on Dietary Choices

The follow-up experiment was held 3 months after the initial experiment, from Monday 12th September to Friday 16th September 2016. In total 265 participated, representing over 83% of the original sample. Table A.8 in the appendix compares the sample characteristics of the initial participants to those who showed up for the follow-up session in September. The two samples are statistically very similar to one another, with none of the characteristics exhibiting any significant changes across the two samples, which is unsurprising given the low dropout rate. Therefore, this similarity across samples helps to allay any concerns regarding attrition bias, which may limit the validity of any analysis of long-term impacts resulting from our information treatments.

Once again, participants were paid £50 compensation for attending, and in total we had 4 sessions a day with a maximum of 18 participants per session. Upon arrival, participants' weight and waist size were measured. The first stage of the follow-up consisted of a survey where the main aim was to gauge any dietary or lifestyle changes that had occurred since June, while the second stage was the same food frequency questionnaire undertaken 3 months earlier. The third stage was a 24-hour dietary recall, using the computer-based INTAKE24, and the experiment concluded with the food choice tool from the previous

[18]We also estimate the interaction of the time and information treatments. This is shown in Table A.7 in the online Appendix. We only find those in the long time group who also received tailored information chose food baskets higher in sugar, none of the other interactions between the two different treatments were statistically significant.

session, albeit with no variation in time availability. One participant was again picked from each session at random in order to receive his/her chosen food and drinks basket.

In the follow-up session we are mainly interested in analyzing whether exposure to health information has any long-term impact on people's food choices. Therefore, we shall once again be estimating equation (1), with the only difference being that for the follow-up session we are interested in different dependent variables. These are:

- Nutritional composition of the participants' food/drink basket in the follow-up session;
- The difference in participants' biometric measures: BMI and waist size (in inches);
- The participants' reported dietary and/or lifestyle changes as reported in the follow-up session questionnaire.

Food Choices — We start with participants' food choices as elicited from the supermarket tool. The aim here is to check whether the results observed in the initial experiment still hold now, namely in relation to the reduced fat and saturated fat content, as well as lower proportion spent on unhealthy items, observed among the generic information treatment relative to the control group. We therefore focus on the same outcome variables that were analyzed in the previous section. We will also use the same set of controls used before.

The results are presented in Table 8. We do not find participants chose healthier baskets in either the generic or tailored treatment in the follow up session three months later. None of the coefficients on the generic treatment are statistically significant. In the tailored treatment, in contrast, the baskets are lower in calories however this is somewhat offset by the basket being lower in fibre - leaving the overall (un)healthiness of the basket unchanged.

Returning to Figure 2 we again examine participants' subjective beliefs regarding their health status in September. All participants were asked to rate their current health status during the follow-up session on a scale from 0 to 100, ranging from 'Very Unhealthy' to 'Very Healthy'. We can thus trace the evolution of participants' health beliefs both before

and after treatment (in June), as well as 3 months later in September. These health beliefs are shown in the final column of the figure. We can see that for both the control and generic information groups, subjective health beliefs have stayed largely constant over time, with no statistically-significant changes across the three periods for both groups. In contrast, participants in the tailored information group, who had become more optimistic after receiving the information, are no longer so in the follow up, which suggests that the positive effect generated by the good news has somewhat worn off over time.

Changes in Body Measurements, Self-Reported Diet and Lifestyle Changes — We now analyze whether the health information had any impact in terms of participants' body measurement and whether participants reported any changes to their diet or lifestyle in the 3 months since the original experiment.

As part of the first stage of the follow-up experiment, participants were asked to indicate whether they had undertaken any dietary or lifestyle changes over the 3-month period since the initial session. They were given a list of options to select from that included:

- Do some form of moderate physical exercise for at least 30 minutes on most days
- Stop smoking
- Cut down the amount of alcohol I drink
- Lose some weight
- Take a multivitamin (like a B complex vitamin)
- Eat 2 or more servings of fish per week
- Eat 5 or more servings of fruit and veg per day
- Eat 3 or more servings of whole grains on most days
- Reduce consumption of saturated fats like red meat, cheese and whole milk

- Reduce consumption of trans-saturated fats like cookies, pies, chips, crisps and deep-fried food

We therefore use these responses as our outcome variables, using the same regressors as in Table 8. The results are shown in Table 9. We do not find any statistically-significant differences in the self-reported changes of either the generic or tailored health information groups relative to the no information (control) group. In the final two columns we analyze two new outcome variables representing the relative change in BMI and waist size from the initial measurement to the follow-up. We then use the same regressors used in Table 8 to analyze these changes across our information treatments. We observe no statistically significant difference in either BMI or waist size in either information treatment. Our sample's initial average BMI and waist size in June was 25.27 and 33.31 inches respectively, with these figures barely changing 3 months later (25.30 and 33.34 inches respectively). Furthermore, the BMI average is just on the borderline between healthy and overweight as prescribed by nutritionists, while the waist average is below the recommended limit of 37 inches for men and just above the 31.5 inches for women. Hence, there may have been limited scope for our participants to lose any weight given that their starting point was already relatively healthy to begin with.

VI. Concluding Remarks

In this paper we sought to analyze the extent to which it is possible to nudge people into re-optimizing their dietary choices. We introduce two sources of experimental variation, one where we provide easy-to-digest health information, generalized (generic) and personalized (tailored); and a second where we vary the time available to shop for a basket of food.

The results show that participants in the generic health information group selected food baskets that, on average, contained less total fat and less saturated fat (approximately 20% less) relative to the no information group, and spent 34% less on unhealthy items. We also find a (weaker) effect of providing tailored information on the foods chosen, although the picture is less clear. We find a significant effect on the amount of calories chosen, but no significant effect on other measures of the nutrient profile of the baskets.

Further analysis suggests that the majority of tailored health participants received *positive* news from the health assessment tool regarding their relative risk of developing both heart disease and diabetes (i.e. below average). However, we find no difference in responses to whether the news was good or bad. That is, beliefs do not appear to play a significant role on choices.

Our second result is that nudging people into spending more time on their dietary choices has little impact on how healthy those choices were.

Our findings indicate that the majority of generic information participants were already familiar with the material presented, since they did not alter their own health perceptions after reading this information. These results support the idea that generic health information influenced people's choices via salience by reminding them of how to reduce their heart disease and diabetes risk, rather than due to the novelty of the details presented. In fact, we find that on average people are not too optimistic about their health, and the tailored treatment does not reveal information that should trigger an improvement in dietary choices.

Finally, we analyze the long-term impact of our health information intervention on people's food choices, any dietary or lifestyle changes undertaken since the experiment, as well as their body measurements. In most cases, we found no significant difference in food choices across all three treatment groups, suggesting that the impact of generic health information is largely instantaneous with no longer-term effects. We did find that tailored information participants picked lower calorie baskets on average relative to the control group, although this was mainly driven by the healthier participants in our sample. Similarly, we also find no statistically-significant differences with regards to any lifestyle or dietary changes, and no difference in BMI or waist size.

Overall, the results presented in this paper have important implications for the design of future health information campaigns. There is now a growing trend towards providing tailored health recommendations. Here we find that participants have relatively pessimistic beliefs about their health and likelihood of developing diabetes or heart disease. As a consequence, these tailored tools may not have the effects intended by their design-

ers, although here we find no evidence of a backfiring effect. Perhaps surprisingly, we find that generic information does affect choices in the short run. These results are in line with recent evidence on the effectiveness of salient information made available at the time of purchase, such as calorie information on product labels (see for example Wisdom et al., 2010 and Bollinger et al. (2011))

References

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association 103*(484), 1481–1495.

Baer, H. J., L. I. Schneider, et al. (2013). Use of a web-based risk appraisal tool for assessing family history and lifestyle factors in primary care. *Journal of general internal medicine 28*(6), 817.

Barrett-Connor, E., J. Z. Ayanian, et al. (2011). *A Nationwide Framework for Surveillance of Cardiovascular and Chronic Lung Diseases*. Institute of Medicine of the National Academies.

Beach, L. R. and T. R. Mitchell (1978). A contingency model for the selection of decision strategies. *Academy of management review 3*(3), 439–449.

Becker, G. and K. Murphy (1988). A theory of rational addiction. *The Journal of Political Economy 96*(4), 675–700.

Belot, M., N. Belin, J. James, and V. Skafida (2018). The formation and malleability of dietary habits: A field experiment with low income families. *IZA Discussion Paper Series* (11317).

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological) 57*(1), 289–300.

Bollinger, B., P. Leslie, and A. Sorensen (2011). Calorie posting in chain restaurants. *American Economic Journal: Economic Policy 3*(1), 91–128.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Carrera, M., S. A. Hasan, and S. Prina (2016). Tailored health information and eating behavior: Effects of a health risk assessment on employees with different ex-ante awareness of their health risks. *Unpublished manuscript*.

Chang, K. C.-M., J. T. Lee, E. P. Vamos, M. Soljak, D. Johnston, K. Khunti, A. Majeed, and C. Millett (2016). Impact of the national health service health check on cardiovascular disease risk: a difference-in-differences matching analysis. *Canadian Medical Association Journal*, cmaj–151201.

Conlisk, J. (1996). Why bounded rationality? *Journal of economic literature 34*(2), 669–700.

Crossley, T. F. and S. Kennedy (2002). The reliability of self-assessed health status. *Journal of health economics 21*(4), 643–658.

Dijker, A. J. and W. Koomen (1996). Stereotyping and attitudinal effects under time pressure. *European Journal of Social Psychology 26*(1), 61–74.

Dupas, P. (2011). Do teenagers respond to hiv risk information? evidence from a field experiment in kenya. *American Economic Journal: Applied Economics 3*(1), 1–34.

Ford, J. K., N. Schmitt, S. L. Schechtman, B. M. Hults, and M. L. Doherty (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes 43*(1), 75–117.

Forouzanfar, M. H., L. Alexander, H. R. Anderson, V. F. Bachman, S. Biryukov, M. Brauer, R. Burnett, D. Casey, M. M. Coates, A. Cohen, et al. (2015). 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 386*(10010), 2287–2323.

Gibson, E. L. (2006). Emotional influences on food choice: sensory, physiological and psychological pathways. *Physiology & behavior 89*(1), 53–61.

Gøtzsche, P. C., K. J. Jørgensen, and L. T. Krogsbøll (2014). General health checks don't work. *BMJ 348*.

Laraia, B., L. C. Vinikoor-Imler, and A. M. Siega-Riz (2015). Food insecurity during pregnancy leads to stress, disordered eating, and greater postpartum weight among overweight women. *Obesity 23*(6), 1303–1311.

Levine, J. A. (2011). Poverty and obesity in the us. *Diabetes 60*(11), 2667–2668.

Lundborg, P. and B. Lindgren (2002). Risk perceptions and alcohol consumption among young people. *Journal of Risk and Uncertainty 25*(2), 165–183.

Mani, A., S. Mullainathan, E. Shafir, and J. Zhao (2013). Poverty impedes cognitive function. *science 341*(6149), 976–980.

Milkman, K. L., T. Rogers, and M. H. Bazerman (2010, Mar). I'll have the ice cream soon and the vegetables later: A study of online grocery purchases and order lead time. *Marketing Letters 21*(1), 17–35.

Oster, E. (2015). Diabetes and diet: Behavioral response and the value of health. Technical report, National Bureau of Economic Research.

Reutskaja, E., R. Nagel, C. F. Camerer, and A. Rangel (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *The American Economic Review 101*(2), 900–926.

Riley, R., N. Coghill, A. Montgomery, G. Feder, and J. Horwood (2016). Experiences of patients and healthcare professionals of nhs cardiovascular health checks: a qualitative study. *Journal of Public Health 38*(3), 543–551.

Romano, J. P. and M. Wolf (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association 100*(469), 94–108.

Romano, J. P. and M. Wolf (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica 73*(4), 1237–1282.

Sanderson, S. C., J. Waller, M. J. Jarvis, S. E. Humphries, and J. Wardle (2009). Awareness of lifestyle risk factors for cancer and heart disease among adults in the uk. *Patient education and counseling 74*(2), 221–227.

Scheibehenne, B., L. Miesler, and P. M. Todd (2007). Fast and frugal food choices: Uncovering individual decision heuristics. *Appetite 49*(3), 578–589.

Spiteri, J., J. James, and M. Belot (2019). A computer-based incentivized food basket choice tool-presentation and evaluation. *14(1):*.

Svenson, O. and A. Edland (1987). Change of preferences under time pressure: Choices and judgements. *Scandinavian Journal of Psychology 28*(4), 322–330.

Van Herpen, E. and H. C. Van Trijp (2011). Front-of-pack nutrition labels. their effect on attention and choices when consumers have varying goals and time constraints. *Appetite 57*(1), 148–160.

Verplanken, B. and H. Aarts (1999). Habit, attitude, and planned behavior: Is habit an empty construct or an interesting case of goal-directed automaticity? *European Review of Social Psychology 10*(1), 101–134.

Viscusi, W. and J. K. Hakes (2008). Risk beliefs and smoking behavior. *Economic Inquiry 46*(1), 45–59.

Viscusi, W. K. (1990). Do smokers underestimate risks? *Journal of political Economy 98*(6), 1253–1269.

Viscusi, W. K. (1991). Age variations in risk perceptions and smoking decisions. *The Review of Economics and Statistics*, 577–588.

Winter, J. and A. Wuppermann (2014). Do they know what is at risk? health risk perception among the obese. *Health economics 23*(5), 564–585.

Wisdom, J., J. S. Downs, and G. Loewenstein (2010). Promoting healthy choices: Information versus convenience. *American Economic Journal: Applied Economics 2*(2), 164–178.

Wood, W. and D. Neal (2016). Healthy through habit: Interventions for initiating maintaining health behavior change. *Behavioral Science and Policy 2*(1), 71–83.

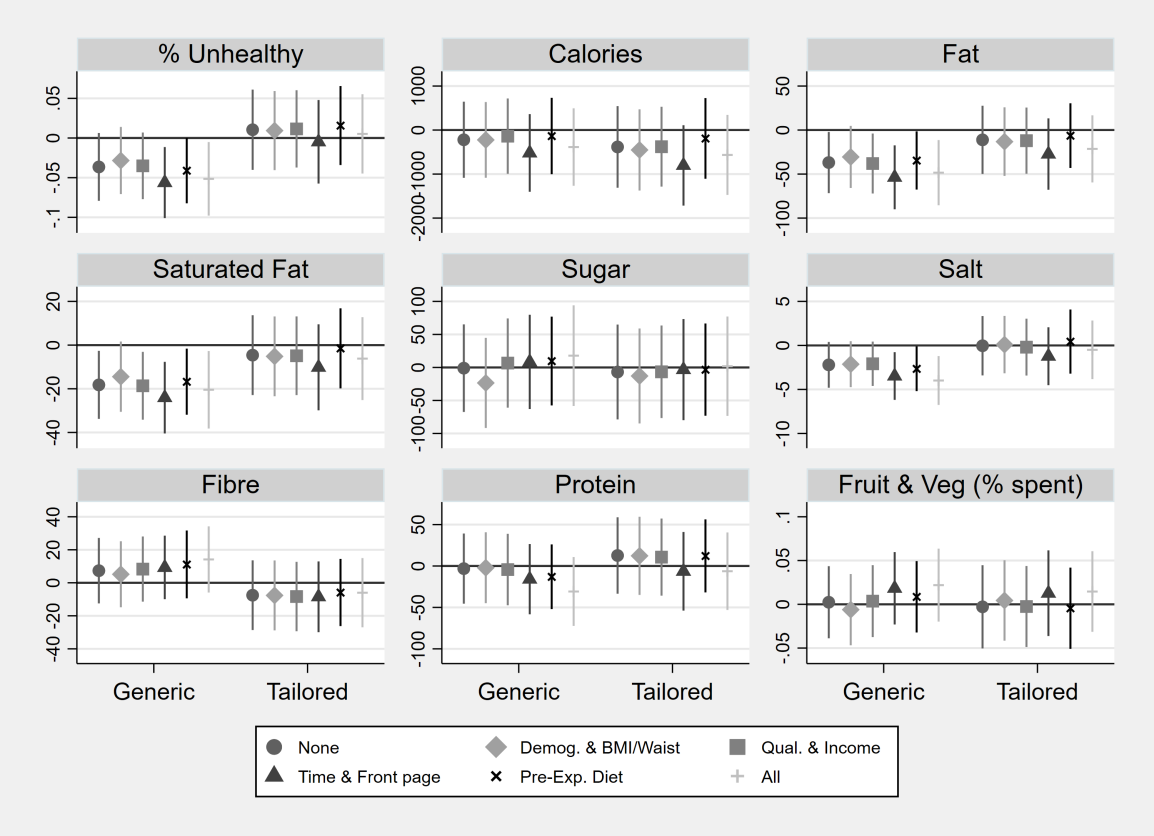
Wood, W. and D. Rünger (2016). Psychology of habit. *Annual Review of Psychology 67*, 289–314.

Young, A. (2019). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The Quarterly Journal of Economics 134*(2), 557–598.

Zhao, M., Y. Konishi, and P. Glewwe (2013). Does information on health status lead to a healthier lifestyle? evidence from china on the effect of hypertension diagnosis on food consumption. *Journal of Health Economics 32*(2), 367–385.

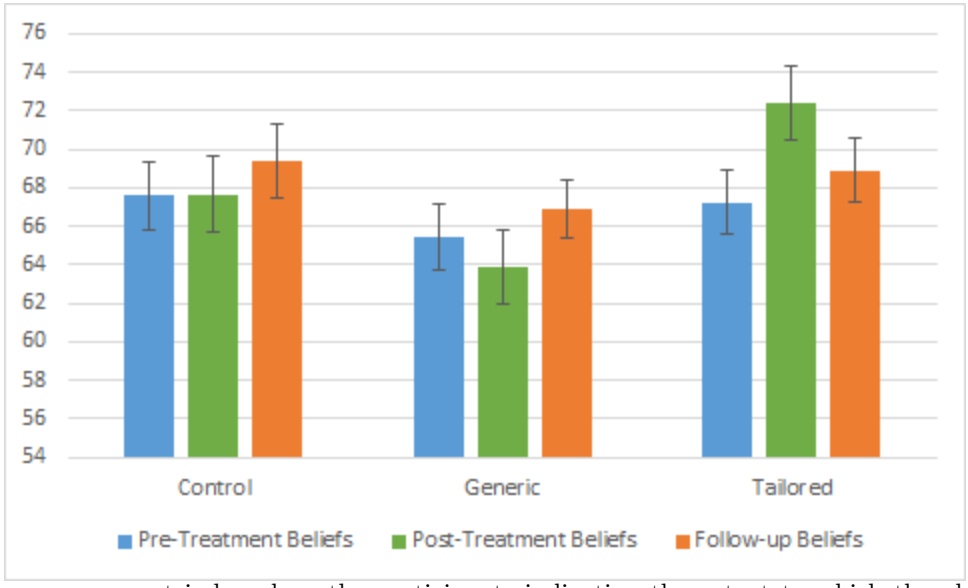
Figures and Tables

Figure 1: The Impact of the Information Treatments on Dietary Choices with various control variables



Note: Each shape per dependent variable comes from a separate regression with different sets of controls. "Demog. & BMI/Waist" includes: indicators for being male, white, married, and controls for age, BMI, and waist. "Qual. Income" includes: indicators for having a postgraduate degree, undergraduate degree, A-level qualifications, being employed, being unemployed, and a set of income categories. "Time Front page" includes: a set of indicators for when the participant's lab session took place, indicators for the first page they saw as part of the tool, and an dummy indicating whether they were hungry. "Pre-Exp. Diet" includes: an indicator for having no dietary restrictions or being a vegetarian, dummies that capture their regular diet (eating fish twice a week, eating fruit and vegetables 5 times a day, eating wholegrains 3 times a day, eating refined grains 3 times a day, eating foods high in saturated fat at least twice a day, eating foods high in transfats daily, drinking alcohol daily) and indicators for whether they currently smoke or have tried to quit smoking. "All" includes: all of the above.

Figure 2: Comparison of Pre-Treatment, Post-Treatment and Long-Run Beliefs of Living a Healthy Lifestyle by Information Group



Note: The measurement is based on the participants indicating the extent to which they believe that they are leading a healthy life? on a scale from 0 to 100 ranging from 'Very Unhealthy' to 'Very Healthy'. Pre-treatment is taken from the first session before the participants in the tailored health information treatment received any information. Post-treatment was elicited after the information at the end of the session and follow-up was measured in the laboratory sessions three months after the first session.

Table 1: Summary Statistics and Balancing

	All Participants		Information treatment			Info. p-value	Time treatment		Time p-value
	Mean	SD	Control	Generic	Tailored	All equal	Short	Long	Short=Long
	(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
Demographics/Body size									
Male	0.39	0.49	0.41	0.31	0.47	0.04	0.41	0.38	0.71
White	0.88	0.33	0.84	0.90	0.88	0.48	0.91	0.85	0.10
Single	0.70	0.46	0.68	0.75	0.66	0.32	0.67	0.72	0.39
Age	36.0	11.6	38.1	33.6	36.7	0.02	36.1	35.9	0.85
BMI	25.0	5.0	25.4	25.3	24.3	0.21	24.5	25.4	0.09
Waist	33.0	7.2	33.5	32.0	33.6	0.19	32.7	33.4	0.38
Qualifications/Employment									
Postgrad Degree	0.36	0.48	0.36	0.38	0.33	0.79	0.35	0.36	0.91
Undergraduate Degree	0.37	0.48	0.36	0.38	0.38	0.93	0.40	0.35	0.34
A-level	0.15	0.36	0.19	0.12	0.16	0.37	0.12	0.18	0.18
Employed	0.56	0.50	0.52	0.56	0.58	0.69	0.50	0.62	0.04
Unemployed	0.08	0.27	0.10	0.08	0.06	0.67	0.10	0.06	0.13
Income									
> £25,000	0.07	0.25	0.09	0.05	0.06	0.62	0.07	0.06	0.79
£20,000-25,000	0.28	0.45	0.24	0.31	0.30	0.60	0.25	0.32	0.16
£15,000-19,999	0.21	0.41	0.19	0.21	0.22	0.85	0.20	0.22	0.64
£10,000-14,999	0.22	0.41	0.24	0.23	0.19	0.55	0.23	0.21	0.72
£5,000-9,999	0.13	0.33	0.13	0.12	0.13	0.94	0.15	0.10	0.21
Experiment timing									
Time: 9.30am	0.23	0.42	0.32	0.23	0.16	0.02	0.29	0.17	0.01
Time 11.30am	0.25	0.43	0.33	0.15	0.27	0.01	0.32	0.17	0.00
Time 2.30pm	0.24	0.43	0.14	0.30	0.27	0.03	0.27	0.22	0.31
Front page of diet tool									
Meat	0.18	0.39	0.19	0.17	0.19	0.94	0.19	0.17	0.71
Bread	0.16	0.37	0.16	0.15	0.18	0.89	0.16	0.16	0.94
Confectionary	0.19	0.39	0.19	0.19	0.19	0.99	0.20	0.19	0.82
Ready meals	0.14	0.34	0.14	0.14	0.13	0.96	0.13	0.14	0.79
Drinks	0.17	0.38	0.18	0.18	0.17	0.96	0.18	0.17	0.94
Hungry at experiment	0.22	0.41	0.12	0.23	0.29	0.02	0.18	0.26	0.07
Pre-experiment Diet/Health Behaviours									
Regular Diet	0.78	0.41	0.77	0.79	0.79	0.90	0.79	0.78	0.75
Vegetarian	0.15	0.36	0.19	0.15	0.12	0.41	0.14	0.17	0.47
Fish 2/week	0.42	0.49	0.40	0.42	0.44	0.82	0.39	0.46	0.18
Fruit & Veg 5/day	0.48	0.50	0.54	0.39	0.52	0.05	0.45	0.51	0.33
Wholegrain 3/day	0.40	0.49	0.48	0.35	0.39	0.18	0.42	0.38	0.41
Refined grain 3/day	0.42	0.49	0.44	0.45	0.36	0.34	0.41	0.42	0.84
Sat fat 2/day	0.42	0.49	0.50	0.39	0.40	0.22	0.42	0.43	0.84
Trans fat daily	0.35	0.48	0.38	0.34	0.34	0.84	0.35	0.35	0.99
Drinks alcohol daily	0.18	0.39	0.16	0.17	0.21	0.55	0.18	0.18	0.94
Currently smokes	0.21	0.41	0.21	0.16	0.26	0.21	0.22	0.21	0.82
Quit smoking	0.29	0.46	0.29	0.28	0.31	0.84	0.27	0.31	0.45
N	309		90	111	108		153	156	

Note: This table presents summary statistics for the two treatment arms and the control group. Column (5) displays the p-value from a test of equal means of the three groups for the information treatment, and column 8 displays the p-value from a test of equal means of the three groups for the time treatment.

Table 2: Joint distribution of Pre-Information belief and risk from YDR

a) Heart disease

		Pre-Information Belief							
		v. much below avg.	much below avg.	below avg.	avg.	above avg.	much above avg.	v. much above avg.	Total
Risk Score	v. much below avg.	**11.3**	11.0	17.8	11.7	5.8	0.3	0.0	57.9
	much below avg.	2.3	**2.6**	4.9	7.8	3.9	0.3	0.0	21.7
	below avg.	0.3	1.3	**1.9**	3.2	2.6	1.0	0.0	10.4
	avg.	0.0	0.0	0.3	**1.0**	1.6	0.0	0.0	2.9
	above avg.	0.0	0.3	0.3	2.3	**1.3**	0.7	0.3	5.2
	much above avg.	0.0	0.0	1.0	0.3	0.3	**0.0**	0.0	1.6
	v. much above avg.	0.0	0.0	0.0	0.3	0.0	0.0	**0.0**	0.3
	Total	13.9	15.2	26.2	26.5	15.5	2.3	0.3	100.0

b) Diabetes

		Pre-Information Belief							
		v. much below avg.	much below avg.	below avg.	avg.	above avg.	much above avg.	v. much above avg.	Total
Risk Score	v. much below avg.	**11.7**	9.1	10.0	10.0	4.9	1.0	0.0	46.6
	much below avg.	2.3	**4.9**	6.5	5.2	3.6	0.7	0.3	23.3
	below avg.	0.7	1.9	**2.6**	3.9	1.9	0.0	0.3	11.3
	avg.	0.7	0.0	1.3	**1.0**	0.7	0.3	0.0	3.9
	above avg.	0.3	0.7	1.6	1.9	**2.9**	0.0	0.0	7.4
	much above avg.	0.3	0.3	0.3	1.6	2.3	**1.0**	0.0	5.8
	v. much above avg.	0.0	0.0	0.0	0.3	1.0	0.3	**0.0**	1.6
	Total	15.9	16.8	22.3	23.9	17.2	3.2	0.6	100.0

Note: Risk score is calculated from the "your disease risk tool". Pre-Information belief was elicited via the survey question: "Please indicate the extent to which you believe that you may be diagnosed with one of the following diseases over the next 10 years compared to the average person your age and gender in Scotland?".

Table 3: The Impact of the Information Treatments on Dietary Choices

	Unhealthy (%)	Calories (kcal)	Fat (g)	Saturated fat (g)	Sugar (g)	Salt (g)	Fibre (g)	Protein (g)	Spend F&V (%)
Generic	-0.0459**	-362.0	-43.81**	-18.92**	-0.415	-3.193**	9.470	-10.23	0.00655
	(0.0231)	(449.1)	(19.08)	(8.960)	(36.29)	(1.401)	(10.06)	(21.78)	(0.0209)
Wild cluster p-val	*0.022*	*0.542*	*0.025*	*0.081*	*0.993*	*0.036*	*0.438*	*0.661*	*0.792*
RI p-val	*0.048*	*0.425*	*0.025*	*0.040*	*0.991*	*0.020*	*0.346*	*0.645*	*0.758*
FDR q-val		*0.674*	*0.094*	*0.095*	*0.991*	*0.094*	*0.674*	*0.852*	*0.862*
FWER p-val		*0.855*	*0.126*	*0.165*	*0.992*	*0.127*	*0.819*	*0.944*	*0.944*
Tailored	-0.0062	-815.2*	-26.97	-9.809	-3.836	-1.284	-7.855	-7.060	0.0146
	(0.0269)	(460.3)	(20.83)	(10.25)	(38.95)	(1.644)	(10.95)	(23.86)	(0.0241)
Wild cluster p-val	*0.849*	*0.090*	*0.229*	*0.402*	*0.934*	*0.511*	*0.353*	*0.749*	*0.582*
RI p-val	*0.814*	*0.079*	*0.202*	*0.341*	*0.920*	*0.446*	*0.473*	*0.766*	*0.536*
FDR q-val		*0.621*	*0.727*	*0.727*	*0.922*	*0.727*	*0.727*	*0.878*	*0.727*
FWER p-val		*0.357*	*0.667*	*0.870*	*0.944*	*0.912*	*0.912*	*0.944*	*0.912*
G = T (p-val)	*0.0574*	*0.283*	*0.239*	*0.167*	*0.922*	*0.279*	*0.069*	*0.881*	*0.684*
G = T = 0 (p-val)	*0.0498*	*0.204*	*0.060*	*0.063*	*0.994*	*0.076*	*0.184*	*0.895*	*0.830*
R-squared	0.087	0.094	0.110	0.097	0.037	0.050	0.020	0.075	0.097
Mean (control)	0.201	10,858	262.3	92.56	645.6	21.55	184.2	481.2	0.346

Note: Observations for all columns equal to 309. The dependent variables are based on the totals of the basket. All regressions include controls, these include age, being male, a set of indicators for the time of the experimental session and a dummy indicating whether the participant was hungry or not. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019). FDR q-val calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). FWER correct p-value refers to the p-value using the step-down methods of Romano and Wolf (2005a,b). All p-values are calculated using 5000 replications. An item is classified as 'unhealthy' is based on the UK Food Standards Agency's nutrient profiling technique. Points for each item are allocated on the basis of the nutrient content of 100g of a food or drink. Points are awarded for energy, saturated fat, total sugar and sodium (A-nutrients), and for fruit, vegetables and nut content, fibre and protein (C-nutrients). The points from C-nutrients are then subtracted from the score for A-nutrients to calculate a final score. Unhealthy items are foods with 4 or more points and drinks with 1 or more points.

Table 4: The Impact of Information on Dietary Choices with the Tailored Treatment Separated by Good and Not Good News

	Unhealthy (%)	Calories (kcal)	Fat (g)	Saturated fat (g)	Sugar (g)	Salt (g)	Fibre (g)	Protein (g)	Spend F&V (%)
Generic	-0.0457**	-358.1	-43.72**	-18.95**	0.204	-3.206**	9.300	-10.06	0.00610
	(0.0231)	(450.1)	(19.10)	(8.955)	(36.36)	(1.410)	(10.07)	(21.83)	(0.0209)
Wild cluster p-val	*0.022*	*0.544*	*0.025*	*0.078*	*0.996*	*0.036*	*0.450*	*0.670*	*0.803*
RI p-val	*0.049*	*0.431*	*0.025*	*0.039*	*0.997*	*0.020*	*0.355*	*0.652*	*0.775*
FDR q-val		*0.684*	*0.094*	*0.094*	*0.996*	*0.094*	*0.684*	*0.860*	*0.882*
FWER p-val		*0.860*	*0.127*	*0.163*	*0.996*	*0.128*	*0.829*	*0.947*	*0.947*
Tailored: Good message	-0.00839	-873.6*	-28.30	-9.419	-13.01	-1.089	-5.338	-9.516	0.0212
	(0.0295)	(485.7)	(22.12)	(11.12)	(41.66)	(1.960)	(11.83)	(25.66)	(0.0265)
Wild cluster p-val	*0.794*	*0.055*	*0.226*	*0.446*	*0.802*	*0.616*	*0.519*	*0.732*	*0.481*
RI p-val	*0.768*	*0.073*	*0.209*	*0.413*	*0.753*	*0.586*	*0.654*	*0.710*	*0.426*
FDR q-val		*0.584*	*0.755*	*0.755*	*0.755*	*0.755*	*0.755*	*0.755*	*0.755*
FWER p-val		*0.343*	*0.685*	*0.918*	*0.951*	*0.951*	*0.951*	*0.951*	*0.918*
Tailored: Not good message	-0.000331	-662.3	-23.48	-10.83	20.21	-1.794	-14.45	-0.625	-0.00278
	(0.0363)	(735.5)	(27.09)	(11.75)	(60.35)	(2.051)	(15.87)	(35.64)	(0.0330)
Wild cluster p-val	*0.993*	*0.536*	*0.310*	*0.276*	*0.733*	*0.457*	*0.417*	*0.981*	*0.902*
RI p-val	*0.994*	*0.379*	*0.399*	*0.370*	*0.736*	*0.409*	*0.376*	*0.986*	*0.938*
FDR p-val		*0.620*	*0.620*	*0.620*	*0.984*	*0.620*	*0.620*	*0.986*	*0.986*
FWER p-val		*0.894*	*0.894*	*0.894*	*0.977*	*0.894*	*0.894*	*0.994*	*0.994*
Gen. vs Good News	*0.116*	*0.259*	*0.333*	*0.210*	*0.731*	*0.325*	*0.166*	*0.982*	*0.506*
Gen. vs Not Good News	*0.168*	*0.663*	*0.373*	*0.388*	*0.722*	*0.471*	*0.112*	*0.775*	*0.766*
Good News vs Not Good News	*0.826*	*0.773*	*0.845*	*0.894*	*0.585*	*0.777*	*0.568*	*0.804*	*0.478*
Joint all	*0.111*	*0.334*	*0.132*	*0.134*	*0.956*	*0.162*	*0.302*	*0.962*	*0.846*
R-squared	0.087	0.095	0.110	0.097	0.039	0.050	0.021	0.076	0.098
Mean (control)	0.201	10,858	262.3	92.56	645.6	21.55	184.2	481.2	0.346

Note: Observations for all columns equal to 309. All regressions include controls that are described in the notes to table 3. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019). FDR q-val calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). FWER correct p-value refers to the p-value using the step-down methods of Romano and Wolf (2005a,b). All p-values are calculated using 5000 replications. All p-values are calculated using 5000 replications. Good, and not good are messages based on tailored health information that was provided by the your disease risk tool and the prior beliefs of developing the diseases. A good heart disease message is one where the tool provided a lower risk than the individual selected in the survey before using the health information tool. We categorise "Good" news (71.3%) for those who were given good news for both heart disease and diabetes, or received good news for one disease and expected news (the individual's prior about their health risk was the same as the risk provided by the tool) for the other. The "not good" news is made up of "Mixed" news (14.8%) which includes those who received good news and bad news, or expected and bad news. "Expected" (6.5%) and "Bad" (7.4%) news are where both pieces of news were in the respective category. See notes to table 3 for classification of unhealthy items.

Table 5: The Effect of the Information Treatments on Changes in Lifestyle Rating & Changes in Risk Beliefs

	Δ Lifestyle	Δ Heart Risk	Δ Diabetes Risk	New Information
Generic	-2.078	-0.0807	-0.150	0.0601
	(1.773)	(0.130)	(0.152)	(0.0574)
Wild cluster p-val	*0.184*	*0.598*	*0.292*	*0.363*
RI p-val	*0.241*	*0.536*	*0.320*	*0.301*
FDR q-val	*0.384*	*0.513*	*0.384*	*0.384*
FWER p-val	*0.633*	*0.633*	*0.633*	*0.633*
Tailored	4.885***	-1.099***	-0.486**	0.105*
	(1.706)	(0.149)	(0.193)	(0.0604)
Wild cluster p-val	*0.005*	*0.000*	*0.026*	*0.208*
RI p-val	*0.004*	*0.000*	*0.013*	*0.088*
FDR q-val	*0.002*	*0.001*	*0.034*	*0.126*
FWER p-val	*0.012*	*0.000*	*0.028*	*0.083*
G = T (p-val)	*0.000*	*0.000*	*0.087*	*0.421*
G = T = 0 (p-val)	*0.000*	*0.000*	*0.044*	*0.222*
R-squared	0.065	0.210	0.038	0.046
Mean (control)	68.2	3.17	3.19	0.167

Note: Observations for all columns equal to 309. All regressions include controls that are described in the notes to table 3. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019). FDR q-val calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). FWER correct p-value refers to the p-value using the step-down methods of Romano and Wolf (2005a,b). All p-values are calculated using 5000 replications. All p-values are calculated using 5000 replications. The dependent variable in columns (1) - (3) are the change in lifestyle, heart risk and diabetes risk after the information treatment compared to before the information treatment. Column (4) is a dummy variable that takes a one if the participant indicated that the information provided was new, and zero if they indicated it was not new. The row G vs T is the *p-value* of the test of the difference of the parameters of generic info and tailored info. Mean (control) shows the mean of variable prior to the information being given (or reading the architecture article in the case of the control group) for the lifestyle, heart risk and diabetes risk variables. For "new info" mean (control) is just the mean of the dependent variable of the control group.

Table 6: The Impact of the Information Treatments breaking the Treatments by old and new information

	Unhealthy (%)	Calories (kcal)	Fat (g)	Saturated fat (g)	Sugar (g)	Salt (g)	Fibre (g)	Protein (g)	Spend F&V (%)
Generic (Old)	-0.0505**	-387.5	-49.60**	-20.15**	1.833	-4.645***	13.37	-25.75	0.0212
	(0.0248)	(478.8)	(20.16)	(9.668)	(38.02)	(1.443)	(10.68)	(22.86)	(0.0216)
Wild cluster p-val	*0.045*	*0.537*	*0.010*	*0.070*	*0.967*	*0.006*	*0.341*	*0.341*	*0.359*
RI p-val	*0.041*	*0.427*	*0.016*	*0.036*	*0.964*	*0.001*	*0.210*	*0.267*	*0.328*
FDR q-val		*0.479*	*0.058*	*0.102*	*0.962*	*0.012*	*0.418*	*0.418*	*0.438*
FWER p-val		*0.662*	*0.075*	*0.174*	*0.964*	*0.007*	*0.619*	*0.650*	*0.662*
Generic (New)	-0.0296	-324.5	-25.72	-15.83	6.751	1.979	-4.076	35.21	-0.0402
	(0.0299)	(715.8)	(25.20)	(11.22)	(62.09)	(2.820)	(14.06)	(34.15)	(0.0296)
Wild cluster p-val	*0.345*	*0.635*	*0.373*	*0.217*	*0.920*	*0.512*	*0.721*	*0.140*	*0.208*
RI p-val	*0.345*	*0.649*	*0.330*	*0.177*	*0.917*	*0.506*	*0.779*	*0.313*	*0.185*
FDR q-val		*0.868*	*0.616*	*0.616*	*0.913*	*0.773*	*0.883*	*0.616*	*0.616*
FWER p-val		*0.949*	*0.801*	*0.607*	*0.951*	*0.898*	*0.951*	*0.801*	*0.630*
Tailored (Old)	-0.0111	-944.9*	-37.76	-13.58	28.14	-2.757*	-3.218	-42.05*	0.0397
	(0.0305)	(492.1)	(23.16)	(12.07)	(41.86)	(1.499)	(11.54)	(24.91)	(0.0264)
Wild cluster p-val	*0.745*	*0.107*	*0.135*	*0.287*	*0.625*	*0.104*	*0.730*	*0.111*	*0.185*
RI p-val	*0.720*	*0.054*	*0.111*	*0.274*	*0.501*	*0.065*	*0.779*	*0.095*	*0.131*
FDR q-val		*0.208*	*0.208*	*0.348*	*0.574*	*0.208*	*0.781*	*0.208*	*0.213*
FWER p-val		*0.281*	*0.352*	*0.564*	*0.728*	*0.307*	*0.781*	*0.352*	*0.380*
Tailored (New)	0.00717	-462.2	2.293	0.452	-91.05	2.688	-20.38	87.93**	-0.0535
	(0.0388)	(744.9)	(26.94)	(11.87)	(61.95)	(4.296)	(17.81)	(34.30)	(0.0335)
Wild cluster p-val	*0.833*	*0.674*	*0.918*	*0.968*	*0.064*	*0.650*	*0.325*	*0.022*	*0.013*
RI p-val	*0.851*	*0.531*	*0.937*	*0.970*	*0.149*	*0.558*	*0.269*	*0.014*	*0.114*
FDR q-val		*0.714*	*0.970*	*0.970*	*0.382*	*0.714*	*0.506*	*0.088*	*0.382*
FWER p-val		*0.886*	*0.989*	*0.989*	*0.452*	*0.886*	*0.621*	*0.069*	*0.410*
Old G v New G (p-val)	0.460	0.930	0.292	0.678	0.936	0.018	0.201	0.066	0.024
Old G v New T (p-val)	0.151	0.922	0.047	0.089	0.137	0.119	0.059	0.001	0.026
New G v Old T (p-val)	0.556	0.389	0.630	0.856	0.733	0.095	0.950	0.025	0.008
Old T v New T (p-val)	0.668	0.526	0.152	0.318	0.063	0.224	0.337	0.000	0.009
R-squared	0.088	0.096	0.119	0.101	0.052	0.083	0.027	0.128	0.129

Note: Observations for all columns equal to 309. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions include controls that are described in the notes to table 3. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019). FDR q-val calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). FWER correct p-value refers to the p-value using the step-down methods of Romano and Wolf (2005a,b). All p-values are calculated using 5000 replications. (Old) are those in the treatment groups who indicated after the treatment that the information was not new, (new) are those who indicated that the information was new. There are 90 participants in the Generic (Old) group, 21 in the Generic New group, 80 in the Tailored (Old) group, 28 in the Tailored (New) group, and 90 in the control group.

Table 7: The Impact of the Time Availability treatment on Dietary Choices

	Unhealthy (%)	Calories (kcal)	Fat (g)	Saturated fat (g)	Sugar (g)	Salt (g)	Fibre (g)	Protein (g)	Spend F&V (%)
Long time	-0.00886	-52.23	-4.273	-3.284	30.81	0.271	1.393	2.409	0.00466
	(0.0202)	(394.1)	(15.02)	(6.894)	(32.43)	(1.430)	(8.801)	(18.37)	(0.0182)
Wild cluster p-val	*0.674*	*0.915*	*0.769*	*0.622*	*0.220*	*0.878*	*0.909*	*0.900*	*0.780*
RI p-val	*0.662*	*0.906*	*0.781*	*0.645*	*0.341*	*0.845*	*0.878*	*0.896*	*0.797*
FDR p-val		*0.896*	*0.896*	*0.896*	*0.896*	*0.896*	*0.896*	*0.896*	*0.896*
FWER p-val		*0.999*	*0.999*	*0.992*	*0.892*	*0.999*	*0.999*	*0.999*	*0.999*
Mean (short time)	0.193	10679	246	86.13	631.4	20.27	183.6	482.9	0.346
R-squared	0.073	0.084	0.092	0.082	0.040	0.038	0.010	0.075	0.096

Note: Observations for all columns equal to 309. All regressions include controls, these include age, being male, a set of indicators for the time of the experimental session and a dummy indicating whether the participant was hungry or not. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019). FDR q-val calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). FWER correct p-value refers to the p-value using the step-down methods of Romano and Wolf (2005a,b). All p-values are calculated using 5000 replications. An item is classified as 'unhealthy' is based on the UK Food Standards Agency's nutrient profiling technique. Points for each item are allocated on the basis of the nutrient content of 100g of a food or drink. Points are awarded for energy, saturated fat, total sugar and sodium (A-nutrients), and for fruit, vegetables and nut content, fibre and protein (C-nutrients). The points from C-nutrients are then subtracted from the score for A-nutrients to calculate a final score. Unhealthy items are foods with 4 or more points and drinks with 1 or more points.

Table 8: The Impact of the Information Treatments on Dietary Choices 3 Months Later

	Unhealthy (%)	Calories (kcal)	Fat (g)	Saturated fat (g)	Sugar (g)	Salt (g)	Fibre (g)	Protein (g)	Spend F&V (%)
Generic	-0.00493	-464.4	-6.880	-5.776	-41.64	-1.394	-13.07	14.03	-0.00250
	(0.0230)	(445.3)	(16.39)	(7.390)	(39.53)	(1.354)	(10.07)	(22.73)	(0.0270)
Wild cluster p-val	*0.790*	*0.282*	*0.602*	*0.303*	*0.293*	*0.217*	*0.232*	*0.542*	*0.927*
RI p-val	*0.825*	*0.300*	*0.685*	*0.440*	*0.305*	*0.319*	*0.198*	*0.546*	*0.925*
FDR q-val		*0.608*	*0.772*	*0.696*	*0.608*	*0.608*	*0.608*	*0.718*	*0.926*
FWER p-val		*0.791*	*0.851*	*0.791*	*0.851*	*0.791*	*0.791*	*0.851*	*0.911*
Tailored	0.00216	-1,214***	-11.50	-5.149	-40.93	-1.954	-24.27**	-5.307	0.00758
	(0.0251)	(460.2)	(16.59)	(7.081)	(42.15)	(1.439)	(10.00)	(24.81)	(0.0263)
Wild cluster p-val	*0.919*	*0.017*	*0.465*	*0.439*	*0.148*	*0.202*	*0.015*	*0.842*	*0.774*
RI p-val	*0.929*	*0.009*	*0.486*	*0.471*	*0.334*	*0.176*	*0.017*	*0.832*	*0.766*
FDR q-val		*0.064*	*0.652*	*0.652*	*0.652*	*0.470*	*0.064*	*0.831*	*0.831*
FWER p-val		*0.053*	*0.742*	*0.765*	*0.838*	*0.520*	*0.140*	*0.900*	*0.900*
G = T (p-val)	*0.741*	*0.0748*	*0.762*	*0.922*	*0.985*	*0.655*	*0.259*	*0.405*	*0.672*
G = T = 0 (p-val)	*0.942*	*0.0276*	*0.786*	*0.702*	*0.522*	*0.385*	*0.0545*	*0.676*	*0.908*
R-squared	0.078	0.090	0.078	0.087	0.042	0.056	0.046	0.050	0.078

Note: Observations for all columns equal to 256. All regressions include controls, these include age, being male, a set of indicators for the time of the experimental session and a dummy indicating whether the participant was hungry or not. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019). FDR q-val calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). FWER correct p-value refers to the p-value using the step-down methods of Romano and Wolf (2005a,b). All p-values are calculated using 5000 replications. An item is classified as 'unhealthy' is based on the UK Food Standards Agency's nutrient profiling technique. Points for each item are allocated on the basis of the nutrient content of 100g of a food or drink. Points are awarded for energy, saturated fat, total sugar and sodium (A-nutrients), and for fruit, vegetables and nut content, fibre and protein (C-nutrients). The points from C-nutrients are then subtracted from the score for A-nutrients to calculate a final score. Unhealthy items are foods with 4 or more points and drinks with 1 or more points.

Table 9: The Impact of the Information Treatments on Self-Reported Health Behaviours and Health Measurements 3 Months Later

	Any	Stop Smoking	More Exercise	Reduce Alcohol	Lose weight	Take vitamin	More Fish	More F&V	More Grains	Reduced Sat. Fat	Reduced Transfat	% Δ BMI	% Δ waist
Generic	0.108*	-0.00319	0.00203	0.0701	0.0184	0.00336	-0.0620	-0.0202	-0.105*	0.0831	0.0193	0.00284	-0.00917
	(0.0583)	(0.0349)	(0.0795)	(0.0796)	(0.0706)	(0.0622)	(0.0721)	(0.0741)	(0.0620)	(0.0760)	(0.0815)	(0.0111)	(0.0224)
Wild cluster p-val	*0.0298*	*0.904*	*0.976*	*0.260*	*0.842*	*0.938*	*0.174*	*0.836*	*0.106*	*0.188*	*0.825*	*0.827*	*0.735*
RI p-val	*0.0712*	*0.934*	*0.977*	*0.384*	*0.791*	*0.953*	*0.375*	*0.782*	*0.0942*	*0.277*	*0.805*	*0.800*	*0.723*
FDR q-val		*0.980*	*0.980*	*0.978*	*0.980*	*0.980*	*0.978*	*0.980*	*0.905*	*0.978*	*0.980*		
FWER p-val		*0.995*	*0.995*	*0.991*	*0.993*	*0.995*	*0.995*	*0.995*	*0.433*	*0.963*	*0.993*		
Tailored	0.0239	-0.0177	-0.0978	0.0652	-0.0351	0.00137	-0.120*	-0.0443	-0.0760	0.0317	-0.0731	0.00650	0.000508
	(0.0643)	(0.0328)	(0.0774)	(0.0760)	(0.0663)	(0.0598)	(0.0702)	(0.0721)	(0.0588)	(0.0742)	(0.0787)	(0.0105)	(0.0253)
Wild cluster p-val	*0.580*	*0.384*	*0.153*	*0.413*	*0.536*	*0.981*	*0.0600*	*0.606*	*0.135*	*0.679*	*0.443*	*0.630*	*0.987*
RI p-val	*0.706*	*0.609*	*0.195*	*0.387*	*0.612*	*0.979*	*0.0844*	*0.536*	*0.193*	*0.663*	*0.343*	*0.548*	*0.987*
FDR q-val		*0.744*	*0.69*	*0.744*	*0.744*	*0.982*	*0.69*	*0.744*	*0.69*	*0.744*	*0.744*		
FWER p-val		*0.993*	*0.767*	*0.983*	*0.993*	*0.993*	*0.837*	*0.983*	*0.837*	*0.993*	*0.981*		
G = T (p-val)	*0.0860*	*0.623*	*0.168*	*0.948*	*0.396*	*0.973*	*0.368*	*0.729*	*0.568*	*0.483*	*0.221*	*0.561*	*0.530*
G = T = 0 (p-val)	*0.0852*	*0.823*	*0.296*	*0.613*	*0.682*	*0.999*	*0.229*	*0.825*	*0.230*	*0.543*	*0.431*	*0.742*	*0.779*
R-squared	0.033	0.073	0.026	0.023	0.012	0.008	0.043	0.013	0.023	0.019	0.016	0.048	0.020

Note: Observations for all columns equal to 256. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions include controls that are described in the notes to table 3. All regressions include controls that are described in the notes to table 3. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019). FDR q-val calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). FWER correct p-value refers to the p-value using the step-down methods of Romano and Wolf (2005a,b). All p-values are calculated using 5000 replications. All p-values are calculated using 5000 replications. The dependent variables are based on answers to the following question: Which, if any, of the following changes have you made over the last 3 months (i.e. since the initial session)? i) Stop smoking (Stop Smoking), ii) Do some form of moderate physical exercise for at least 30 minutes on most days (More Exercise), iii) Cut down the amount of alcohol I drink (Reduce Alcohol), iv) Lose some weight (Lose weight), v) Take a multivitamin - like a B complex vitamin (Take vitamin), vi) Eat 2 or more servings of fish per week (More Fish), vii) Eat 5 or more servings of fruit and veg per day (More F&V), viii) Eat 3 or more servings of whole grains on most days (More Grains), ix) Reduce consumption of saturated fats like red meat, cheese and whole milk (Reduced Sat. Fat), x) Reduce consumption of trans-saturated fats like cookies, pies, chips, crisps and deep-fried food (Reduced Transfat).

ONLINE APPENDIX

Supplementary Material

Facilitating Healthy Dietary Habits: An Experiment with a Low Income Population

Michèle Belot[1], Jonathan James[2], and Jonathan Spiteri[3]

[1]European University Institute
[2]Department of Economics, University of Bath
[3]Department of Economics, University of Malta

This Appendix reports additional analysis and results discussed in the main text, which could not be included due to space concerns. To locate the material more easily in the context of the paper, in what follows we refer to the specific sections of the text. We also provide additional material. We begin by providing information where the analysis in the paper deviates from the pre-analysis plan.

0. Deviations from pre-analysis plan[1]

Changes to the experimental design

Treatments

The PAP uses different labels for the treatment varying the time availability (high and low priority) rather than in the paper (short and long time).

Information intervention

We announced that our control group would receive news taken from local and international media outlets, but when piloting those, we realized it was hard to find emotionally neutral news. We therefore opted for an article not related to health, on architecture.

Time intervention

[1]The pre-analysis plan can be found here `https://doi.org/10.1257/rct.1189-1.0`

The pre-analysis plan announces that the shopping tool will include 100 food items, while we ended up including 120 food and drink items.

Incentives

The pre-analysis plan announces that at the end of each session, 1 subject per session will be picked at random and his/her food basket will be delivered to his/her home address within one week. To ensure that their choices were not contaminated by their current food and drink stocks at home, we decided to postpone the delivery date to at least two weeks after participation.

Randomization

We announced that the treatment would be randomized at the session level. To ensure maximum balance across days and times, we opted for pre-allocating the time slots to specific treatments ourselves.

Sample

We intended to have 300 participants, we ended up with a sample of 309

Changes to the Analysis

Due to of concerns related to possible imbalances in characteristics of participants across sessions, we decided to include the following controls in all regressions: Age, Gender, a set of indicators for the time of the experimental session and dummy indicating whether the participant was hungry or not. Adding these controls does not matter for the results (significance and magnitude of coefficients see Figure 1 in the main text).

Tables 4, 5 and 6 were not specified in the pre-analysis plan and are carried out in order to help understand the main findings.

Long term analysis

Long term analysis: We announced the following outcome variables:

1. Number of fruit and vegetables;

2. Total nutrient content of the participants' daily dietary intake, where the nutrients considered are calories, carbohydrates, fats, saturated fats, sugar, salt, fibre and protein (a separate total will be calculated for each nutrient).

We decided to report the percentage spent on fruit and vegetables instead of the number, to be consistent with the short-tern analysis.

We also announced that we would be looking at self-reported dietary or lifestyle changes that the participants will undertake in the 3 months after the initial intervention. We reported the follow-up questionnaire contained two questions related to this matter. The first question states: "Looking back over the last 3 months, have you made any changes to your diet or lifestyle habits?", which elicits a simple 'Yes' or 'No'. The second question follows-up on this point: "If yes, please indicate the change in your diet or lifestyle that you have undertaken from the list below", which then lists the following options: "Stop smoking, Do more exercise, Eat more fruit and vegetables, Eat less junk food and processed foods, Eat less sugar, Eat less red meat, Take vitamins and other supplements, Drink less alcohol". Therefore, we can use this data to obtain two outcome variables:

The options asked were not exactly as stated in the pre-analysis plan. Instead we asked whether participants had made any of the following changes to their lifestyle:

- Do some form of moderate physical exercise for at least 30 minutes on most days
- Stop smoking
- Cut down the amount of alcohol I drink
- Lose some weight
- Take a multivitamin (like a B complex vitamin)
- Eat 2 or more servings of fish per week
- Eat 5 or more servings of fruit and veg per day
- Eat 3 or more servings of whole grains on most days
- Reduce consumption of saturated fats like red meat, cheese and whole milk
- Reduce consumption of trans-saturated fats like cookies, pies, chips, crisps and deep-fried food

In addition, we have decided to also report the changes in BMI and waist circumference.

Power calculations

Power calculations were done before the experiment and analysis, but we did not include them in the pre-analysis plan. See section III.G in the main text for a discussion of the power analysis.

III. Information Intervention

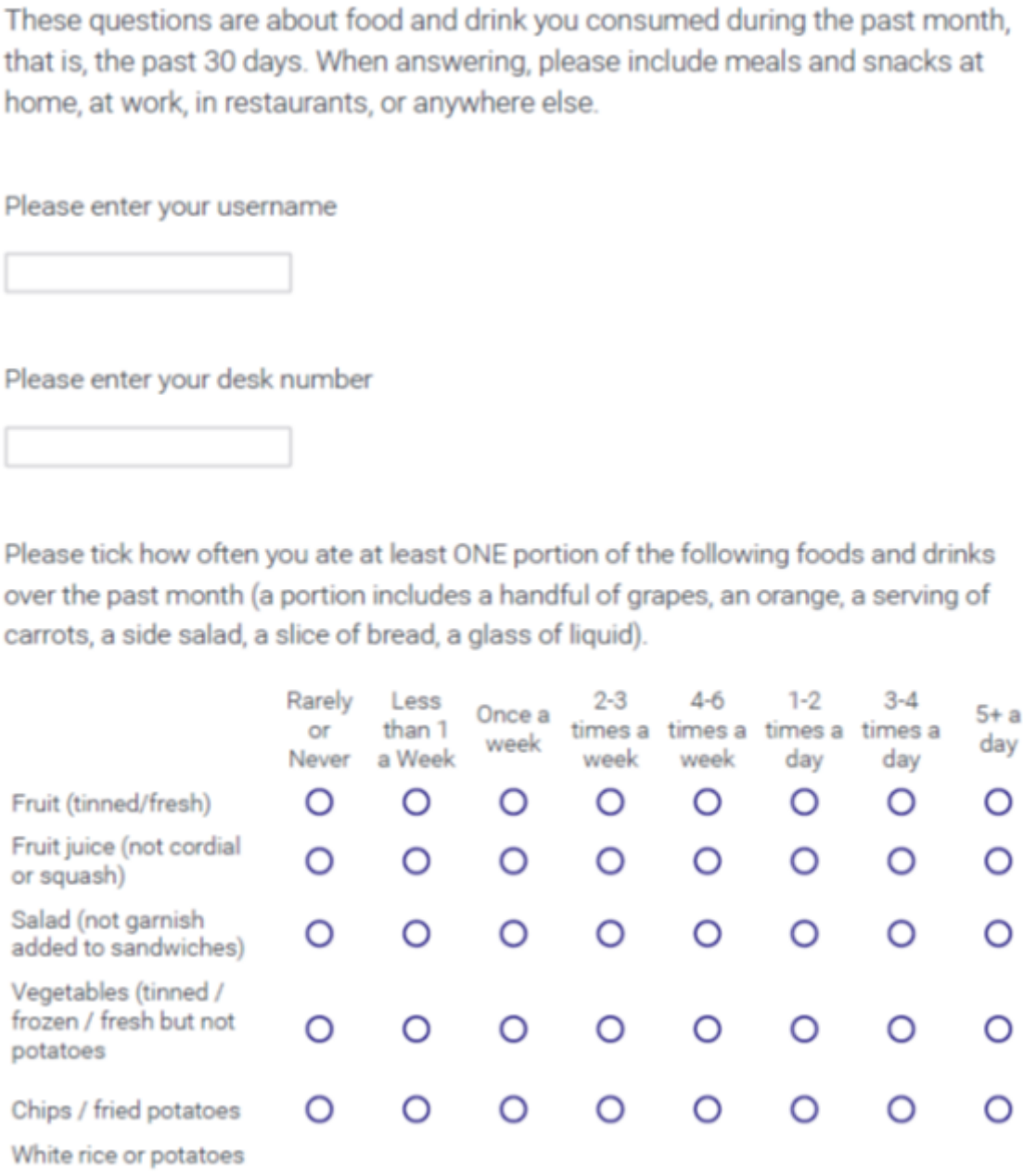
III.B. Procedure

Table 1: Allocation of treatments to sessions

Day	**Time**	**Treatments**
Monday	9.30am	Control - High
Tuesday	11.30am	Control - High
Wednesday	5.30pm	Control - High
Wednesday	11.30am	Control - Low
Thursday	2.30pm	Control - Low
Friday	9.30am	Control - Low
Monday	11.30am	Generic - High
Wednesday	9.30am	Generic - High
Thursday	5.30pm	Generic - High
Friday	2.30pm	Generic - High
Monday	5.30pm	Generic - Low
Tuesday	2.30pm	Generic - Low
Thursday	9.30am	Generic - Low
Monday	2.30pm	Tailored - High
Tuesday	5.30pm	Tailored - High
Friday	11.30am	Tailored - High
Tuesday	9.30am	Tailored - Low
Wednesday	2.30pm	Tailored - Low
Thursday	11.30am	Tailored - Low
Friday	5.30pm	Tailored - Low

Food Frequency Questionnaire

Figure A.1: Food Frequency Questionnaire



These questions are about food and drink you consumed during the past month, that is, the past 30 days. When answering, please include meals and snacks at home, at work, in restaurants, or anywhere else.

Please enter your username

Please enter your desk number

Please tick how often you ate at least ONE portion of the following foods and drinks over the past month (a portion includes a handful of grapes, an orange, a serving of carrots, a side salad, a slice of bread, a glass of liquid).

	Rarely or Never	Less than 1 a Week	Once a week	2-3 times a week	4-6 times a week	1-2 times a day	3-4 times a day	5+ a day
Fruit (tinned/fresh)								
Fruit juice (not cordial or squash)								
Salad (not garnish added to sandwiches)								
Vegetables (tinned / frozen / fresh but not potatoes								
Chips / fried potatoes								
White rice or potatoes								

like baked, boiled, mashed or sweet potatoes	○	○	○	○	○	○	○	○
Beans or pulses like baked beans, lentils, chick peas, dahl	○	○	○	○	○	○	○	○
Tea or coffee sweetened with sugar or honey, including iced tea and frappuccino	○	○	○	○	○	○	○	○

Please tick how often you ate at least ONE portion of the following foods and drinks over the past month (a portion includes a 30g bowl of cereal, 25g of cheese, a packet of crisps, 2 biscuits, a slice of bread, a glass of liquid).

	Rarely or Never	Less than 1 a Week	Once a week	2-3 times a week	4-6 times a week	1-2 times a day	3-4 times a day	5+ a day
Fibre-rich breakfast cereal like Weetabix, Fruit 'n Fibre, Porridge, Muesli	○	○	○	○	○	○	○	○
Wholemeal bread or chapattis	○	○	○	○	○	○	○	○
Brown rice / bulgur wheat / quinoa	○	○	○	○	○	○	○	○
Cheese / yoghurt	○	○	○	○	○	○	○	○
Crisps / savoury snacks	○	○	○	○	○	○	○	○
Sweet biscuits, cakes, chocolate, sweets	○	○	○	○	○	○	○	○
Ice cream / cream	○	○	○	○	○	○	○	○
Non alcoholic fizzy drinks / pop (not sugar free or diet)	○	○	○	○	○	○	○	○

Please tick how often you ate at least ONE portion of the following foods and drinks over the past month.

	Rarely or Never	Less than 1 a Week	Once a week	2-3 times a week	4-6 times a week	At least everyday
Beef, Lamb, Pork, Ham - steaks, roasts, joints,	○	○	○	○	○	○

mince or chops						
Chicken or Turkey - steaks, roasts, joints, mince or portions (not in batter or breadcrumbs)	○	○	○	○	○	○
Sausages, bacon, corned beef, meat pies/pasties, burgers	○	○	○	○	○	○
Chicken/turkey nuggets. turkey burgers, chicken pies, or in batter or breadcrumbs	○	○	○	○	○	○
White fish in batter or breadcrumbs - like fish 'n chips	○	○	○	○	○	○
White fish not in batter or breadcrumbs	○	○	○	○	○	○
Oily fish - like herrings, trout, salmon, sardines, mackerel, fresh tuna (not tinned tuna)	○	○	○	○	○	○

What milk do you USUALLY use or drink, such as in hot & cold drinks or on cereal (including tea, coffee, hot milk, milkshakes, or on cereal)

Whole / full-fat milk

Semi-skimmed milk

Skimmed Milk

Rarely / never use milk

Other (please specify)

III.C. Information Intervention

Generic Information Document and Sample Tailored Information Results (for Heart Disease)

Figure A.2A: Generic Information

Heart Disease

The average person in Scotland over the age of 18 has a 7% chance of contracting coronary heart disease over the next 10 years of his or her life.

Coronary heart disease (CHD) occurs when the flow of oxygen-rich blood to your heart is blocked or reduced by a build-up of fatty material (atheroma) in the coronary arteries.

As they narrow because of a build-up of atheroma, the blood supply to your heart will be restricted. This can cause angina (chest pain). If a coronary artery becomes completely blocked, it can cause a heart attack.

To reduce the risk of contracting heart disease, the following recommendations are typically suggested by medical experts:

Do some form of moderate physical exercise for at least 30 minutes on most days

Stop smoking

Control one's weight in order to maintain a healthy body-mass index (BMI)

Take a multivitamin (like a B complex vitamin)

Eat 2 or more servings of fish per week

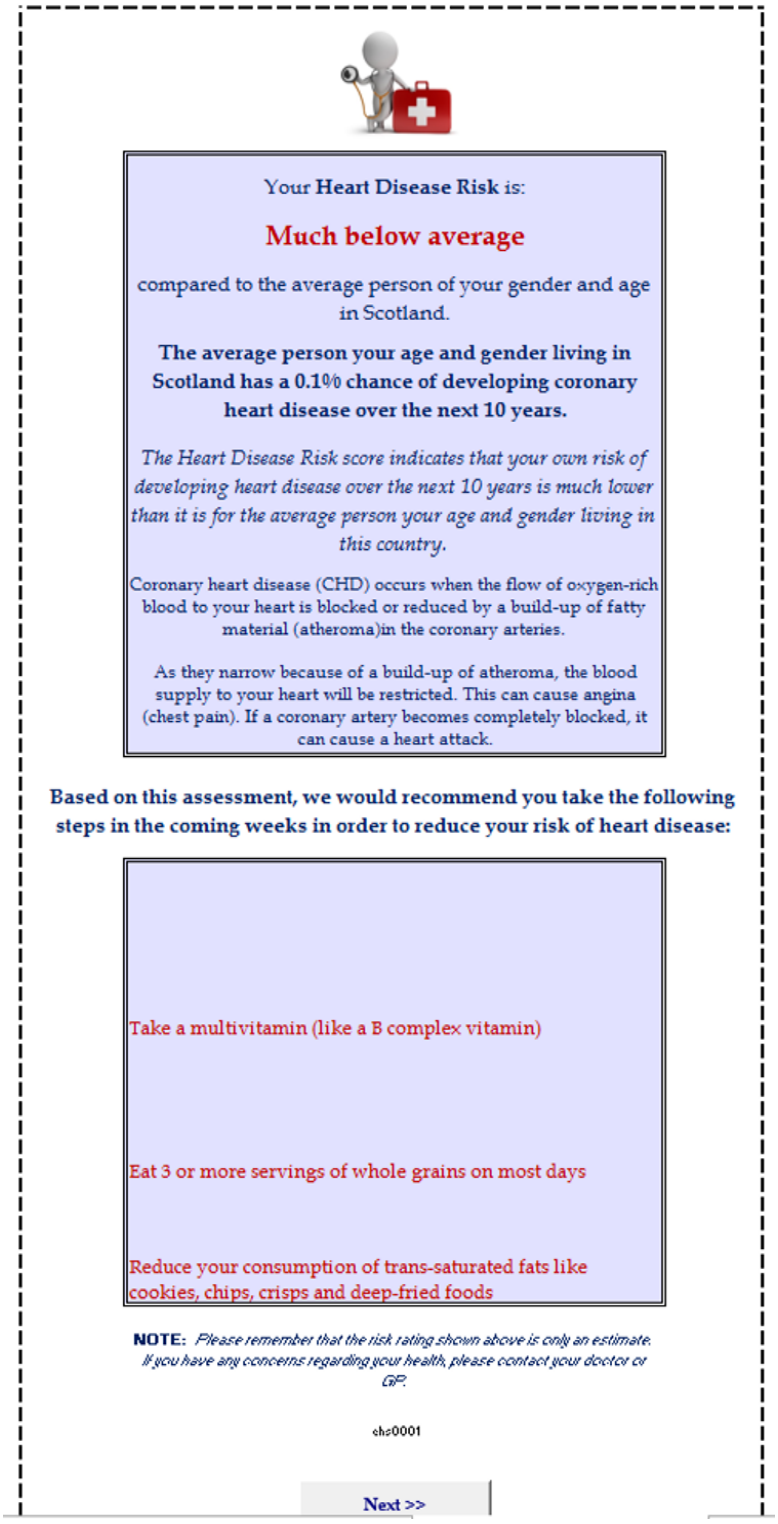
Eat 5 or more servings of fruit and vegetables per day

Eat 3 or more servings of whole grains on most days

Reduce consumption of saturated fats like red meat, cheese and whole milk

Reduce consumption of trans-saturated fats like cookies, chips, crisps and deep-fried foods

Figure A.2B: Tailored Information Results (for Heart Disease)



Your **Heart Disease Risk** is:

Much below average

compared to the average person of your gender and age in Scotland.

The average person your age and gender living in Scotland has a 0.1% chance of developing coronary heart disease over the next 10 years.

The Heart Disease Risk score indicates that your own risk of developing heart disease over the next 10 years is much lower than it is for the average person your age and gender living in this country.

Coronary heart disease (CHD) occurs when the flow of oxygen-rich blood to your heart is blocked or reduced by a build-up of fatty material (atheroma)in the coronary arteries.

As they narrow because of a build-up of atheroma, the blood supply to your heart will be restricted. This can cause angina (chest pain). If a coronary artery becomes completely blocked, it can cause a heart attack.

Based on this assessment, we would recommend you take the following steps in the coming weeks in order to reduce your risk of heart disease:

Take a multivitamin (like a B complex vitamin)

Eat 3 or more servings of whole grains on most days

Reduce your consumption of trans-saturated fats like cookies, chips, crisps and deep-fried foods

NOTE: *Please remember that the risk rating shown above is only an estimate. If you have any concerns regarding your health, please contact your doctor or GP.*

chs0001

Next >>

III.D. Time Intervention

Table A.2: Food and drink items in the food choice tool

Fruit and Veg	**Meat and Fish**	**Bread and Grains**
Royal Gala Apples, x5	Chicken Breast Fillets, 460g	Sliced White Bread, 800g
Fairtrade Bananas, x5	**Chicken Kiev, Garlic, x2**	Sliced 50/50 Bread, 800g
Oranges, x6	Chicken Goujons, 245g	Sliced Wholemeal Bread, 800g
Red Grapes, 500g	**Chicken Pie, 550g**	White Baguettes, x2
Conference Pears, x4	Beef Rump Steak, 250g	Brown Baguettes, x2
Raspberries, 150g	Beef Mince, 500g	**White Rolls, x8**
Strawberries, 400g	**Steak Burgers, 340g**	**Wholemeal Rolls, x8**
Peaches, x4	**Steak and Ale Pie, 550g**	**Crumpets, 400g**
Kiwi Fruit, x4	**Pork Chops, 450g**	Plain Naan Bread, 260g
Lemons, x5	**Pork Sausages, 400g**	Brown Soda Bread, 400g
Cherry Tomatoes, 650g	**Mini Pork Pies, 300g**	Tortilla Wraps, x8
White Mushrooms, 300g	**Smoked Bacon, 300g**	Basmati Rice, 1kg
Maris Piper Potatoes, 2.5kg	**Lamb Chops, 275g**	Brown Basmati Rice, 1kg
Sweet Potatoes, 1.25kg	Salmon fillets, 240g	Penne Pasta, 1kg
Mixed Peppers, x3	Cod fillets, 250g	Wholewheat Penne, 1kg
Carrots, 1kg	Sea Bass Fillets, 180g	Spaghetti, 500g
Onions, 1kg	**King Prawns, 150g**	Wholewheat Spaghetti, 500g
Fine Beans, 200g	Breaded Cod, 350g	Cous Cous, 1kg
Broccoli, 335g	**Smoked Haddock with Cheese, 400g**	Quinoa, 300g
Sweetcorn, x2	**Salmon en Croute, 380g**	Egg Noodles, 375g
	Cod Fish Fingers, 480g	
	Mackerel in Garlic Butter, 340g	

Confectionery	**Ready Meals**	**Drinks**
Dairy Milk Chocolate, 200g	**Cheese & Tomato Pizza, 10"**	**Blackcurrant Squash, 850ml**
70% Dark Chocolate, 100g	**Pepperoni Pizza, 10"**	Orange & Passion Fruit Drink, 4x275ml
Mars Bars, x4	Beef Lasagne, 430g	Coconut Water, 1L
Snickers, x4	Macaroni Cheese, 430g	**Sports Drink, 1L**
M&M's Peanut, 165g	**Chicken & Bacon Pasta Bake, 430g**	**Energy Drink, 1L**
Nutella, 400g	Cottage Pie, 450g	**Ginger Beer, 1.5L**
Starmix Candy, 215g	Beef Stew, 450g	**Tonic Water, 1L**
Jelly Babies, 190g	Chicken Tikka & Rice, 500g	**Coke, 1.75L**
Marshmallows, 200g	Beef Burrito, 400g	**Irn Bru, 2L**
Chocolate Chip Brioche, x8	Chicken Chow Mein, 450g	**Fanta, 2L**
Chocolate Digestives, 300g	Beef Satay, 380g	Orange Juice, 1L
Chocolate HobNobs, 262g	Chicken Ramen, 380g	Mango & Passion Fruit Smoothie, 750ml
Shortbread Fingers, 400g	**Pigs in Blankets, 260g**	Still Water, 6x500ml
Custard Creams, 400g	Vegetarian Cannelloni, 430g	**Chocolate Milkshake, 1L**
Chocolate Cookies, 175g	**Vegetable Spring Rolls, 60g**	Soya Milk, 1L
Pringles Original, 190g	Vegetable Biryani, 500g	
Salt & Vinegar Chips, 150g	Tomato & Mozzarella Bake, 430g	
Cheese & Onion Crisps, 6x25g	Lentil Cottage Pie, 400g	
Corn Chips, 200g	**Mushroom Risotto, 430g**	
Croissants, x8	**Tomato Soup, 600g**	

Note: Items in **bold** are classified as being 'unhealthy'. This classification is based on the UK Food Standards Agency's nutrient profiling technique developed by the UK's Food Standards Agency (FSA). Points for each item are allocated on the basis of the nutrient content of 100g of a food or drink. Points are awarded for energy, saturated fat, total sugar and sodium (A-nutrients), and for fruit, vegetables and nut content, fibre and protein (C-nutrients). The points from C-nutrients are then subtracted from the score for A-nutrients to calculate a final score. The unhealthy items are then classified as foods with 4 or more points and drinks with 1 or more points. For full details of how the points are calculated please see `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216094/dh_123492.pdf`

III.E. Post-Treatment and Follow-up

Figure A.3: INTAKE24 User Interface

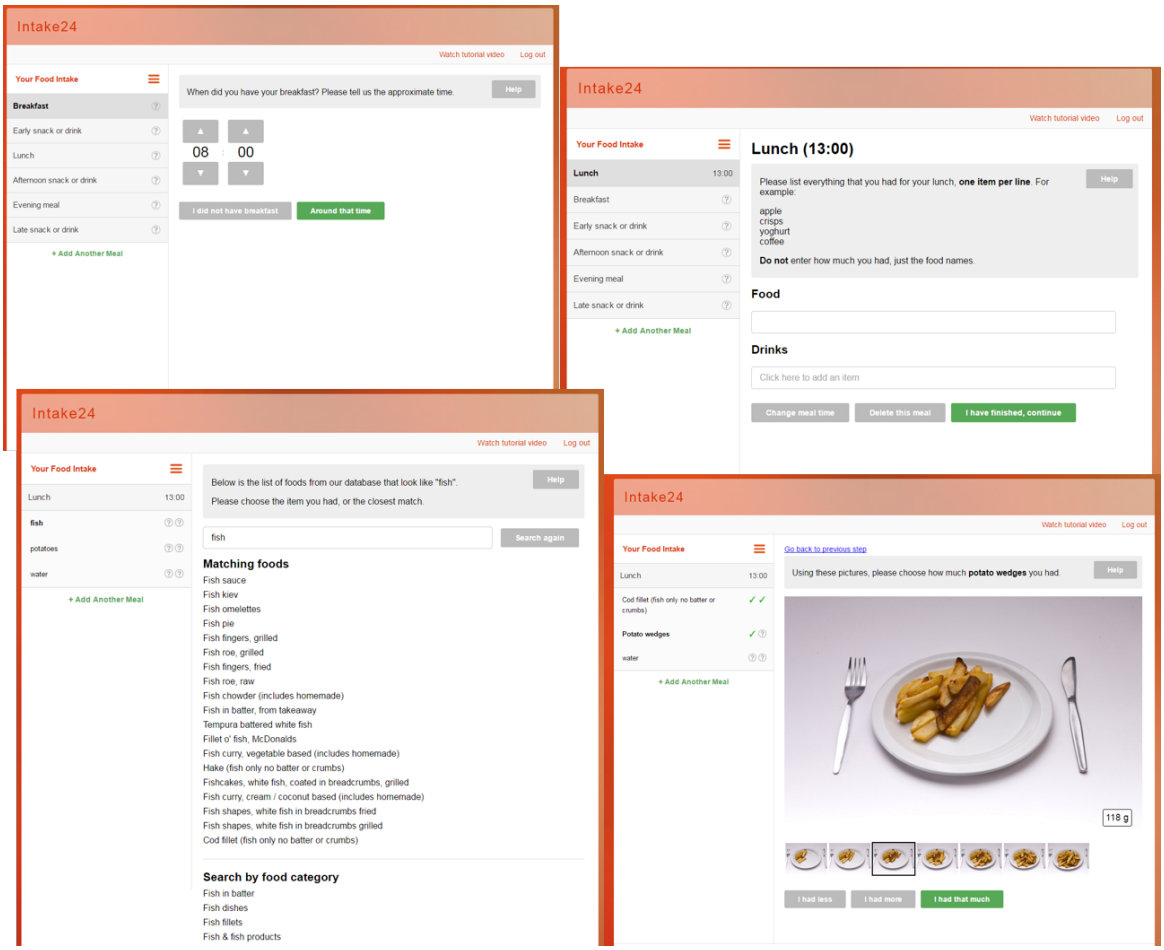


Table A.3: Comparison of Participant Food Choices to UK Household Choices

	ONS Data	Food Choice Tool
Fruit and Vegetables	27.2%	34.6%
Meat and Fish	34.5%	34.1%
Bread and Grains	14.4%	12.1%
Confectionery	12.8%	6.9%
Ready Meals	N/A	5.2%
Drinks	7.5%	7.1%

Note: Figures reported are the proportion of total food and drink expenditure spent on each food category. ONS data are from ONS (2018). Note that data on ready meal spending from ONS data was unavailable due to unclear categorsation of expenditures.

V.A. Summary Statistics

Table A.4: Comparison of the experimental sample to Scottish Health Surveys

	Experimental Sample (mean)	SHS: All		p-val diff.	SHS: <£26,000 Household Income		p-val diff.
		N	(mean)		N	(mean)	
Male	0.39	47,187	0.44	*0.12*	20,050	0.42	*0.36*
White	0.88	47,049	0.97	*0.00*	19,998	0.98	*0.00*
Single	0.70	47,174	0.17	*0.00*	20,045	0.21	*0.00*
Age	36.01	47,187	51.74	*0.00*	20,050	55.15	*0.00*
BMI <25	0.59	39,291	0.32	*0.00*	16,897	0.31	*0.00*
BMI 25 − 29	0.26	39,291	0.38	*0.00*	16,897	0.37	*0.00*
BMI > 29	0.16	39,291	0.30	*0.00*	16,897	0.33	*0.00*
Weight	72.45	39,572	77.95	*0.00*	17,080	77.00	*0.00*
Waist	33.03	8,280	37.00	*0.00*	3,528	37.47	*0.00*
Family history of heart disease	0.19	18,851	0.28	*0.00*	7,438	0.32	*0.00*
Smoker	0.21	47,084	0.23	*0.35*	20,011	0.30	*0.00*
Daily Vitamin	0.27	23,183	0.28	*0.75*	9,408	0.27	*0.85*
Fruit & Veg 5 or more	0.48	47,155	0.23	*0.00*	20,043	0.19	*0.00*
Daily alcohol	0.18	39,655	0.09	*0.00*	15,907	0.10	*0.00*

Note: Means for Scottish Health Survey (SHS) samples calculated using data from 2008-2015 for all variables except Family history of heart disease which was available for 2012-2015 only. SHS samples include all those ages 18+. There are some differences in the measurement of the variables in the experimental sample compared to the SHS due to questions from Your Disease Risk differing slightly in the wording compared to the SHS. "Family history of heart disease": YDR: *Has anyone in your immediate family (mother, father, sister, brother) had a heart attack?* SHS: *Whether either parent or siblings had heart disease/stroke before age 60.* "Fruit & Veg 5 or more": YDR: *Do you eat 5 or more servings of fruit and vegetables a day?* SHS: *Portions of fruit (inc. fruit juice) and veg eaten yesterday.* "Daily alcohol": YDR: *How many servings of alcohol do you have on a typical day?* SHS: *Frequency drank any alcoholic drink last 12 months.* "Daily Vitamin": YDR: *Do you take a multivitamin or a B complex supplement on most days?* SHS: *Take vitamins/minerals to improve health.*

V.B. Information Treatment

Table A.5: The Impact of the Information Treatments on Dietary Choices (nutrients per 100g)

	Unhealthy (%)	Calories (kcal)	Fat (g)	Saturated fat (g)	Sugar (g)	Salt (g)	Fibre (g)	Protein (g)	Spend F&V (%)
Generic	-0.0459**	-10.25	-0.662	-0.216	-0.506	-0.0740***	0.0916	-0.453	0.00655
	(0.0231)	(8.528)	(0.511)	(0.242)	(0.598)	(0.0242)	(0.0836)	(0.380)	(0.0209)
Wild cluster p-val	*0.022*	*0.184*	*0.199*	*0.424*	*0.566*	*0.002*	*0.148*	*0.242*	*0.792*
RI p-val	*0.048*	*0.237*	*0.204*	*0.419*	*0.421*	*0.001*	*0.294*	*0.238*	*0.758*
FDR p-val		*0.439*	*0.439*	*0.455*	*0.455*	*0.020*	*0.439*	*0.439*	*0.754*
FWER adjusted p-val		*0.645*	*0.591*	*0.680*	*0.680*	*0.018*	*0.645*	*0.645*	*0.752*
Tailored	-0.00617	-5.404	-0.0293	0.0651	-0.196	-0.0236	0.0205	0.187	0.0146
	(0.0269)	(9.526)	(0.621)	(0.318)	(0.618)	(0.0301)	(0.105)	(0.492)	(0.0241)
Wild cluster p-val	*0.849*	*0.567*	*0.958*	*0.872*	*0.829*	*0.535*	*0.855*	*0.735*	*0.582*
RI p-val	*0.814*	*0.581*	*0.962*	*0.858*	*0.763*	*0.447*	*0.855*	*0.704*	*0.536*
FDR p-val		*0.962*	*0.962*	*0.962*	*0.962*	*0.962*	*0.962*	*0.962*	*0.962*
FWER adjusted p-val		*0.962*	*0.987*	*0.987*	*0.987*	*0.936*	*0.987*	*0.987*	*0.962*
G = T (p-val)	*0.0574*	*0.471*	*0.124*	*0.165*	*0.478*	*0.106*	*0.298*	*0.127*	*0.684*
G = T = 0 (p-val)	*0.0498*	*0.440*	*0.160*	*0.251*	*0.628*	*0.00929*	*0.341*	*0.217*	*0.830*
R-squared	0.087	0.083	0.068	0.068	0.041	0.070	0.040	0.053	0.097

Note: Observations for all columns equal to 309. The dependent variables are based on the average nutrients per 100g per item. All regressions include controls, these include age, being male, a set of indicators for the time of the experimental session and a dummy indicating whether the participant was hungry or not. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019). FDR q-val calculated using the method from Benjamini and Hochberg (1995) and Anderson (2008). FWER correct p-value refers to the p-value using the step-down methods of Romano and Wolf (2005a,b). All p-values are calculated using 5000 replications. An item is classified as 'unhealthy' is based on the UK Food Standards Agency's nutrient profiling technique. Points for each item are allocated on the basis of the nutrient content of 100g of a food or drink. Points are awarded for energy, saturated fat, total sugar and sodium (A-nutrients), and for fruit, vegetables and nut content, fibre and protein (C-nutrients). The points from C-nutrients are then subtracted from the score for A-nutrients to calculate a final score. Unhealthy items are foods with 4 or more points and drinks with 1 or more points.

Dietary Recommendations and Food Choices — A another aspect to consider in relation to the tailored information cohort is the impact of individual recommendations on their food choices. As detailed earlier, apart from receiving the risk of contracting heart disease and diabetes, subjects in this group also received a set of recommendations in order to reduce this risk, based on the responses provided in the YDR tool. It would therefore be interesting to analyse whether these personalised recommendations had any influence on the food choices made by tailored information subjects. For the purposes of this section, we only consider those recommendations that can be linked to the nutrients considered in this paper, namely those that recommend an increase in fruit and vegetable intake, a reduction in saturated and trans-saturated fat consumption, as well as the recommendation to lose weight. The rationale behind this analysis is that although the actual heart/diabetes risk scores may have had no impact on subjects' food choices due to their above-average health status, individual recommendations may have resulted in systematic differences in the nutritional composition of their food baskets.

The results are shown in Table A.5 below, where we run the usual set of nutrient regressions (including controls), only this time including additional dummies denoting the recommendations cited above, interacted with the tailored information treatment dummy.[2] We find that the individual recommendations largely had no effect on food choices in the tailored information treatment. Except for the recommendation to reduce saturated fat which led to an increase in fat and saturated fat, conversely, in column (4) we observe that the saturated fat recommendation led to a negative and statistically-significant reduction of intake of sugar. In column (5) we find that those tailored subjects who received the saturated fat recommendation spent a larger proportion of their budget on unhealthy items. In addition, we also find that a recommendation to lose weight led to lower expenditure on fruit and vegetables in our tailored group, which is also contrary to expectations. These findings must be interpreted with caution due to the relatively small sample of our tailored information cohort and the limited variation therein.

[2]We only do this for the tailored information group since generic health information subjects were all shown the entire list of recommendations, rather than those specific to their own requirements.

Table A.6: Dietary Recommendations and Food Choices

	Unhealthy (%)	Calories (kcal)	Fat (g)	Saturated fat (g)	Sugar (g)	Salt (g)	Fibre (g)	Protein (g)	Spend F&V (%)
Generic	-0.0473**	-357.7	-45.06**	-19.49**	0.916	-3.139**	10.01	-8.865	0.00749
	(0.0231)	(453.2)	(19.14)	(8.938)	(36.46)	(1.409)	(10.10)	(21.86)	(0.0209)
Wild cluster p-val	*0.0190*	*0.544*	*0.0154*	*0.0602*	*0.982*	*0.0364*	*0.399*	*0.700*	*0.761*
RI p-val	*0.0396*	*0.432*	*0.0194*	*0.0304*	*0.981*	*0.0226*	*0.321*	*0.694*	*0.725*
Tailored	-0.0777**	-939.2	-63.67***	-28.20**	97.07*	-5.445**	14.42	-34.20	0.0779**
	(0.0320)	(594.7)	(22.47)	(11.17)	(53.85)	(2.398)	(15.43)	(31.26)	(0.0331)
Wild cluster p-val	*0.113*	*0.184*	*0.0198*	*0.0822*	*0.0874*	*0.0114*	*0.401*	*0.247*	*0.0988*
RI p-val	*0.0150*	*0.121*	*0.00480*	*0.0104*	*0.0770*	*0.0238*	*0.357*	*0.278*	*0.0182*
Tailored*Fruit & Veg	0.0471	98.36	11.39	6.777	-79.82	5.089**	-14.26	51.77	-0.0468
	(0.0306)	(659.0)	(22.38)	(9.617)	(51.09)	(2.486)	(13.96)	(33.62)	(0.0316)
Wild cluster p-val	*0.0840*	*0.797*	*0.369*	*0.273*	*0.160*	*0.00640*	*0.224*	*0.0986*	*0.171*
RI p-val	*0.133*	*0.886*	*0.623*	*0.507*	*0.132*	*0.0432*	*0.335*	*0.138*	*0.147*
Tailored*Saturated Fat	0.0942***	-195.0	51.91**	27.07**	-144.4***	3.069	-36.61**	-13.33	-0.0741**
	(0.0356)	(638.0)	(23.13)	(11.61)	(52.69)	(3.284)	(14.21)	(33.60)	(0.0305)
Wild cluster p-val	*0.00140*	*0.736*	*0.0378*	*0.0206*	*0.0156*	*0.575*	*0.0318*	*0.762*	*0.0732*
RI p-val	*0.00900*	*0.759*	*0.0278*	*0.0206*	*0.00820*	*0.387*	*0.0136*	*0.699*	*0.0220*
Tailored*Trans Fat	0.00486	585.8	14.62	6.369	8.607	1.564	7.404	25.12	-0.0105
	(0.0343)	(716.6)	(23.54)	(10.37)	(54.48)	(3.533)	(14.43)	(36.07)	(0.0336)
Wild cluster p-val	*0.964*	*0.753*	*0.514*	*0.552*	*0.899*	*0.920*	*0.682*	*0.372*	*0.663*
RI p-val	*0.893*	*0.433*	*0.554*	*0.556*	*0.882*	*0.681*	*0.621*	*0.494*	*0.758*
Tailored*Weight Loss	0.124**	-472.6	68.98	27.91	-112.4	0.735	-42.32**	-1.351	-0.101**
	(0.0567)	(1,259)	(42.42)	(18.69)	(119.5)	(3.708)	(20.67)	(52.22)	(0.0481)
Wild cluster p-val	*0.162*	*0.543*	*0.205*	*0.166*	*0.564*	*0.851*	*0.163*	*0.978*	*0.182*
RI p-val	*0.0790*	*0.718*	*0.156*	*0.188*	*0.392*	*0.853*	*0.0730*	*0.980*	*0.0686*
R-squared	0.137	0.098	0.135	0.123	0.077	0.077	0.055	0.089	0.137

Note: Observations for all columns equal to 309. The dependent variables are based on the totals of the basket. All regressions include controls, these include age, being male, a set of indicators for the time of the experimental session and a dummy indicating whether the participant was hungry or not. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019).

V.B. Time Availability Treatment

Interaction between Information and Time Treatments — We also take a look at how our two interventions interact with one another to influence dietary choices. Although we have found no statistically-significant impact of time availability on the nutritional composition of food baskets, it is possible that we may observe differences across the information treatments depending on whether they had more or less time to ponder their choices. Specifically, it might be the case that tailored information participants assigned to the long time treatment may have selected healthier options from the food choice tool since they had enough time to digest the information provided and thus select accordingly, in contrast to the short time treatment group who only had 3 minutes to make their choices. This may also be true of the generic health information group, with those in the long time treatment driving the statistically-significant results reported above. Thus, we interact the long time availability dummy with both our generic information and tailored information dummies, and include them in our main regression model.

The results are shown in Table A.6. It is clear that the vast majority of the interaction terms included do not have any statistically-significant impact on food choices. The only significant result obtained is for the tailored health information/high time treatment with regards to sugar, which suggests that this cohort actually selected more sugary items in their food baskets. Nonetheless, the lack of significant results elsewhere further underlines the minimal influence of time availability on people's food choices in this experiment.

Table A.7: The impact of information and time on dietary choices (interaction of treatments)

	Unhealthy (%)	Calories (kcal)	Fat (g)	Saturated fat (g)	Sugar (g)	Salt (g)	Fibre (g)	Protein (g)	Spend F&V (%)
Generic	-0.0456	97.93	-43.86	-19.89	-16.66	-1.268	19.12	14.91	-0.0250
	(0.0319)	(626.8)	(26.61)	(12.16)	(49.90)	(1.664)	(13.56)	(28.74)	(0.0289)
Wild cluster p-val	*0.354*	*0.845*	*0.210*	*0.268*	*0.821*	*0.222*	*0.278*	*0.712*	*0.622*
RI p-val	*0.168*	*0.886*	*0.116*	*0.114*	*0.738*	*0.473*	*0.162*	*0.613*	*0.391*
Tailored	0.00536	-920.5	-22.23	-7.310	-65.11	0.412	-15.19	30.34	-0.0114
	(0.0384)	(636.8)	(29.70)	(14.47)	(47.93)	(2.633)	(14.42)	(30.55)	(0.0331)
Wild cluster p-val	*0.899*	*0.126*	*0.489*	*0.649*	*0.233*	*0.879*	*0.112*	*0.455*	*0.685*
RI p-val	*0.895*	*0.150*	*0.473*	*0.634*	*0.174*	*0.893*	*0.288*	*0.324*	*0.726*
Long Time*Tailored	-0.0277	203.5	-11.88	-6.514	148.1**	-3.703	16.53	-83.94*	0.0604
	(0.0481)	(885.4)	(33.97)	(15.55)	(72.07)	(3.652)	(21.83)	(47.70)	(0.0470)
Wild cluster p-val	*0.565*	*0.834*	*0.614*	*0.641*	*0.0126*	*0.415*	*0.462*	*0.127*	*0.306*
RI p-val	*0.559*	*0.818*	*0.723*	*0.673*	*0.0380*	*0.327*	*0.444*	*0.0772*	*0.201*
Long Time*Generic	-0.00222	-845.5	-0.543	1.534	39.62	-3.938	-16.97	-53.84	0.0636*
	(0.0381)	(777.5)	(30.56)	(14.14)	(61.46)	(2.422)	(19.04)	(39.74)	(0.0367)
Wild cluster p-val	*0.952*	*0.402*	*0.987*	*0.929*	*0.529*	*0.160*	*0.500*	*0.208*	*0.206*
RI p-val	*0.950*	*0.283*	*0.988*	*0.916*	*0.510*	*0.106*	*0.367*	*0.187*	*0.0858*
Long Time	0.00647	44.66	1.057	-0.744	-39.21	3.101	-1.978	52.06	-0.0371
	(0.0344)	(644.1)	(27.97)	(12.81)	(48.94)	(2.103)	(15.70)	(32.84)	(0.0333)
Wild cluster p-val	*0.736*	*0.963*	*0.954*	*0.949*	*0.507*	*0.289*	*0.900*	*0.163*	*0.446*
RI p-val	*0.855*	*0.950*	*0.971*	*0.957*	*0.417*	*0.145*	*0.903*	*0.119*	*0.277*
R-squared	0.088	0.100	0.111	0.099	0.054	0.057	0.029	0.087	0.106

Note: Observations for all columns equal to 309. The dependent variables are based on the totals of the basket. All regressions include controls, these include age, being male, a set of indicators for the time of the experimental session and a dummy indicating whether the participant was hungry or not. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild clustered bootstrapped p-values are clustered at the session level using the procedure by Cameron et al. (2008). RI p-value refers to the equivalent p-value using a Randomization Inference procedure specifically, the randomization-t p-value from Young (2019).

V.D. The Long-Term Impact of Information on Dietary Choices

Table A.8: Comparison of Sample to Scottish Population & Follow-up sample

	Mean		p-values
	Full Sample	Follow-up Sample	Difference (p-value)
No Information	0.291	0.285	0.394
Generic Info	0.359	0.356	0.397
Tailored Info	0.350	0.359	0.387
Long Time	0.505	0.516	0.386
Short Time	0.495	0.484	0.386
% Males	0.395	0.379	0.370
Married	0.696	0.676	0.350
Employed	0.557	0.582	0.332
Unemployed	0.081	0.086	0.390
Student	0.252	0.223	0.283
Income above £25,000	0.068	0.070	0.396
Income £20,000-25,000	0.285	0.305	0.349
Income £15,000-20,000	0.207	0.215	0.389
Income £10,000-15,000	0.220	0.211	0.385
Income £5,000-10,000	0.126	0.125	0.398
Age	36.0	37.0	0.246
Weight	72.5	73.1	0.346
BMI	25.0	25.3	0.306
Waist	33.0	33.3	0.358
Family History Heart Disease	0.191	0.207	0.356
Regular Diet	0.783	0.773	0.384
Vegetarian	0.152	0.156	0.395
Fruit and Vegetables 5 a day	0.479	0.477	0.398
Daily Alcohol Serving	0.181	0.211	0.270
Smoker	0.210	0.207	0.397
N	309	256	

References

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association 103*(484), 1481–1495.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological) 57*(1), 289–300.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Romano, J. P. and M. Wolf (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association 100*(469), 94–108.

Romano, J. P. and M. Wolf (2005b). Stepwise multiple testing as formalized data snooping. *Econometrica 73*(4), 1237–1282.

Young, A. (2019). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The Quarterly Journal of Economics 134*(2), 557–598.