



Shopper's behavioural responses to 'front-of-pack' nutrition logo formats: GDA Diet-Logo vs. 6 alternative Choice-Logos

Laurent Muller, Bernard Ruffieux

► To cite this version:

Laurent Muller, Bernard Ruffieux. Shopper's behavioural responses to 'front-of-pack' nutrition logo formats: GDA Diet-Logo vs. 6 alternative Choice-Logos. 2020. hal-02479464

HAL Id: hal-02479464

<https://hal.science/hal-02479464>

Preprint submitted on 14 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GAEL

Grenoble Applied Economic Laboratory

Consumption – Energy - Innovation

**Shopper's behavioural responses to 'front-of-pack'
nutrition logo formats:**

GDA Diet-Logo vs. 6 alternative Choice-Logos

**Müller, Laurent
Ruffieux, Bernard**

February, 2020

JEL: D120, D180, C910, C930.



Shopper's behavioural responses to 'front-of-pack' nutrition logo formats:

GDA Diet-Logo vs. 6 alternative Choice-Logos

Laurent Muller ¹; Bernard Ruffieux ¹

Abstract: A Framed Field Experiment was implemented in France in order to compare the relative behavioural responses and then the induced nutritional effectiveness of seven front-of-pack logo formats: The *Guideline Daily Amounts* (GDA), and six 'choice logos' such as *Green Keyhole* or *Traffic Lights*. From a consumer point of view, while GDA requires a demanding global-diet heuristic, the choice logos require an easy product comparison heuristic. Our six 'choice logos' are different after 3 criteria: aggregate *vs.* analytical information, shelf *vs.* all products point of reference, multicolour logos *vs.* 'only green' logos. We measure the effect of each logo on the nutritional quality of actual consumers' shopping baskets in a controlled experimental shop. We use a standard criterion aggregating the overall density of free sugar, saturated fatty acid and salt. We find that different logo formats generate different nutritional impacts. Some choice logos have better nutritional impacts than GDA. Aggregate and multicolour logos induce best responses. Shelf-referenced logos are not more efficient but they trigger very different behavioural trajectories than logos referenced on all-products.

February 2020

Keywords: food consumption, nutritional labelling, food policy, framed field experiment, behavioural economics

JEL codes: D120, D180, C910, C930.

¹ Univ. Grenoble Alpes, CNRS, INRAE, Grenoble INP, GAEL, 38000 Grenoble, France

1. Introduction

Since the 1970's, when official reports began to recommend nutritional guidelines for consumers, food packaging labelling has been recognized as a relevant political tool that might be used to encourage consumers to adopt a healthier diet. Presently, most political authorities recognize this labelling policy as a key consensual instrument (Drichoutis, Lazaridis & Nayga, 2006; Grunert & Wills, 2007). Several front-of-pack labelling systems have been proposed such as Traffic Lights and colour-coded Guideline Daily Amounts (UK), Heart Check (USA), Green Keyhole (Sweden), Pick the Tick (Australia and New Zealand), Health Check (Canada). How do consumers respond to these different systems and why? We study here the different behavioural impact induced by different labelling systems, as the actual nutritional impact of a given system on consumer choices might differ from the impact of another system. Our aim here is to test the actual current shopper response to various front-of-pack labelling systems. Seven contrasted systems are put in competition in a laboratory mini-shop where ordinary consumers come for their actual food shopping.

The seven systems that we put in competition had been chosen for their contrasted expected consumer responses. These responses depend first on the consumer preferences, but also on the consumer rationale, technically speaking on the 'heuristic' used to choose products in the shop. Szanyi (2010), Kleef and Dagevos (2015) and Muller & Prévost (2016) expressed the need to relate labels use with explicit psychological phenomena. Here we explicit a series of heuristic proposals that justify our selection of logo formats and might explain our experimental results on consumers actual shopping responses. Each format is designed according to three building blocks or criteria: the unit ranked, the reference set and the scope of colours.

Once the 7 systems designed, we then use a framed field experiment (Harrison & List, 2004)

to put these systems in competition. To do that, we observe shoppers' behaviour before and after each nutritional logo has been applied to all the products of our laboratory mini-shop. Hence, we are able to directly measure the effects of each logo system on the nutritional quality of shoppers' baskets contents. Following the WHO's recommendations, nutritional quality is assessed by measuring the overall density of free sugar, saturated fatty acid and salt. Finally, this study sets out to answer three questions. (a) Do front-of-pack logos significantly improve the nutritional quality of diets? (b) Do different logo systems generate different behavioural responses and therefore contrasted nutritional performance? (c) And finally, what element might a system include to induce best nutritional response?

In the next section we present the six proposals regarding to behavioural responses to logo formats that are tested here. Second, we explain our methodological choices comparatively to the existing literature in section 3. We then describe the experiment and the results in sections 3 and 4. Finally, we discuss the results and conclude in section 5.

2. Logo formats and heuristic proposals

2.1. From back-of-pack panels to front-of-pack logos

Standard economic theory assumes that individuals maximize their well-being subject to a set of constraints. In food contexts, this means that, given the information and income available, consumers are making their best possible food choices. Amongst other things, they arbitrate between short-term hedonic pleasure and longer-term gains of health and wellness. With such rational heuristic, the task of public deciders should only consist in fixing externalities (*e.g.* social cost of obesity) by altering relative prices (for example through taxes and subsidies) and in providing the consumer with all the useful nutritional information – as it is not directly accessible through the consumption of the product. On this last point, considerable efforts have been made to fully inform consumers. In most countries, the back of food packs now contains,

in addition to the list of ingredients, a nutrition panel giving precise information on the product content in nutrients. Our *Homo Æconomicus* has thus everything in hands to select his optimized diet knowing his preferences and constraints. Nevertheless, such nutrition panels have in fact produced very poor behavioural responses by consumers: no improvements have been observed in deep-rooted food consumption trends. While most consumers clearly welcome the presence of such precise information (Philipson, 2005; INPES, 2008; McCullum & Achterberg, 1997; Chan, Patch & Williams, 2005; Wansink, 2003), they do not find such panels easy to understand and to use (Grunert, Wills & Frenandez-Celemin, 2010; Cowburn & Stockley, 2005). Most consumers do not even try to read it (Drichoutis, Nayga & Lazaridis, 2011; Levy, & Fein, 1998; Roe, Levy, & Derby, 1999).

For a food consumer, nutrition is only one out of many important characteristics such as price, taste, ease-of-use, safety, and, at least for some, sustainability. Due to the fact that the health impact is not prominent, uncertain and delayed, other characteristics such as taste and price are regarded as more straightforward and thus prioritized in the heuristic used by regular shoppers. As food shoppers are usually in a hurry, they may be unmindful of health issues, attracted by colours, trademarks, and the marketing environment (Wansink, 2004). Attention matters. Given this shopping environment, some stakeholders have suggested bringing synthetic and essential information to the front of the packs in order to generate higher consumer awareness. A first objective is to catch consumer attention: salient logos may be displayed front-of-pack and visible at a glance. Therefore, our first proposal is the following.

- Proposal 1. Front-of-pack logos improve the nutritional quality of food baskets

By catching attention and stressing nutritional concern, front-of pack logos may significantly improve the nutritional quality of the consumer basket. A typical consumer is making fast shopping decisions with poor consideration to nutritional issues (Szanyi, 2010). *Attention* may

then play a key role for a better consideration of nutritional quality in consumer's preferences (van Kleef & Dagevos, 2015). Front-of-pack logos may increase the likelihood of capturing the shoppers' attention (Bialkova & Van Trijp, 2010; Van Herpen & Van Trijp, 2011). Once the attention is captured, the weight of the health component in food choice should increase relatively to other food attributes. Consumers are far from ignorant of nutritional issues, they may then identify what is nutritionally good or bad for them and somewhat improve their diet when aware. In order to be fully convinced by the proposal, we may suggest a *deduction ad absurdum*: a concerned, ready to invest in efforts, but ignorant of nutritional issues consumer would use a back-of-pack panels. We know they are not efficient (Cowburn & Stockley, 2005).

2.2. Diet Logo v.s. Choice Logo

To capture attention, the “2010 First Lady Michelle Obama initiative” moved GDA from back to front-of-pack (“facts up front!”)¹. Front-of-pack GDAs convert the analytical, precise information of the back-of-pack panels into consumer more friendly, synthetic information. They provide guidelines that intend to help consumers to consider nutrition in the context of their overall daily diet. The heuristic issue here is that optimizing a daily diet may happen to be highly complex. The maximization program includes thousands of products and dozens of nutrients, needing computing abilities and very good memory. Let's consider a simple case: buying a pizza. The shopper needs to think globally in order to identify the marginal impact on his daily diet, and this nutrient by nutrient. Even if the consumer restricts his nutrition optimization program to the three major nutrients to limit (sugar, fat and salt) the computation skill is certainly beyond reach for any consumer. While *Daily Values* ease decision fatigue by providing the daily percentage of each nutrient based on are commended diet, making food

¹ <http://www.gmaonline.org/issues-policy/health-nutrition/facts-up-front-front-of-pack-labeling-initiative/>

decisions based on the GDA recommendations still requires global, slow and effortful decision process. Therefore, our second proposal is the following.

- Proposal 2. GDA induce lower nutritional impact than choice logos

GDA is the right decision-support to assist extremely rational behaviour. Due to *limited computing capability* and *lack of global thinking*, GDA may not be the most efficient logo system for fast thinking consumers. In Crosetto, Muller & Ruffieux (2016)'s experiment, collectedly participants invited to think slowly and deep were invited to build diets that fit the daily-recommended intake for each nutrient. Success rate was higher with GDA than with the coarser Traffic Lights. By providing the appropriate numerical data, GDA is better equipped to resolve an optimization problem under constraints. Participants took all the necessary time and provided the necessary effort to make such computations. In a similar experimental context, when setting a time constraint and, therefore stress participants to decide fast, preventing from careful computations, GDA's edge disappeared or even reversed to the benefit of Traffic Lights. The use of simpler heuristics to shorten and ease decision-making is all the more likely to occur in non-laboratory setting. Indeed, supermarkets are certainly not the right place for slow and demanding heuristics. It has been estimated that people make an average of 200 to 300 decisions regarding food consumption in any given day (Wansink & Sobal, 2007), food consumers do not compute all available information (Just & Payne, 2009) and most decisions are taken on the basis of local thinking (Ariely, 2009). Such deviations from rational thinking are well documented (DellaVigna, 2009). Speed (ready-access), perception and saliency are important decision determinants (Bialkova & van Trijp, 2010). From here, logo system designers and stakeholders may wish to take advantage of these behavioural traits to efficiently guide consumers towards healthier diets. Accordingly, logos ought to be prescriptive, giving clear hints about the nutritional quality of the product compared to others avoiding a detour to the

global diet. *Traffic Lights* or *Green Keyhole* are examples of such systems. Both differ from GDA by providing normative assessments regardless the global diet. We call them *Choice logos*. Their use of graphic symbols to convey their judgment are better recognized than GDA's use of words and numbers (Geiger, 1998; Viswanathan et al., 2009).

2.3. *Three criteria defining the choice logos*

Traffic Lights and Green Keyhole differ, at least, by three criteria: the unit graded, the reference set and the scope of colour. First, while Traffic Lights rate each nutrient, the unit evaluated with Green Keyhole is the product as a whole. Second, while Green Keyhole only signals the best products, Traffic Lights signal all products with green, amber or red marks. Third, Green Keyhole compares products within a given shelf, whereas the ratings of a Traffic Light relates to all products. Therefore, our proposal 3, and our proposals 4, 5 and 6, criterion per criterion, are the following.

- Proposal 3. Consumer responses differ according to the choice logo format

Different choice logo formats may induce different behavioural responses in magnitude and quality. Some system may require more cognitive efforts or leave the consumer with more tricky trade-offs to solve. Consumer may the give up or follow various heuristics of change. As a result, the nutritional impacts induced may differ. We present the expected effect for each criterion.

- Proposal 4. Unit graded: Grading the whole products induces greater nutritional improvement than grading nutrient by nutrient

As we saw with GDA, more information is better only for those who are able to process it. The degree of aggregation of information may differ, leading to either an analytical evaluation or an overall healthiness evaluation (Lytton 2010). A logo may evaluate each nutrient or evaluate

the whole product with a unique global indicator aggregating the nutrients contents (like the *Smart choice* or the *Keyhole* logos). Simple aggregated logos are easier to understand and use (Wanskink, Sonka & Hasler, 2004). They are desired by consumers (Feunekes, Gortemaker, Willems, Lion & van den Kommer, 2008). A global unique logo allows avoiding difficult trade-offs when a product includes both good and bad evaluations. Reduced efforts may then stimulates more changes. On the other hand, an aggregated logo may induce misleading nutrient evaluation (Andrews, Burton & Kees, 2011) and hence useless, even counterproductive efforts if hidden trade-offs appear (*e.g.* negative correlations among nutrients)².

- Proposal 5. Reference set: Absolute grading based on all products induces greater nutritional improvement than relative grading based on shelves

Would it be more efficient to rank products in reference to the whole food supply or to a category of products (in a supermarket, each shelf comprises close substitutes). According to Higginson, Kirk, Rayner, & Draper (2002) and Feunekes, Gortemaker, Willems, Lion, & van den Kommer (2008), it remains unclear for consumers whether, facing a given label, food is evaluated in absolute or relative to a particular category. A logo referring to all products will less differentiate the nutritional quality of products that belong to the same food category. But shelf-by-shelf logo will enhance within category substitutions, to the detriment of between categories substitutions. Consumers may over-evaluate relatively good products within an unhealthy category. For instance, (Steenhuis et al. 2010) found that chocolate cake are perceived as less unhealthy with a choice logo per category than without logo. Furthermore, substituting across shelves (for instance substituting crisps for radish) may induce more nutritional gain than substituting within a shelf (for instance substituting crisps to lighter crisps).

² Breaking down nutrients may also give more transparency and thus appear as more credible.

A counter argument may be that intra-category substitutions may require less effort for consumers.

- Proposal 6. Scope of colour: Red and green induces greater nutritional improvement than green only systems

Should a system signal only good products (using a green label) or should all products be signal by an explicit signal from green to red? A red indicator, with its off-putting effect, may make a strong emotional impression on shoppers, transitory but yet significant (Elliot, Maier, Moller, Friedman & Meinhardt, 2007). *Velleity* plays a big part, and a balanced diet is often postponed to tomorrow (Ariely, 2010). With an emotional red patch, some substitutions or some relinquishments may be less subject to procrastination. Some participants may also value *compliance*, while others may refuse it (Kelman, 1958).

3. Methodological background of the experiment

This study adds to the existing literature by proposing a novel experimental design that (i) allows the observation of the whole food baskets rather than just a specific items, (ii) is based on real purchase rather than stated intentions and (iii) offers *ceteris paribus* conditions for a straightforward comparison between several competing logo systems. While many studies have investigated consumers' attention, understanding or attitude towards front-of-pack labelling, very few studied the impact of front-of-pack labels on actual purchase behaviour (Kleef and Dagevos 2015). The vast majority of the literature focuses on one product or one product family (Borgmeier & Westenhoefer, 2009; Steenhuis, Kroeze, Vyth, Valk, Verbauwen & Seidell, 2010; Sacks, Rayner & Swinburn, 2009; Vyth, Steenhuis & Heymans, 2011). Not considering all possible substitutions, such studies does not allow the examination of purchasing behaviour in a global way (Balcombe, Fraser, & Falco 2010). In the present experiment, participants could choose among 273 items representative of the French food supply.

Until now, most studies comparing logo formats are self-reported consumer studies (Balcombe, Fraser & Di Falco, 2010; Borgmeier & Westenhoefer, 2009; Feunekes, G., Gortemaker, I., Willems, A., Lion, R. & van den Kommer, 2008). In such studies, respondents may potentially over report their use of labels in order to give socially desirable answers (Cowburn & Stockley, 2005; Grunert & Wills, 2007; Malam et al., 2009). In the present experiment, decisions are ‘consequential’: participants really purchase a subset of the items they have selected during experimental session.

Finally, in most studies, the causality of the relationship between nutritional logos and purchasing behaviour is unclear. While natural field experiments (studies that observe purchase behaviours in actual points of purchase, e.g. Borgmeier and Westenhoefer, 2009, Vermeer et al., 2011, and Vyth et al., 2011) better simulate the real life shopping than researches conducted in laboratory settings, their natural environments provide too much explicative variables that could mitigate the interpretations of the results. Laboratory settings allow the construction of proper counterfactual scenarios and hence enable comparative statics for ascertaining treatment effects. van Herpen and Trijp (2011) argue that the ultimate test would consist in comparing sales data prior and after the introduction of nutrition labels like in Sacks, Rayner, and Swinburn (2009) and in Sutherland, Kaley, and Fischer (2010). However, the latter studies do not keep track of individual trajectories and thus do not control for individual differences. The present protocol architecture used in the present study includes observations of the same individuals ‘before’ and ‘after’ a logo system is introduced. By allowing for *within subjects* measurements of behavioural responses, we control for sampling and context variability. To our best knowledge, such design has only been used to examine the effect of a price policy (Muller, Lacroix, Lusk & Ruffieux, 2016). Artificial settings may have another advantage over sale studies: experimenters are not limited by what exist.

4. Materials and methods

4.1 Implementing an E-Shop in the Lab

This framed field experiment uses an e-shopping mock-up that includes a total of 273 food products³ in 35 familiar food categories⁴. Defining the list of food items available and sorting these items into categories is critical for the relevance and generalization of the results. Unfortunately and unavoidably, it contains *ad hoc* features. In this experiment, categories correspond to the usual classification used in self-service grocery stores in France. It also fits the standard classification proposed by OQALI⁵. Each category includes six, nine or twelve products⁶. With the help of renowned French nutritionists and consumption data, food items have been chosen among the most frequently bought products in France, so as to model the existing range of nutritional quality in each food category. Products were proposed at current outside market prices. Posted prices had been observed in a local supermarket at the time of the sessions. Participants were aware of that.

Each participant was seated alone in front of a computer and was handed a paper catalogue containing all the 273 food products. The catalogue is a 35-page A4 format colour booklet. Each page comprises all the products of one same category. Each product is associated with a coloured front-of-pack picture with its name, its price and a bar code (see appendix A1). By

³ Condiments and cooking fat are excluded.

⁴ The 35 product categories (and number of items in each category) are: Meat (9); Cooked meat (9); Sausages (6); Fish and Seafood (9); Fresh Vegetables (9); Fresh Fruit (9) Canned Vegetables (9); Cooked Potatoes (9), Pasta and Rice (6), Cooked Vegetables (6), Ready-Made Meals with Meat (9); Ready-Made Meals with Cheese (6), Ready-Made Meals with Fish (6) Pizzas (6); Pies and Quiches (6) Sandwiches (6) Snacks (9), Salads (6) Soups (6), Milk (9), Dairy Products (9), Cream (6) Pudding (6); Flavoured Milk Dessert (6), Dry Cheese (9), Bread and Buns (9), Pastries and Cakes (6) Biscuits (9) Chocolate Bars (6), Breakfast Cereals (9) Spreads (6) ; Chocolate (9); Fruit Desserts (9); Juices and Nectars (12), Sodas and Syrups (12).

⁵ OQALI is a French public agency whose mission is to observe food quality, and in particular nutritional quality (http://www.oqali.fr/oqali_eng).

⁶ We use only three multiples in order to ease our ranking configuration (1/3 best, 1/3 worst and 1/3 neutral).

reading any product code with an easy-to-us bar code reader, the user makes this product pop up on the computer screen. She may then use the computer keyboard to buy one or more units of the selected item. On the right side of the screen the work-in-progress basket appears. It includes the name of items already selected, the price of each item and the total amount already spent. Any selected item may easily be removed from the basket during the shopping stage⁷.

4.2. Seven logo formats in competition in the Lab

In this study, we test seven logo formats: One diet logo (GDA) and six choice logos. The latter differ according to the three criteria defined in subsection 2.3. Each of these three criteria takes two options: *product* (P) vs. *nutrients* (N) for the unit graded, *category* (C) vs. *overall* (O) for the reference set and *green* (G) vs. *traffic lights* (T) for the scope of colour. With option N, each nutrient of a given product is graded as such. We only consider the density (g per 100g) of the following three nutrients: salt, free sugar and saturated fatty acid. These nutrients are consensually recognized among nutritionists as the nutrients to be limited⁸. Therefore, the nutrients logos include three qualifications (colour patches in our experiment), one for each of our three nutrients. With option P, one aggregated grade is used: the LIM indicator (Darmon, Vieux, Maillot, Volatier & Martin, 2009)⁹. With option C, each product is graded according to its product category. For example, a ‘croissant’ is ranked in the category ‘Viennese pastries and cakes’. With option O, a product is ranked referring to all food products in our experimental E-Shop. With option T, the following sorting rule will be used: the best 1/3 will be green, the

⁷ The front-of-pack pictures of each of our selected 273 items, the logos, the graphical user interface, and the store paper catalogues were designed and produced in our lab, specifically for the purpose of this experiment.

⁸ WHO (Global Strategy on Diet, Physical Activity and Health 2004); USDA (U.S. Department of Health and Human Services, 2010. *The Surgeon General's Vision for a Healthy and Fit Nation*. Rockville, MD: U.S. Department of Health and Human Services, Office of the Surgeon General) and PNNS (<http://www.mangerbouger.fr/bien-manger/les-9-reperes/>)

⁹ See subsection 2.4 for details.

worst 1/3 will be red and the remaining 1/3 will be colourless¹⁰. Finally, only the best third is rewarded with a green dot with option G. Each choice logo is defined as a combination of three options, one for each criterion. Six of the eight possible formats that this system generates have been selected; the other two are too complex to be promising¹¹. Thus a selection of six logo formats with which to compare GDA is proposed (See table 1).

4.3 The experiment

Participant recruitment was done via telephone, Internet and flyers. The experiment was conducted with 364 adults in the greater Grenoble area in France (see appendix B for a sample description). Participants had to be 18 years old or older, to have at least one child living in the household, and to be a regular food shopper for the entire household. They were aware that the research was public, aimed at fundamental research and that no private corporation was involved in any way in the research, including its objectives and funding. They were also aware that they would have the opportunity to buy food products for research purposes. However, the nutritional aim of the study was not mentioned and participants were not told that the French Ministry of Health had funded the research. All sessions were held between 6 September and 9 October 2010. Sessions took place in the experimental laboratory of the Grenoble Institute of Technology. Forty-four sessions were organized, each dedicated to one of the 7 treatments (one per logo format). A session lasted two hours.

At the outset of the experiment, participants were given €25 as a fixed compensation for participating in the study. In stage 1, they were asked to shop for food in order to feed their

¹⁰ Note that the fresh fruits category and the fresh vegetable category, though present in our e-shop, were not labelled in any format treatment.

¹¹ The Steering Committee of the study included scientists, government officials, and representatives from consumer associations. It validated: (i) the logo formats selection, (ii) the catalogue including products and categories, (iii) the architecture of the experiment, (iv) the criteria used to recruit the subjects.

household members over two days following the experiment. They were free to choose any quantity of any items from the food catalogue. To avoid hypothetical bias (Carson & Groove, 2007; Carlsson, 2011), decisions were made incentive compatible: Participants were informed that at the end of the session they would have to buy a significant sub-set of the products they have chosen during the session, *i.e.* around one eighth. They did not know until the very end of the experiment, which products were actually on sale. At this point, participants were unaware of the purpose of the experiment and that further stages were to come. In the absence of logos, we refer participants' basket in stage 1 as the *reference basket*. In stage 2, one logo format is introduced and explained to the participants. Logos are then applied exhaustively to the 273 products and are visible online and in new catalogues (see appendix A2). Everything else remains unchanged. Participants are then invited to revise their reference basket by keeping, removing, adding or substituting the products selected in stage 1. This new basket built is called *logo basket*. Again, the task is incited: as in stage 1, participants are informed they would have to buy a significant sub-set of their basket but that only one of the two baskets (reference basket or logo basket) would be randomly chosen for the actual selling at the end of the session. Finally, participants are asked to fill a questionnaire in stage 3. At the end of a session, each participant bought one eighth of the products from their randomly drawn basket. They paid for these products at the prices posted in the catalogue and went home with them. Subjects spent on average less than €5.

4.4 Data analysis

This study aims at measuring the nutritional impact of different logo formats on food shopping baskets. We first measure the relative distance (in %) between the reference basket and the logo basket for each subject, and thus the changes, *ceteris paribus*, induced by the logo (*within subject* method). On this basis, one can measure the relative effectiveness of the seven logos by

comparing the extent of changes between logo formats (*between subject* method). This is possible because each treatment differs only in respect to the logo. We consider not only average distances but also individual dispersion. In particular, individuals who improve, do not alter and reduce the nutritional quality (perverse effect) of their baskets are distinguished. Within subjects analysis examines proposal 1 and between subjects analysis looks at proposals 2 to 6. Non-parametric tests are used. With matched data (distance between reference basket and logo basket per individual) we use the Wilcoxon matched-pairs signed-ranks (WSR). For unmatched data (distance between reference basket and logo basket per logo, option or subjects' characteristics), we use the Mann-Whitney test (MW). Finally, we use the Fisher Exact test (FE) when proportions are compared.

In order to estimate the nutritional quality of a shopping basket, the LIM score proposed by Darmon, Vieux, Mailliot, Volatier & Martin (2009) is used. The LIM score is a standard index used by nutritionists to estimate the mean percentage of the maximal recommended values for free sugar, salt and SFA. In other words, the LIM score averages the content per 100g of free sugar, salt and SFA weighted by the nutrients' daily maximal recommended values. It is calculated as follows: $LIM = 100 \times \{ (free\ sugar)/50 + (Sodium/3153) + (SFA/22) \} / 3$. Our change indicator between stage 1 and stage 2 is the distance in percentage between the LIM based on the reference basket and the LIM based on the logo basket.

5. Results

The statistical results are summarized in Table 2 for the impact per logo on the LIM score and in Tables 3 (a and b) for the impact per logo on food items. In the following, the 6 proposals of section 2 are tested.

In line with Proposal 1, the implementation of a logo induces significant nutritional improvement. As measured in the lab, logos improve the average nutritional quality of household baskets whatever the format. On average and over all the seven treatments, results exhibit a statistically significant (WSR, $p\text{-value}=0.000$) LIM decrease of -8.7%. 68.1% of the participants improved the nutritional quality of their basket, 12.1% of the participants remain unchanged and 19.8% reduced the nutritional quality of their basket.

GDA is neither the best nor the worst possible format. As expected by Proposal 2, GDA does not induce the greatest nutritional improvement. However it is the best among nutrient formats. GDA average performance ranks fourth among the seven formats with an average LIM decrease of -10.6%. With it 76.9% of participants improved their nutritional performance while 15.5% deteriorated it. GDA is better than the other two nutrient logos (NCG and NOG): its overall impact is significantly higher (-10.6% against -4.8%) and perverse effects affect significantly fewer participants (15.4% against 25.5%, F, $p\text{-value}=0.094$).

Logo format does matter. Consistently with Proposal 3, the 6 choice logos tested reveal highly contrasting nutritional impacts. Behavioural responses strongly depend on the format implemented, and some are much more efficient than others. The two nutrient choice logos are the least effective; the two traffic light logo formats are the most effective (top graph in Figure 1).

Product logos induce greater nutritional impact compared to nutrient logos (second graph in Figure 1). This validates Proposal 4. When the entire product is qualified with an aggregated index, a format is twice as efficient as when each nutrient is qualified separately: The overall impact on the LIM is -10.2% on average for the two product green logos (POG and PCG) and -4.8% for the two nutrient green logos (NOG and NCG) (MW, $p\text{-value}=0.000$). Product logos induce the largest improvement (LIM decrease over -20%) for a greater proportion of

participants: 81.1% against 53.1% for the nutrient logos (FE, p-value=0.000). With the latter more participants do not change their basket – neutral effect – (21.4% vs. 6.3%, FE p-value=0.001) and even more participants decrease the nutritional quality of their baskets (25.5% vs. 6.3%, FE p-value=0.011).

Reference set options induce balanced nutritional impact (third graph in Figure 1). This contradicts Proposal 5. The difference between category logos and overall logos is not significant. On average overall logos induces a -7.5% LIM decrease and category logos a -6.8%. While their overall LIM decrease is similar, category logos and overall logos have generated different changes in purchasing behaviours between carts: Category logos induce more substitutions than overall logos, but most substitutions with category logos are intra-category substitutions and most substitutions with overall logos are inter-category substitutions. On average subjects substitute 2.9 products with category logos and only 1.7 substitutions with overall logos. However 78% of substitutions with category logos – compared with 48% with overall logos – take place within the same categories and such category substitutions generate weaker nutritional gains per unit of substitution. Both options generate perverse effects but for different reasons. With overall logos, participants with perverse effects are clearly non-compliant: they decrease the share of green products. With category logos participants with perverse effects behave with compliancy: the frequency of green products still increases. However, such behaviour would have led to a decrease in green products with an overall option as participants have a tendency to replace bad products from good categories with good products from bad categories (see Table 3B for details).

In line with Proposal 6, Traffic light logos induce greater nutritional impact (bottom graph in Figure 1). On average traffic-light logos are nutritionally more effective than green logos. Nevertheless traffic-light logos also induce nutritional deterioration of baskets for a higher

proportion of participants comparatively to green logos (23% against 13% - FE, p -value=0.070). The overall difference between the two options is 40%: Traffic light logos lead to an average LIM decrease of -14.2% against -10.2% for the green logos. The difference is significant according to average, but not according to rank (t-test, p -value=0.070; Mann Whitney, p -value=0.424). On the one hand traffic light logos are more efficient for subjects with favourable effects (LIM decreases by -21.5% against -13.7% with green logos - MW, p -value=0.080). On the other hand, unintended effects are greater with traffic light logos (LIM increases by +6.0% against +4.9% (MW, p -value=0.084). Traffic light logos generate more extreme effects both ways, more favourable effects but also more perverse effects.

6. Discussion, limitations and conclusion

The present experimental study shows that logos have an impact on the nutritional quality of shopping carts. The extent of the impact must however be mitigated given the non-natural environment induced by the laboratory. Due to the 'before-after' design of the experiment, logos are made salient and the nutritional purpose of the experiment becomes obvious to the participants after the first phase and may cause social desirability bias in responses despite the monetary incentives. This may diminish the external validity of our results: In real setting (supermarkets), shoppers' attention and time are limited due to a stressful environment. Hence, results have to be taken as an upper limit behavioural response. In their studies based on sales data, Sutherland, Kaley, and Fischer (2010) and van't Riet (2013) showed significant but small changes. In their natural field experiment, Borgmeier and Westenhoefer (2009), Sacks, Rayner & Swinburn (2009), Steenhuis et al. (2010), Vyth et al. (2011) and Vermeer et al. (2011) did not find any visible effect. Nevertheless, these studies focused on a restricted set of items and are unable to extrapolate their findings if the labelling would cover a wider range of products. Furthermore, the main purpose of the present study is to appraise the relative performances of

logos according to specific criteria that define them. In that prospect, the potential lack of external validity of our experimental setting is here compensated by strong internal validity: By allowing *ceteris paribus* conditions, the experiment controls for individual and environmental variance.

The criteria tested are the unit graded, the reference set and the scope of colour. Our most salient result is that a simple choice-logo format that assesses the product rather than its multiple nutrients has a better nutritional impact. Simple labels are preferred (Feunekes, Gortemaker, Willems, Lion & van den Kommer, 2008; Möser, Hoefkens, Van Camp & Verbeke, 2009) and require more efforts (Feunekes, Gortemaker, Willems, Lion & van den Kommer, 2008). Nonetheless, GDA, despite its complex underlying heuristic, exhibits good performance in the present study. While less effective than synthetic logos, GDA display better results than nutrient-based choice logo. Kelly et al. (2009) and Balcombe, Fraser & Falco (2010) for instance show that a multiple traffic light system induce healthier choices than GDA. Unfortunately, this former system has not been tested in the present study.

Whether the products are rated with reference to its product category or rather to the overall product set do not seem to make a difference on the averaged global nutritional impact. When looking at the individual level however, behavioural responses to these two options are significantly different. Leading to intra-category substitutions, category logos generate larger responses. On the other hand, intra-category substitutions have a smaller nutritional impact than inter-category substitutions. With quantity and quality effects cancelling each other out, the global efficiencies of the two options are eventually similar. The choice of the reference set could however make a difference through the behaviours of firms. We may think that category logos would imply more competition and thus be more inclined to enhance the nutritional

quality of food. While Roberto et al. (2012) and Vyth et al. (2010) underline the importance of labelling on product development, this is out of reach of the present paper.

Finally, traffic light logos that signals both best and worst products have, on average, a better effect on nutritional quality than ‘only green’ logos. Nonetheless traffic light formats have significant unintended effects on a large number of participants. These unintended effects are not only due to non-compliant responses. For instance, we have seen that participants could be compliant to category logos by increasing the consumption of products labelled ‘green’ and decreasing the consumption of products labelled ‘red’ (where applicable) and still deteriorate the nutritional quality of their baskets. Along compliancy refusal, food categorisation may cause perverse effects when one substitutes red products from healthy categories for green products from unhealthy ones. Nevertheless, our data does not allow to satisfactorily rationalizing why the introduction of red logos exacerbates such outcomes. Grunert, Wills & Fernández-Celemín (2010) underlines potential over interpretations from consumers of red coloured logos that could lead to inappropriate generalizations. By over interpreting the severity of red logos, consumers may yield to inappropriate inter-category substitutions. Another explanation involves health halos and the intuition that unhealthy food are tasty and conversely (Aaron, Mela & Evans 1994; Raghunathan, Naylor & Hoyer 2006). If consumers indeed implicitly correlate negatively health and taste, red-labelled products hence become attractive for consumers whose healthiness of food is not a concern.

Many choices we have made concerning the experimental design may be questioned and criticized. For instance,

- By focusing exclusively on shopping behaviour, we excluded consumption, the eating process, which of course is the very determinant of a diet.

- Many other formats were possible, we stick to the one ‘in the air’ of French public authorities.
- Our 273-product selection was small (compared to real supermarkets) and certainly *ad hoc*, though selected very carefully, it limits the generalisation of our results. Our categories are also certainly *ad hoc* though sorted according to a standard (OQALI).
- Incentives may be judged insufficient, though meticulously calibrated and welcomed by participants. It was practically difficult to do more without counter-productive effects or rejection
- Nutrition back-of-pack panels were not available to participants.
- Our nutritional score – the LIM score – only takes into account bad nutrients (fat, sugar and salt) and overlook good nutrients such as vitamins, fibre, proteins, etc. This weakens any inference about the global quality of diet. Nutritional score that also includes fruit & vegetables, nuts, fibre and proteins such as the UK Ofcom Nutrient Profiling Model (Rayner, Scarborough, and Lobstein 2009) may have been more suitable to assess the global quality of diets.
- Practice effect or fatigue may possibly appear and affect the second task. We did not control for such artefacts (by for instance simply repeating the same task without logos). While possible practice effect or fatigue may influence the absolute impact of logos, the ranking should however not be affected since practice and fatigue are expected to be the same whatever the treatment ¹².

¹² If by any chance a logo induces more fatigue than another one, this effect is explicative rather than artefactual.

Despite these limitations, we have proposed here a novel experimental protocol capable of measuring, without noise and without hypothetical bias, the effect of global labelling policies that would be very difficult to evaluate in a real-world marketplace. We have obtained significant results that inform the future of the debate on front-of-pack logo formats. First, because logos matter and because the nutritional impact of their implementation is positive, one should certainly continue to support the induction of some sort of front-of-pack logo. Second, as changes in behaviour and induced changes in nutritional quality of food strongly depend on the constituents of the logo, logo formats matter.

Acknowledgements

This study was funded by the French Ministry of Health and the French National Research Agency (ANR - AlimInfo). We thank Nicole Darmon, Michel Chauillac and all participants of this study's Scientific Committee for precious comments. We also thank Jean-Loup Dupuis, Cédric Lanu, Mariane Damois, Coline Schneider, Jérémie Schneider, Alexandre Bechetoille and Vinciane Reboud for their help in setting up the experiment.

References

- Aaron, J., Mela, D., & Evans, R. (1994). The Influences of Attitudes, Beliefs and Label Information on Perceptions of Reduced-Fat Spread. *Appetite*, 22(1): 25–37.
- Andrews, J.C., Burton, S. & Kees, J. (2011). Is simpler always better? Consumer evaluations of front-of-package nutrition symbols. *Journal of Public Policy & Marketing*, 30(2), 175-190.
- Ariely, D. (2009). Predictably Irrational, The Hidden Forces That Shape Our Decisions. Harper Collins.
- Ariely, D. (2010). The Upside of Irrationality: The Unexpected Benefits of Defying Logic at Work and at Home. Harper Collins.
- Balcombe, K., Fraser, I., & Di Falco, S. (2010). Traffic lights and food choice: a choice experiment examining the relationship between nutritional food labels and price. *Food Policy*, 35, 211-220.
- Bialkova, S., & van Trijp, H. (2010). What determines consumer attention to nutrition labels? *Food Quality and Preference*, 21, 1042-1051.
- Borgmeier, I., & Westenhoefer, J. (2009). Impact of different food label formats on healthiness evaluation and food choice of consumers: a randomized-controlled study. *BMC Public Health*, 9, 184.
- Carlsson, F. (2011). Non-Market Valuation: Stated Preference Methods in *The Oxford Handbook of Food Consumption and Policy*, ed. Jayson Lusk, Jutta Roosen and Jason Shogren, 716-746. Oxford University Press.
- Carson, R., & Groove, T. (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics*, 7, 181– 210.
- Chan, C., Patch, C., & Williams, P. Australian consumers are sceptical about but influenced by claims about fat on food labels. *European Journal of Clinical Nutrition*, 59, 148-151.

Cowburn, G., & Stockley, L. (2005). Consumer understanding and use of nutrition labelling: a systematic review. *Public Health Nutrition*, 8, 21-28.

Crosetto, P., Muller, L., & Ruffieux, B. (2016). Helping consumers with a front-of-pack label: numbers or colours? Experimental comparison between Guideline Daily Amount and Traffic Light in a diet-building exercise. Forthcoming in *Journal of Economics and Psychology*.

Darmon, N., Vieux, F., Maillot, M., Volatier, J., & Martin, A. (2009). Nutrient profiles discriminate between foods according to their contribution to nutritionally adequate diets: a validation study using linear programming and the SAIN, LIM system. *American Journal of Clinical Nutrition*, 89, 1227-1236.

DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, 47, 315–372.

Drichoutis, A., Lazaridis, P., & Nayga, R. (2006). Consumers' Use of Nutritional Labels: A Review of Research Studies and Issues. *Academy of Marketing Science Review*, 9, 1-22.

Drichoutis, A., Nayga, R., & P. Lazaridis. (2011). Nutritional Labeling, in *The Oxford Handbook of Food Consumption and Policy*, ed. Jayson Lusk, Jutta Roosen and Jason Shogren, 716-746. Oxford University Press.

Elliot, A.J., Maier, M.A., Moller, A.C., Friedman, R. and Meinhardt, J. (2007). Color and psychological functioning: the effect of red on performance attainment. *Journal of experimental psychology: General*, 136(1), 154.

Feunekes, G., Gortemaker, I., Willems, A., Lion, R. & van den Kommer. (2008). Front-of-pack nutrition labelling: testing effectiveness of different nutrition labelling formats front-of-pack in four European countries. *Appetite*, 50, 57-70.

- Geiger, C. (1998). Health Claims: History, Current Regulatory Status, and Consumer Research. *Journal of the American Dietetic Association*, 98(11): 1312–1322.
- Grunert, K., & Wills, J. (2007). A review of European research on consumer response to nutrition information on food labels. *Journal of Public Health*, 15, 385–399.
- Grunert, K., Wills, J., & Fernandez-Celemin, L. (2010). Nutrition knowledge, and use and understanding of nutrition information on food labels among consumers in the UK. *Appetite*, 55, 177-189.
- Harrison, G., & List, H. 2004. Field Experiments. *Journal of Economic Literature*, 42, 1009-1055.
- Higginson, C., Kirk, T., Rayner, M., & Draper, S. (2002). How Do Consumers Use Nutrition Label Information? *Nutrition & Food Science*, 32(4): 145–152.
- INPES. (2008). Baromètre Santé Nutrition. INPES Edition.
- Just, D., & Payne, C. (2009) Obesity: Can Behavioral Economics Help? *Annals of Behavioral Medicine*, 38(1): 47–55.
- Kelly, B., Hughes, C., Chapman, K., Chun-Yu Louie, J., Dixon, H., Crawford, J., King, L., Daube, M., & Slevin, T. (2009) Consumer Testing of the Acceptability and Effectiveness of Front-of-Pack Food Labelling Systems for the Australian Grocery Market. *Health Promotion International*, 24(2): 120–129.
- Kelman, C. (1958). Compliance, Identification, and Internalization: Three Processes of Attitude Change. *The Journal of Conflict Resolution*, 2, 51-60.
- Levy, A., & Fein, B. (1998). Consumers' ability to perform tasks using nutrition labels. *Journal of Nutrition Education*, 30, 210-217.
- Kleef, E.V. and Dagevos, H., 2015. The growing role of front-of-pack nutrition profile labeling: A consumer perspective on key issues and controversies. *Critical reviews in food*

science and nutrition, 55(3), pp.291-303.

Lytton, T. (2010). Signs of Change or Clash of Symbols - FDA Regulation of Nutrient Profile Labeling. *Health Matrix: Journal of Law-Medicine*, 20: 93.

McCullum, C., & Achterberg, C. (1997). Food shopping and label use behaviour among high school-aged adolescents. *Adolescence*, 32, 181-197.

Möser, A., Hoefkens, C., Van Camp, J., & Verbeke, W. (2009) Simplified Nutrient Labelling: Consumers' Perceptions in Germany and Belgium. *Journal Für Verbraucherschutz Und Lebensmittelsicherheit*, 5(2): 169–180.

Muller, L., Lacroix, A., Lusk, J., & Ruffieux, B. (2016). Distributional Impacts of Fat Taxes and Thin Subsidies. Forthcoming in *Economic Journal*.

Muller, L., & Prévost, M. (2106). What cognitive sciences have to say about the impacts of nutritional labelling formats. Forthcoming in *Journal of Psychology and Economics*.

Philipson, T. (2005). Government perspective: food labelling. *American Journal of Clinical Nutrition*, 82, 262-264.

Raghunathan, R., Naylor, R., & Hoyer, W. (2006) The Unhealthy = Tasty Intuition and Its Effects on Taste Inferences, Enjoyment, and Choice of Food Products. *Journal of Marketing*, 70(4): 170–184.

Rayner, M., Scarborough, P., & Lobstein, T. (2009). The UK Ofcom nutrient profiling model.

Roberto, C., Shivaram, M., Martinez, O., Boles, C., Harris, J., Brownell, K. (2012) The Smart Choices Front-of-Package Nutrition Label. Influence on Perceptions and Intake of Cereal. *Appetite* 58(2): 651–657.

Roe, B., Levy, A., & Derby, B. (1999). The impact of health claims on consumer search and product evaluation outcomes: results from FDA experimental dat. *Journal of Public Policy and Marketing*, 18, 89-105.

Sacks, G., Rayner, M., & Swinburn, B. (2009). Impact of front-of-pack “traffic-light” nutrition labelling on consumer food purchases in the UK. *Health Promotion International*, 29, 344-352.

Steenhuis, I., Kroeze, W., Vyth, E., Valk, S., Verbauwen, J., & Seidell, J. (2010). The effects of using a nutrition logo on consumption and product evaluation of a sweet pastry. *Appetite*, 55, 707-709.

Sutherland, L., Kaley, L., & Fischer, L. (2010). Guiding Stars: the effect of a nutrition navigation program on consumer purchases at the supermarket. *American Journal of Clinical Nutrition*, 91, 1090-1094.

Szanyi, J.M., 2010. Brain food: Bringing psychological insights to bear on modern nutrition labeling efforts. *Food & Drug LJ*, 65, p.159.

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

van 't Riet, J. (2013) Sales Effects of Product Health Information at Points of Purchase: A Systematic Review. *Public Health Nutrition*, 16(03): 418–429.

Vermeer, W., Steenhuis, I., Leeuwis, F., Bos, A., de Boer, M., Seidell, J. (2011). View the Label before You View the Movie: A Field Experiment into the Impact of Portion Size and Guideline Daily Amounts Labelling on Soft Drinks in Cinemas. *BMC Public Health*, 11(1): 438.

Viswanathan, M., Torelli, C., Xia, L., & Gau, R. (2009). Understanding the Influence of Literacy on Consumer Memory: The Role of Pictorial Elements. *Journal of Consumer Psychology*, 19(3): 389–402.

Vyth, E., Steenhuis, I., Brandt, H., Roodenburg, A., Brug, J., & Seidell, J. (2012). Methodological quality of front-of-pack labelling studies: a review plus identification of research challenges. *Nutrition Reviews*, 70, 709-720.

Vyth, E., Steenhuis, I., & Heymans, M. (2011). Influence of placement of a nutrition logo on cafeteria menu items on lunchtime food Choices at Dutch work sites. *Journal of the American dietetic Association*, 111, 131-136.

Vyth, E., Steenhuis, I., Roodenburg, A., Brug, J., & Seidell, J. (2010). Front-of-Pack Nutrition Label Stimulates Healthier Product Development: A Quantitative Analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 7(1): 1–7.

Wansink, B. (2003). How do front and back labels influence beliefs about health claims? *Journal of Consumer Affairs*, 37, 305-316.

Wansink, B. (2004). Environmental Factors that Increase the Food Intake and Consumption Volume of Unknowing Consumers. *Annual Review of Nutrition*, 24, 455-479.

Wansink, B., Sonka, S.T. & Hasler, C.M. (2004). Front-label health claims: when less is more. *Food Policy*, 29(6), 659-667.

Wansink, B., & Sobal, J. (2007). Kitchenscapes, Tablesapes, Platescapes, and Foodscapes. Influences of Microscale Built Environments on Food Intake. *Environment & Behavior*, 39, 124-142.

Table 1: Three criteria, six options and 7 logo formats

		Options	
Criteria	Criterion 1. Units ranked	Product (P)	Nutrient (N)
	Criterion 2. Reference sets	Category (C)	Overall (O)
	Criterion 3. Scopes of colours	Green (G)	Traffic-light (T)
The 7 Logo formats			
1-GDA	<i>Guideline Daily Amounts</i> displays the percentage of daily-recommended intake values per serving for each of our three nutrients (salt, free sugar and saturated fat).		
2-NCG	<i>Nutrient Category Green</i> displays no symbol, 1, 2 or 3 green dots. An additional green dot is displayed (with the explicit nutrient name) whenever the product is among the best third of its category subset of products; the ranking appears three times, once for salt, free sugar and saturated fat respectively. Otherwise the logo is left blank.		
3-NOG	<i>Nutrient Overall Green</i> displays no symbol, 1, 2 or 3 green dots. An additional green dot is displayed (with the explicit nutrient name) whenever the product is among the best third of the overall set of products in terms of its contents in salt, free sugar and saturated fat. Otherwise the logo is left blank.		
4-PCG	<i>Product Category Green</i> displays no symbol or 1 dot. A green patch is found when the product is in the best third of its category subset of products concerning its average nutritional quality. Otherwise the logo is left blank.		
5-POG	<i>Product Overall Green</i> displays no symbol or 1 dot. A green patch is found when the product is in the best third of the overall set of products in terms of its average nutritional quality. Otherwise the logo is left blank.		
6-PCT	<i>Product Category Traffic-light</i> includes 1 dot that may be green (same as 4-PCG), or red when the product is in the inferior third within its category products subset. No dot is used when the product is in the average one-third group.		
7-POT	<i>Product Overall Traffic-light</i> includes 1 dot that may be green (same as 5-POG), or red when the product is in the inferior third within its Category products subset. No dot is used when the product is in the average one-third group.		

Table 2: Logos' global and individual impact per treatment

Logos	Average LIM decrease, as % of the reference LIM (standard deviation)	Individual change in LIM from reference to logo basket Percentage of participants in each category			
		$-20\% < \Delta < 0\%$ Improvement	$\Delta < -20\%$ <i>Large improvement</i>	$\Delta = 0\%$ Unchanged	$\Delta > 0\%$ Degradation
1-GDA	-10.6% (13.5)*	76.9%	19.2%	7.7%	15.4%
2-NCG	-4.4% (9.2)*	61.4%	5.7%	15.7%	22.9%
3-NOG	-5.2% (15.9)*	45.3%	12.0%	26.7%	28.0%
4-PCG	-9.7% (13.6)*	81.4%	13.6%	5.1%	13.6%
5-POG	-10.8% (16.0)*	80.8%	23.1%	7.7%	11.5%
6-PCT	-13.8% (17.5)*	72.4%	34.5%	3.4%	24.1%
7-POT	-14.6% (20.7)*	74.0%	33.3%	3.7%	22.2%
All logos	-8.7% (15.0)*	68.1%	17.0%	12.1%	19,8%

* Wilcoxon significant at 1%. In **red**: % of participants for whom the presence of a logo leads to a deterioration of the basket nutritional quality.

Tables 3:

a. Average number of items per basket in the reference basket and in the logo basket according to the seven treatments

Reference basket	GDA	PCG	POG	NCG	NOG	PCT	POT
Average number of items per basket in the reference basket	22.3	20.5	23.6	22.2	20.7	19.9	20.4
Logo basket	GDA	PCG	POG	NCG	NOG	PCT	POT
Average number of items per basket in the logo basket	20.4	20.1	22.2	21.9	19.8	18.4	19.0
Average number of items per basket kept from the reference basket	18.5	16.9	20.3	19.3	18.6	15.5	17.1
Average number of items per basket substituted within the same category	1.4	2.6	0.7	1.9	0.9	2.3	1.0
Average number of items per basket substituted across different categories	0.6	0.7	1.2	0.7	0.6	0.6	0.9

b. Changes in the number of products from reference basket to logo basket for the treatments POG, POT, PCG and PCT (as percentages of the number of products in the reference basket)

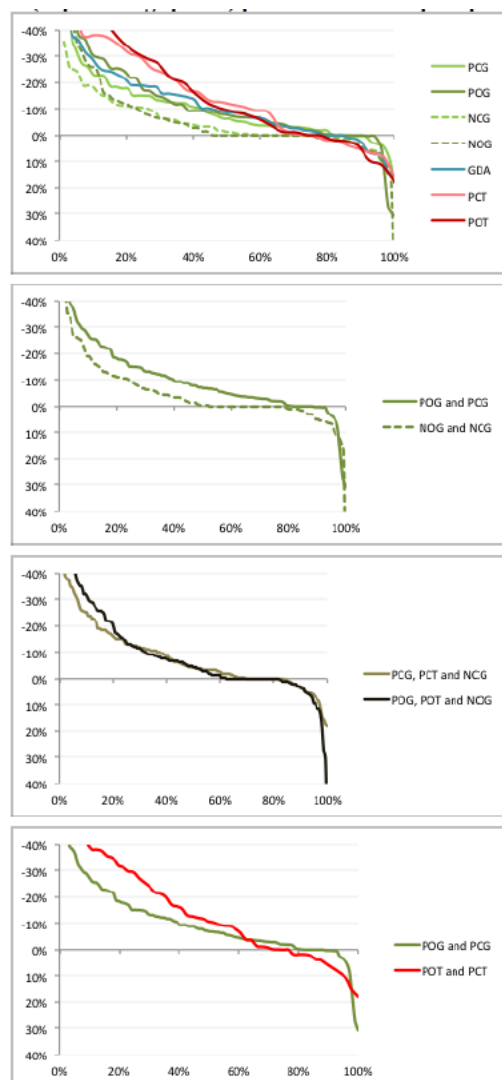
Behavioural Response	Changes according to coloured products with overall reference				Changes according to coloured products with category reference			
	Participants improving the nutritional quality of their basket		Participants lowering the nutritional quality of their basket		Participants improving the nutritional quality of their basket		Participants lowering the nutritional quality of their basket	
	Green	Red	Green	Red	Green	Red	Green	Red
POG	+24%	-18%	-8%	+3%	+7%	-16%	-6%	-3%
POT	+5%	-47%	-6%	-6%	+6%	-25%	+18%	-4%
PCG	+9%	-15%	-4%	0%	+62%	-32%	+37%	-3%
PCT	-2%	-20%	-2%	+3%	+37%	-56%	+39%	-36%

Bold type: changes directly suggested by the logo applied in the treatment.

Red type: changes in the direction opposite to that explicitly suggested by the logo applied in the treatment.

Bold frame: logo analyzed through its structure (e.g. overall logo formats analyzed through overall structure).

Figure 1: Distribution Function of the LIM's variation between the reference basket and the logo basket (in percentage points) per treatment and per option



Appendix A1. Food catalogue with no logo – Example of one page: the meat category.

Viandes

 2€64 STEAKS HACHÉS 15% MG 2 steaks de 125 g la pièce	 3€85 STEAKS HACHÉS PUR BŒUFS 5% MG 2 steaks hachés de 130 g la pièce	
 1€01 CÔTE DE PORC 1 tranche de 130 g en barquette	 8€91 NOIX DE VEAU À RÔTIR 1 tranche de 390 g en barquette	 1€98 ÉMINCÉS DE DINDE BASILIC Au basilic et huile d'olive. 1 boîte de 240 g
 3€86 GIGOT EN TRANCHE 1 tranche de 200 g en barquette	 4€72 ENTRECÔTE DE BŒUF 1 tranche de 230 g en barquette	
 2€80 CORDONS BLEUS DE VOLAILLE CUIITS ET SURGELÉS 1 sachet de 800 g	 2€17 ÉMINCÉS DE POULETS PANÉS Croustillants. 1 boîte de 200 g	

Appendix A2. Food catalogues with logos - Example of the meat category with the PCG logo system at the top left, PCT at the top right, NCG at the bottom left and GDA at the bottom right.



Appendix B. Description of the experimental sample per logo format

(Number of subjects; frequency of women; mean age; mean income per consumption unit per household; mean BMI; average LIM)

Logo format	Number of subjects	Sample composition			
		Women	Age	Income*	BMI
1-GDA	52	75%	39.5	1246€	22.2
2-NCG	70	82%	39.7	1326€	23.8
3-NOG	75	80%	40.5	1250€	24.3
4-PCG	59	74%	41.5	1290€	25.4
5-POG	52	90%	41.3	1155€	23.8
6-PCT	29	83%	43.0	1366€	23.9
7-POT	27	96%	39.2	1134€	23.2
TOTAL	364	82%	40.6	1258€	23.9

* The mean income is per consumption unit per household. A consumption unit is equalled to 1 for the first adult, 0.5 for the following individuals aged 14 or more and 0.3 for children under 14 years.