GUE/NGL

THE LEFT

IN THE EUROPEAN PARLIAMENT

EU-wide Food labelling:

Nutrition and Ecolabel Schemes

September 2021

Document Control Sheet

Title of the Study	EU-wide Food labelling: Nutrition and Ecolabel Schemes
Contractor (service provider)	Analysis Consulting Ltd
Authors	Stelios Gerontidis (Front -of -pack Nutrition Labelling Schemes), Tasos Michalopoulos (Ecolabel Schemes for Food Products)

Disclaimer and Acknowledgement:

This report does not pretend to be a scientific publication, and it does not strictly abide by academic rules concerning plagiarism. The report freely reuses pieces of text from existing reports and publicly accessible websites with relevant information. However, it does provide references to the sources used, although it does not always use quotation marks or other indicators when one or more sentences are literally used.

Contents

Executive Summary

The European Green Deal aspires to transform Europe into the first resource-efficient and climate-neutral continent by 2050. One of the priorities, in this regard, is to transform the entire food system, from production to consumption, to minimise its climate footprint and at the same time provide enhanced information to empower consumers make healthy, balanced, and sustainable dietary choices.

Front-of-pack nutrition labelling is simplified nutrition information provided on the front of food packaging aiming to help consumers with their food choices. Under the current EU rules, the indication of nutrition information on the front-of-pack is not mandatory but could be provided on a voluntary basis. The current EU rules on nutrition information will be updated to include a mandatory EU-wide front-of-pack nutrition labelling. While front-of-pack nutrition labelling schemes are implemented on a voluntary basis in several EU Member States, the transition to a unified and harmonised system at EU level will be complex.

Front-of-pack nutrition labels (FOPNL) are a crucial intervention to support healthy choices and reduce obesity rates by communicating complex nutritional information to consumers in a way that's easy to understand. This can reduce dietary intakes of calories and nutrients, such as salt, saturated fat, and sugar which, if overconsumed, have been linked to chronic disease.

FOPNLs are proposed to improve dietary quality through two main pathways: improving consumer understanding (reliant on consumer engagement) and driving product reformulation.

For the purposes of this study, FOP labels were divided into 5 major categories

1. **Reference Intakes and similar schemes**. These are labels that contain a lot of information about the quantities of food ingredients and let the consumer decide if the food is good for his health, without it being evaluated by the FOP label system.

2. **Color-coded nutrient-based schemes**. In this group of FOP labels, the label has colours depending on the amount of ingredient in the food. These are labels that are more understandable to the consumer and include some "evaluation" of the food.

3**. Overall rating schemes**. In this type of label, the system by which they are created contains a food evaluation algorithm according to the quantities of its ingredients and the consumer is informed through the label about the final result.

4. **Endorsement schemes ('positive logos').** These types of labels are also based on a food evaluation algorithm, but the consumer is informed only when the food is considered to be good for his health, since the food is "rewarded" by obtaining the specific label.

5. **Warning signs**. These are FOP labels that contain food evaluation and warn consumers when a food is considered not to contribute to a healthy diet. On the contrary, it contains in large quantities ingredients that the consumer should consume in small quantities.

The main conclusions from the analysis of the various FOP labels systems are:

Regarding **reference intake schemes**, we notice that several attempts have been made (most outside the EU) and all of them remain active until now. These are labels that are relatively difficult for consumers who do not have the relevant knowledge to understand. Consumers need to be "educated" in order to be able to evaluate which food they will choose.

However, these FOP labels do not involve the "subjective judgment" of any algorithm, so they are fair labelling systems as they list food ingredients in detail and allow consumers based on the knowledge and information, they have to choose both which food to consume and in what quantity. Most have the portion as reference (which is a more reliable than 100g) and are also created according to the rules of a Directive.

Colour-coded nutrient-based schemes are more understandable to consumers because of the

labelling of ingredients in three colours (red, orange, green). An evaluation of each ingredient is applied separately, and this is useful for the consumer, since for health reasons he may want to avoid consuming a large amount of a particular ingredient. It is also positive that some of them have the portion as a reference.

Studies have shown that the fact that there are three colours does not usually encourage consumers to increase their consumption of "green" coloured food but leads them to avoid products with "red" and "orange" colours.

Overall rating schemes are currently the most popular FOP labels in the countries that apply them. Their main advantage is that they are very easily understood by the consumer, since they give him a unique information, that of the "evaluation" of the food and in also coloured with characteristic traffic light colours. This way of labelling helps a very large number of consumers who are not "trained" to decide on the foods they will consume based on the FOP label.

However, there are two main disadvantages to this type of labelling. The first is a reference to the amount of "100g" for the evaluation of the individual ingredients and then the food as a whole. This is not a "fair" system for foods that are consumed in quantities much less than 100g and contain ingredients that are beneficial to health when consumed in such small quantities. The second, is the fact that the consumer is informed only about the final result of the food's evaluation, without being informed about the rating of the individual ingredients. It is also difficult for a consumer to be informed and understand the algorithm with which the final food is evaluated.

In this way it is possible to challenge the rating system. Consumers are also not "educated" in information and critical thinking about the foods they consume and their eating behaviour in general.

Endorsement schemes ('positive logos') are quite developed in number both inside and outside the EU. They are obviously very popular because they do not require "education" of the consumer and are also based on the evaluation of food by the relevant algorithm.

The specific types of FOP label are not given a "negative" rating, but the food is "rewarded" with the positive logo as long as it meets the nutritional value requirements according to the algorithm of each system.

Despite this, the diet is actually at risk from a system purportedly designed to help boost public health by allowing consumers to make more informed choices. With numerous studies indicating that consumers find current nutritional labelling confusing and a barrier to eating a healthy diet, it is not surprising that the EU wants to implement a harmonized front-of-pack (FOP) labelling system by 2022.

There is no such thing as a bad food or a good food, there is no such thing as a food that can be eaten and one that cannot. We need to focus on the quantities as related to the overall food and nutritional daily intake. It is misleading to believe that to combat the obesity epidemic, for instance, we should remove fat or sugars. We need to focus on awareness and education and not on simplification. We cannot hope to win over obesity by misleading people, we need the exact opposite.

It is also suggested that it would be good for the FOP labels to be colourful, because this attracts the attention of the consumers. The very simplified FOPs are easy to read by consumers, but because they are "black box" it is not clear to the consumer how the ratings of each food are obtained.

The best practice is to "train" consumers to read and understand food ingredients labels, that will lead to more informed choices on which food to consume and in what quantity. By understanding the ingredients' data and especially the %RI, consumers will be able to evaluate not only whether they will buy a food or not but also in what quantity they will include it in their daily diet.

Traffic light FOP label is very easy for the consumers, but red (much more) and orange (less) "prevent" them more from consuming a food that falls into this category, than "encourage" them with green colours to consume foods belonging to the relevant category. There are no "bad foods", there are "bad amounts" of eating a specific food product.

Maybe in correspondence with the "enjoy responsibly" in the drinks, there should be an indication

"consume the right amount" on the food packages. By this logic, a FOP could be adopted that would be "traffic light" regarding the daily consuming amount. This can be calculated depending on the level the specific food product "charges" the "battery" regarding the "ideal", or "maximum" or "minimum" amount that a person should consume of a component during the day.

For example, the FOP label should prevent the consumer from eating the largest percentage of his daily fat from a snack and to "determine" the maximum amount to be consumed from that particular snack per day, without generally making it a "forbidden" food.

The average daily intake needs of ingredients are for the "average adult". But in reality, there is no "average consumer". Children have other needs; the elderly have other needs. Maybe it would be good to also refer on the FOP label the up or down %RI deviations for large groups of consumers (children, the elderly, etc).

It is, therefore, advisable to separate the ingredients that are harmful to health when a daily intake is exceeded (such as fat, or sugar) from the ingredients that have beneficial effects to the body and which, in a balanced diet, should be consumed at least a minimum amount.

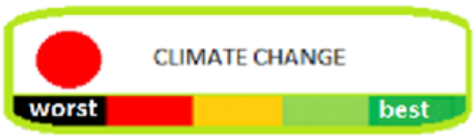
That is, if the FOP will be in a form of "battery" to warn when the battery is "overcharged" for "bad" components and "urge" to "charge" the battery more when it refers to "useful" components.

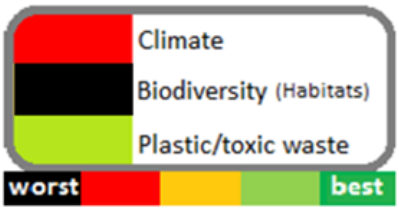
Ecolabeling helps identify food products that have a reduced environmental impact throughout their life cycle, from the extraction of raw material through to production, use and disposal. Recognised throughout Europe, The EU Ecolabel is a voluntary scheme that forms part of overall EU policy to encourage more sustainable consumption and production. The EU Ecolabel scheme is a commitment to environmental sustainability. The criteria have been developed and agreed upon by scientists, NGOs, and stakeholders to create a credible and reliable way to make environmentally responsible choices.

A detailed analysis of the Eco-Score labelling system in terms of relevant functionality features is provided in the report.

Label type Functionality	Eco-Score *(Compulsory version with product differentiation within categories is assumed)*	CDCC
Market-based	no	yes
Continuous incentive for improvement across entire range of substitutable products	no	yes
Incentivizes improvement above top grade	no	yes
Explicit focus on climate change	no	yes
Incentivizes virtuous cycle	**no**	**yes**

A minimal amount of information must be communicated for the CDCC label to be functional.





The present report proposes for consideration an innovative labelling system as a synthesis of functional features of available labelling approaches. The proposed labelling system is Compulsory Dynamic Categorical and Color-coded (CDCC). Contrary to the established eco-labelling paradigm, the information provided by the proposed labelling system does not merely aim to raise a part of production to a higher environmental standard, which in the best cases is sustainable. Instead, the CDCC label is designed to serve the more ambitious labelling goal to establish for environmental attributes the full effect that normal market forces have on the evolution of observable (non-credence) attributes that belong to the more-is-better type (similarly as the less-is-better type, the more-is-better attribute type has a clear normative direction). Namely, to engage the market in a consumer-driven and open-ended race for continuous improvement. To exemplify, consider the evolution of the processor capacity attribute in electronic devices.

Similarly, to other attributes of the same type, the implicit consumer acceptability threshold for processor capacity is dynamic and market defined. It is set by the evolution of processor speed among marketed products, and it is continuously raised by suppliers that compete in investing, inventing, and incorporating innovative technology to improve the speed of their products. In this continuous race for improvement, novel improvements by competitors effectively downgrade in the eyes of consumers the performance of products that until that moment were regarded as satisfactory. This inherently 'free-market' dynamics is spontaneously at work in markets for observable attributes: Products and firms must continuously innovate and improve, or else they will become obsolete and will exit the market. Although this dynamic is business as usual in markets for observables, it sounds alien in markets for environmental attributes that are ruled by the dominant eco-labelling paradigm.

The proposed CDCC labelling approach disrupts the established environmental labelling paradigm by incentivizing for environmental attributes the same market optimization dynamic that applies to observable ones. Simultaneously, it disrupts the established paradigm of negative environmental externalities in the supply sector: It is designed to engage product substitutes into a continuous, market-based, and consumer-driven race for environmental improvement. The incentivized race is open-ended towards the positive direction. As driven by consumer demand and enabled by environmental innovation, it is free to proceed beyond zero impact and into the territory of positive impact. Namely, into the territory of environmental restoration (e.g. carbon offset).

The proposed CDCC labelling approach is designed to use the market as an instrument against its own bad self. **That is, to initiate a continuous competition race in which normal market forces, 'the law of supply and demand', incentivize businesses to compete freely and transparently for environmental reputation in the market**. Towards this goal, it makes consumers aware of the comparative environmental impact of their consumption, so that the purchase of environmentally inferior products will need to happen knowingly, subject also to peer pressure and in light of the climate change effects observed.

Synergies in this process can be expected from related markets, such as the labour and finance markets. These can contribute effects that include reduced ability to recruit highly qualified employees, lower employee morale, and disinvestment by ethical funds, which can enhance the motivation of the worst polluters to improve.

In the light of widespread, rapid, and intensifying climate change (IPCC 9/08/2021), this report proposes that the historic goal of a welcomed EU-wide environmental labelling initiative should be nothing less than to unleash the full force of market forces to incentivize rapid, systemic, tangible, widespread and sufficient environmental improvement.

The disruptive functionality offered by the proposed CDCC labelling approach makes it a powerful tool for achieving this goal, as dictated by science, proclaimed in international and global initiatives, and as owed to life on earth as we know it, including the future generations of humanity.

A significant practical difference that distinguishes the labelling approach recommended in this report for environmental impact (especially GHG emissions) from that recommended for nutritional content is inherent in the nature of these attributes: The ideal content in most food ingredients is often a matter

of optimal dose and also varies with the physiology of the consumer. In contrast, the GHG emissions attribute exhibits a clear normative direction of the "less is better" type: it is always good to emit less GHG. It is better to emit none, and it is even best to be most negative in emissions, i.e. to absorb as much GHC as possible (e.g. through carbon offsetting). Furthermore, it is by now scientifically established and overwhelmingly accepted that the simultaneous deterioration of multiple environmental parameters poses a severe threat, with overarching catastrophic effects of human-induced climate change foreseen at all levels (social, economic, and also on the biodiversity and on broader ecosystems).

These parameters, namely the clear definition of the problem and the clear normative direction of the change needed mark a stark difference between the two issues analysed in this report. In a sense, when dealing with the environmental issue we have the more fortunate situation that the enemy, so to speak, namely greenhouse gasses - GHG, is clear and what needs to be done is beyond doubt as well: to reduce their concentration in the atmosphere as much as possible. This is particularly important in connection to food production, which amounts to about a quarter of total manmade GHG emissions globally. The problem, however, is not as simple when it comes to how to achieve this reduction in the present political and economic context. Despite decades of warnings, by now we have already missed the opportunity to avoid catastrophic effects of climate change and the time available to mitigate them also runs out.[1]

In this regard, the prospect of an EU-wide Front-of-Pack (FoP) nutrition and environmental labelling initiative, even if it applies only to food products, is certainly a move in the right direction. Product labelling is a tool long used in efforts to remove information asymmetry that prevents ethical consumerism from optimizing the levels of unobservable (credence) product attributes in the market. Such labelling initiatives succeeded to bring some improvement through product certification and also through positive spill over effects to the broader production sector.

Recommendations on Front -of -pack (FoP) Nutrition Labelling

There are no "good" and "bad" foods. All foods provide nutrients necessary to the human body. What is often "wrong" is the amount of food consumed. This is because the body is provided with large or excessive amounts of some ingredients with harmful effects on health. It would be good in foods that have a high caloric value or large amounts of specific ingredients to indicate with "emphasis" on the package maximum amount of consumption per meal or per day.

The proposed FOP label system and the methodology of "evaluation" of the food that it will provide should be based on international standards and regulations.

The system should be "appreciable" and use colour gradation, so that it is understandable and "attractive" to as many consumers as possible. It should be noted here that it is very important to "educate" consumers about nutrition so that they can understand even more complex labelling systems and be able to become the same "evaluators" of the food they are going to choose on the shelf. instead of blindly "trusting" the algorithm of an evaluation system.

The system should be "fair" towards food, using as a reference the portion (and not 100gr or 100ml) but also the percentage of supply of individual ingredients in relation to the maximum daily allowable amount.

Recommendations on Environmental labelling for Food Products

Environmental labelling systems should be mandatory. If labelling is optional, only the best products will be labelled and therefore the consumer about will not be informed on the less good

[1] "Climate change widespread, rapid, and intensifying – IPCC", The Intergovernmental Panel on Climate Change, Press Release 9 August 2021
https://www.ipcc.ch/site/assets/uploads/2021/08/IPCC_WGI-AR6-Press-Release_en.pdf,
The full IPCC Sixth Assessment Report is available at: https://www.ipcc.ch/report/ar6/wg1/

environmental products. Thus, if labelling is mandatory the "worst" polluters will be motivated to improve their environmental performance and it will also minimise the possibility of misleading consumers through greenwashing.

Environmental labelling systems should use color-coded multilevel Front-of-Pack marking, which is available in the form of a security sticker (sticker label). The security sticker form (i.e. with notches) facilitates frequent updates without changing the packaging, therefore the visibility is maximised and quick product comparisons without much effort from consumers is facilitated.

Environmental labelling systems should be "Unconcentrated" (at least in terms of greenhouse gas emissions). An overall environmental assessment (score) which covers all the environmental impact categories of each product should not pre presented. It is suggested the product score for the category Climate Change (Greenhouse Gas Emissions - GHG) to be presented separately from the other environmental impacts, in order to maximize the effectiveness of the labelling for this criterion.

Environmental labelling systems should be "Dynamic" and "Comparative" of the available products in each category. The final product rating should not be absolute (as in Eco score) but should correspond to / depend on existing alternatives: in this regard, the least environmentally friendly product in each category should receive the lowest rating (grade) by definition, while the best receives the highest (the other products are distributed intermediate). This maximizes the incentive to improve both the worst and the best products:

(1) Avoid the phenomenon that all available products have e.g. score A (see EU Energy label)

(2) Incentives are created for the best performers to invest in innovation in order to improve even more, moving the whole rating scale to the best score, and thus forcing other products to improve (virtuous circle)

(3) The periodic update of the product evaluation scale and product ratings can be done automatically, impartially, and objectively, based on the evolution of the performance of the best and worst products.

Environmental labelling systems should be updated at regular (frequent) intervals. To avoid the phenomenon observed in the case of the EU Energy label, where scale updates occur every ten years, where as a result the rating scale quickly becomes obsolete and ineffective, updates on environmental labelling systems should be annual.

The differentiation of the environmental impact of competing products during their grading, should be based on scientific data.: Although it is logical that some assumptions may be made for reasons of lack of data and cost reduction, it is important not to see a systemic bias against technologically innovative forms of production that could provide necessary solutions (e.g. carbon capture, synthetic meat, etc.), due to the tight attachment to specific production systems (e.g. organic, biodynamic) as Ecoscore does.

1. Background Information

1.1. Introduction - Scope of the study

The European Green Deal aspires to transform Europe into the first resource-efficient and climate-neutral continent by 2050. One of the priorities, in this regard, is to transform the entire food system, from production to consumption, to minimise its climate footprint and at the same time provide enhanced information to empower consumers make healthy, balanced, and sustainable dietary choices.

The environmental impacts of the production and processing of food, feed and drinks make up between 20% and 30% of the total environmental impacts of consumable goods in the EU.

Front-of-pack nutrition labelling is simplified nutrition information provided on the front of food packaging aiming to help consumers with their food choices. Under the current EU rules, the indication of nutrition information on the front-of-pack is not mandatory but could be provided on a voluntary basis. The current EU rules on nutrition information will be updated to include a mandatory EU-wide front-of-pack nutrition labelling. While front-of-pack nutrition labelling schemes are implemented on a voluntary basis in several EU Member States, the transition to a unified and harmonised system at EU level will be complex.

Ecolabeling helps identify food products that have a reduced environmental impact throughout their life cycle, from the extraction of raw material through to production, use and disposal. Recognised throughout Europe, The EU Ecolabel is a voluntary scheme that forms part of overall EU policy to encourage more sustainable consumption and production. The EU Ecolabel scheme is a commitment to environmental sustainability. The criteria have been developed and agreed upon by scientists, NGOs, and stakeholders to create a credible and reliable way to make environmentally responsible choices.

In this context, the proposed study aims to:

- identify and gather data on the various schemes currently in place in EU Member States. A comparative analysis will complement the compilation of data so as to present the pros and cons of each scheme in a comprehensible manner,
- present a policy recommendation about the appropriate front-of-pack system at EU level, based on science, with a view to providing real information to consumers in particular a label that comprises information vis-a-vis the nutritional value of the food product -also in the context of a healthy and balanced diet-, its health claims, as well its climate and environmental footprint.

2. Front -of -pack (FoP) Nutrition Labelling Schemes

2.1. Introduction

Front-of-pack nutrition labels (FOPNL) are a crucial intervention to support healthy choices and reduce obesity rates by communicating complex nutritional information to consumers in a way that's easy to understand.

This can reduce dietary intakes of calories and nutrients, such as salt, saturated fat, and sugar which, if overconsumed, have been linked to chronic disease.

FOPNLs are proposed to improve dietary quality through two main pathways: improving consumer understanding (reliant on consumer engagement) and driving product reformulation.

For the purposes of this study, FOP labels were divided into 5 major categories

1. **Reference Intakes and similar schemes**. These are labels that contain a lot of information about the quantities of food ingredients and let the consumer decide if the food is good for his health, without it being evaluated by the FOP label system.

2. **Color-coded nutrient-based schemes**. In this group of FOP labels, the label has colours depending on the amount of ingredient in the food. These are labels that are more understandable to the consumer and include some "evaluation" of the food.

3**. Overall rating schemes**. In this type of label, the system by which they are created contains a food evaluation algorithm according to the quantities of its ingredients and the consumer is informed through the label about the final result.

4. **Endorsement schemes ('positive logos').** These types of labels are also based on a food evaluation algorithm, but the consumer is informed only when the food is considered to be good for his health, since the food is "rewarded" by obtaining the specific label.

5. **Warning signs**. These are FOP labels that contain food evaluation and warn consumers when a food is considered not to contribute to a healthy diet. On the contrary, it contains in large quantities ingredients that the consumer should consume in small quantities.

2.2. Landscape of FoP Nutrition Labelling Schemes in Europe and beyond

2.2.1. Reference Intakes and similar schemes

[1] Facts-Up-Front

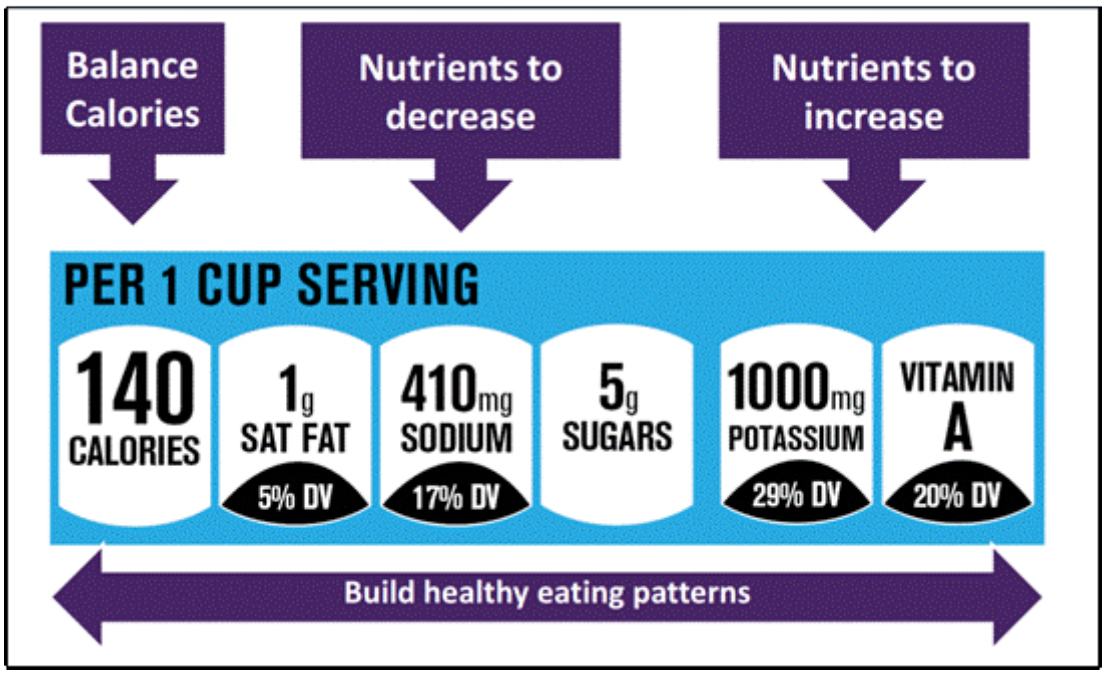
The "Facts Up Front" front-of-package (FOP) nutrition system is currently displayed on packaged foods in the United States. It summarizes important nutrition information from the Nutrition Facts Panel in a clear, simple, and easy-to-use format on the front of food and beverage packages.

The basic Facts Up Front label includes the below icons:

- Nutrients to limit as identified by the Dietary Guidelines for Americans.
- Consistent with the order of appearance on Nutrition Facts Panel.
- Specific serving size (same as Nutrition Facts Panel).

The Optional Icons consist of up to two additional plaques, representing specific additional nutrients required or permitted to be declared in nutrition labelling. These represent up to two "Nutrients to Encourage". Could be potassium, fiber, protein, vitamin A, vitamin C, vitamin D, calcium, iron. All are shortfall nutrients or are required to be on the Nutrition Facts Panel. The nutrients must contain 10% or more DV and be a "good source" to be featured.

It is a monochrome, quite complexity and reductive (non-interpretative) FOP label scheme.



The amounts of nutrients on the label are based on the "portion" suggested by the food manufacturer.

This initiative is being implemented by more than 50 manufacturers, retailers, and wholesalers in the United States on their branded and private label packaged food products.

Consumer awareness. Self-reported awareness of the Facts-Up-Front scheme was 62% among US Caucasians, 75% among Hispanics, and 60% among African Americans, while awareness rates tended to increase based on the education level.

Consumer understanding. The Facts-Up-Front scheme resulted in poor healthfulness judgement accuracy of more and less healthful cereals but better accuracy regarding frozen entrees (relative to chance). Better nutrition knowledge was associated with greater healthfulness judgement accuracy, even when less attention was paid to FOP labels. Attention to some specific nutrients (calories, fat, and sodium) was negatively correlated with healthfulness judgement accuracy, with this effect being more marked for individuals with less nutrition knowledge. Of note, the less complex FOP label (for entrees) performed better than the more complex one (for breakfast cereals).

Purely numerical, reductive FOP schemes can be confusing in cases where the nutritional information is equivocal. For example, 100% fresh orange juice might be richer in nutrients, but also higher in calories, compared to diet soft drinks. Kim et al. (2012) showed that reductive FOP labelling (Facts-Up-Front and Clear on Calories) made consumers rate milk and 100% fresh juice as less healthful and soft drinks and fruit drinks as more healthful compared to a no-label condition. They suggest that the negative information in 100% fresh juice and milk, namely the high calorie content, may have outweighed the positive information on the FOP label. An evaluative FOP scheme would seek to avoid this misunderstanding.

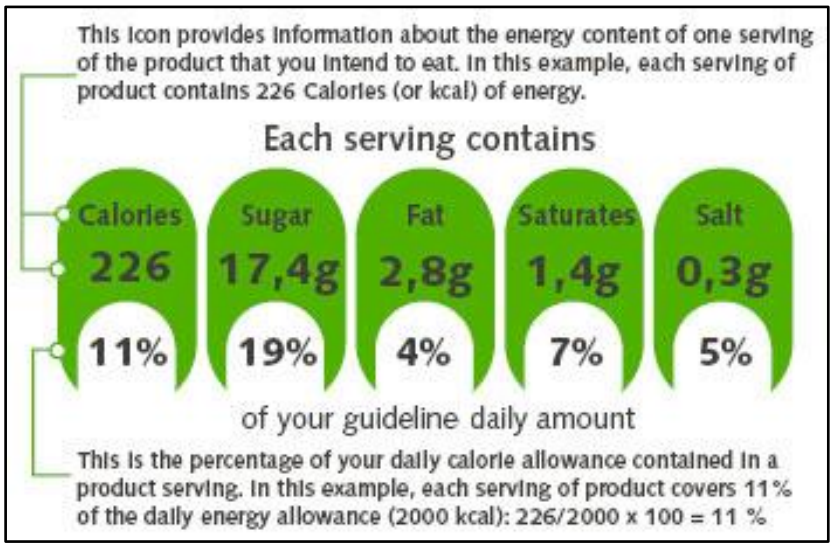
[2] Guideline Daily Amounts (GDA) - Reference Intakes

A Guideline Daily Amount (GDA) is a nutrition facts label that originally began in 1998 as a collaboration between the UK government, the food industry and consumer organizations. The process was overseen by the Institute of Grocery Distribution (IGD). To help consumers understand the nutrition information provided on food labels, science facts were translated into consumer-friendly

information, providing guidelines on package that help consumers put the nutrition information they read on a food label into the context of their overall diet. The GDA system was spread across Europe by some of the biggest international food companies, including big producers of sweets, chips, and soft drinks.

Based on new European rules, the new term "Reference Intakes" ("RI") instead of "GDA" is being introduced, however, the commitment for providing consumers with this information remained the same.

Key features. Nutrition information (energy plus four nutrients: fat, saturated fat, sugars, and salt) is indicated in grams and as percentage of daily reference intake. Portion is indicated as main reference base; 100 g or 100 ml as reference base for additional energy info. It is typically monochrome and constitutes a complex label, with much information, hence it is a reductive (non-interpretative) FOP label scheme.



GDA system is a guide to how much energy and nutrients are present in a portion of a food or beverage. So, the GDA label makes it easier for consumers to see what proportion of their daily nutritional needs is met by a particular food or drink and helps them to choose a balanced diet. The GDA system is intended to be a guide for consumers' orientation on the maximum amount of certain nutrients with negative impact on health, which they should consume daily.

Without using colour coding, GDA system provides information to consumers about the energy and also the content sugars, total fat and saturated fat and salt in a serving size, which is recommended by the manufacturer (CIAA (2006) recommends the serving size should reflect the amount of the product that can reasonably be expected to be consumed on an eating or drinking occasion). The GDA label also shows the contribution (in %) of the portion size to the reference values of energy, total fat, saturated fat, sugars, and salt recommended for an adult (Sanitarium Health & Wellbeing, 2011).

Consumer awareness. According to Bialkova & van Trijp, 2010, attention capture was faster and more accurate when the label was double the standard display size. Another study showed that consumers responded more quickly to color-coded than to monochrome GDA labels (Antúnez et al., 2015).

Babio et al., 2014, concluded that around 90% of respondents preferred a Multiple traffic-lights GDA label to monochrome GDA because it was perceived as more friendly and understandable; less than 3% preferred the monochrome scheme.

Consumer acceptance. Results from various studies indicate that health logos were easy to understand, highly accepted and useful for making decisions. Moreover, the logo was designed as endorsed by a credible institution, which gave the consumers greater confidence. GDA and Rating Stars came out as the least accepted FOP schemes.

Another survey was conducted to test consumer acceptance of four FOP labels: GDA, Multiple Traffic Lights, an early version of the Nutri-Score, the Green Tick, and a 'no label' condition. Acceptability was evaluated by several indicators: attractiveness, liking and perceived cognitive workload. The GDA label was rated as the most attractive and liked label (yet it was rated not easy to identify and understand). The Nutri-Score label, in turn, was rated as the easiest to identify and to understand rapidly (Ducrot et al, 2015a).

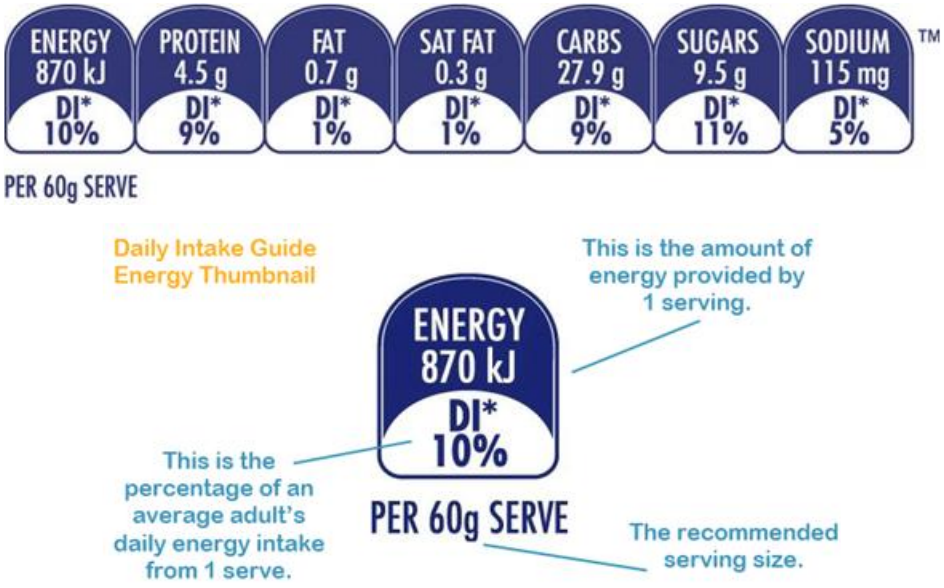
Consumer understanding. Most critics of the GDA system underlined the difficulty of understanding the nutritional information by consumers, particularly those with low level of education and pointed out that the scientific basis for the reference levels does not seem so clear. Because the GDA system does not reflect the quality of nutrients contained in a food product, focusing only on the quantitative aspect of the negative nutrients, most consumers could be confused[2]. Indeed, the GDAs are based on the nutrition requirements for an average adult of healthy weight and average activity level so several consumer segments can make uninformed choices like children, pregnant woman and elderly people.

[3] Daily Intake Guide (DIG)

The aim of the DIG scheme is to assist the food industry in providing key nutrition information in a recognisable, easy to use front of pack format which will assist consumers to make informed dietary choices by helping them to see the relationship between a serve of food and daily nutrition requirements. The scheme was designed during 2005 in Australia, where work involved collating research and methodology, and then in 2006, the Daily Intake Guide labels appeared on food and drink products in Australia and New Zealand.

Daily Intakes (DI), are a set of reference values for acceptable intakes of energy and a variety of nutrients, including protein, fat, saturated fat, carbohydrates, sugars, fibre, and sodium. DI values are based on an average adult's daily requirement of 8700kJ. A persons DIs may be higher or lower depending on his energy needs.

DIs are not recommendations, but rather an acceptable intake which provides a benchmark suitable for the majority of people. The reference values used for the Daily Intake Guide are based on those provided in the Australia New Zealand Food Standards Code (FSC). The FSC has outlined the composition and labelling requirements for food.



[2] www.foodpolitics.com

Key information on DIG labels includes amounts per serving of energy in Kilojoules, Fat, Saturated Fat, Sugars, Sodium (salt). Additional Nutrients that may appear on the label are protein, carbohydrates, vitamins & minerals. The label has 4 mail variations:

- Energy + 4 Key Nutrients (Fat, Saturated Fat, Sugars, Sodium)
- Energy
- Energy + 6 Key Nutrients, Vitamins or Minerals
- Energy + 1 Vitamin or Minera

The DIG scheme is a monochrome, complex label and the amounts of nutrients presented in it are based on the contents of 1 serving. It is a reductive (non-interpretative) FOP label scheme, since the consumer must self- evaluate the number of nutrient intakes from the specific food in relation to the suggested total daily amount.

Consumer acceptance. According to the study of Pettigrew et all, 2017, across the whole sample and among all respondent subgroups (males vs. females; adults vs. children; lower socio-economic status vs. medium-high socio-economic status; normal weight vs. overweight/obese), the Health Star Rating was the most preferred FOP scheme (44%) and the Daily Intake Guide was the least preferred (20%). Reasons related the test the acceptance rate include ease of use, evaluative content, and salience.

Consumer understanding. A couple of other studies [Christoforou et al., 2018; Devi et al., 2014] highlighted the use of FOP nutrition information (in a wide sense, including claims but also health logos and the Daily Intake Guide scheme) as marketing tools, since if applied to products with poor nutrient profiles, such nutrition information has the potential to mislead about the healthfulness of food products.

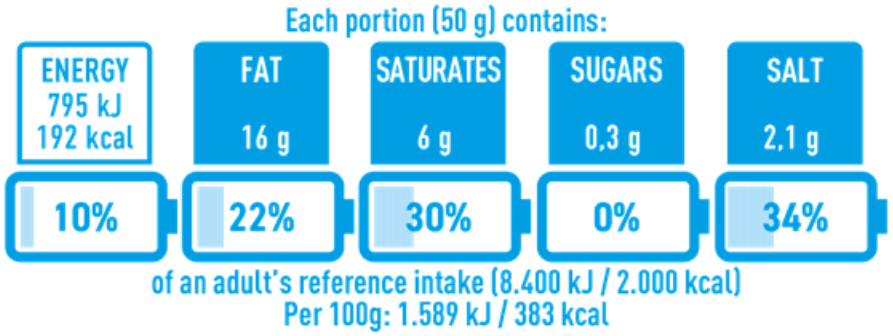
The variants of the traffic-lights labels increased the likelihood of consumers correctly identifying the more healthful food option (five-fold more than monochrome Daily Intake Guide label, and three-fold more than colour-coded Daily Intake Guide label). There were no differences in the number of correct responses between the monochrome and polychrome Daily Intake Guide labels. The traffic-lights scheme was particularly effective in identifying the more healthful option among consumers of lower socio-economic status (Kelly et all, 2009).

[4] Nutriform Battery

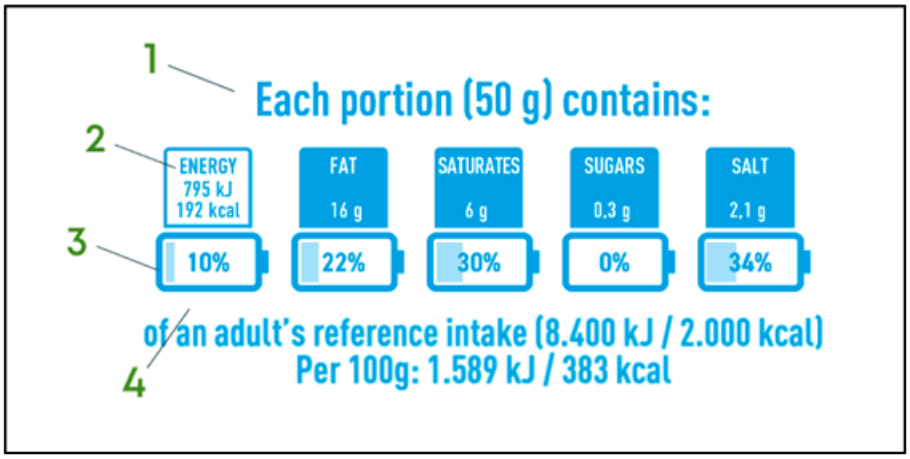
The Nutriform Battery FOP label was created in Italy, from the nutritionists of the Italian Institute of Health (ISS), the Food and Economic Research Council (CREA) and the representatives of category associations of food and agricultural products. This project's main stakeholders were four Ministries of the Italian government (Health, Economic Development, Agriculture and Foreign Affairs).

The reference parameters used to calibrate the single batteries (one for each single critical element, meaning calories, total fats, saturated fats, sugars and salt) are the European parameters set forth by table VIII of EU Regulation no. 1169/2011 – art. 35, which are also in line with LARN parameters (reference levels of nutrients for the Italian population), from which the portions determined on the basis of available scientific nutritional evidence are derived. The battery filling level corresponds in fact to the percentage of that specific nutrient that the recommended portion of the food brings to the consumer's diet, referring to the Reference Intakes (part B of Annex XIII of EU Regulation 1169/2011).

Nutrinform Battery is currently ready to be applied on a voluntary basis. It will be progressively implemented in Italy in the coming months, with a view to a possible harmonization at European level by the end of 2022.



Reading Guide



All the indicated values concern the single portion, and each box indicates the quantity of calories, fats, saturated fats, sugars and salt of the single portion. The energy content is expressed in Joule and Calories. The contents of fats, saturated fats, sugars, and salt are expressed in grams. Inside the ''battery'' symbol there is a percentage of energy, fats, saturated fats, sugars and salt contained in each single portion in relation to the suggested daily intake. The daily quantities recommended in EU are:

Energy: 8400 kJ / 2000 kcal

Fats: 70 g

Saturates: 20 g

Sugar: 90 g

Salt: 6 g

The charged section of the battery indicates graphically the percentage of energy or nutrients contained in the single portion, allowing to quantify it also visually. The sum of the food products consumed eaten during the day can "fill" the battery charge without surpassing a certain limit, in order not to exceed the suggested daily intake quantities.

The Nutrinform Battery labeling idea, considers proportionality, with no demonization of the food contents, including saturated fat, sugar, or salt. An example is that if a consumer consumes a serving of Parmigiano cheese the Battery will show how much that portion counts in the overall daily intake. Nutrinform allows consumers to identify the category of each nutrient and how much it "fills up the battery." In a nutshell, it is a monochrome, quite complex Reductive (non-interpretative) FOP label scheme.

2.2.2. Colour-coded nutrient-based schemes

[1] UK Multiple Traffic Lights (MTL)

The "Traffic Light" system (TL) was developed in UK, in 2006, by the Food Standards Agency (FSA) and enjoyed the support of many non-profit organizations from the food field. In creating this system, the main objective of FSA was to help consumers in making healthy food choices and overcome the difficulties that consumers had with the previous nutrition label systems (Food Standards Agency, 2008). At present, the TL system is used voluntarily, especially in the UK, only by some manufacturers and retailers.

The main features of this FOP label scheme are:

- Nutrition information (energy plus four nutrients fat, saturated fat, sugars, and salt) in grams and as percentage of daily reference intake.
- Traffic light colour coding indicating low (green), medium (amber) and high (red) levels of the nutrients stated.
- Portion as reference base for numerical information; 100 g or 100 ml as reference base for colour coding7 and additional energy info.
- Separate colour thresholds for solid foods and beverages.

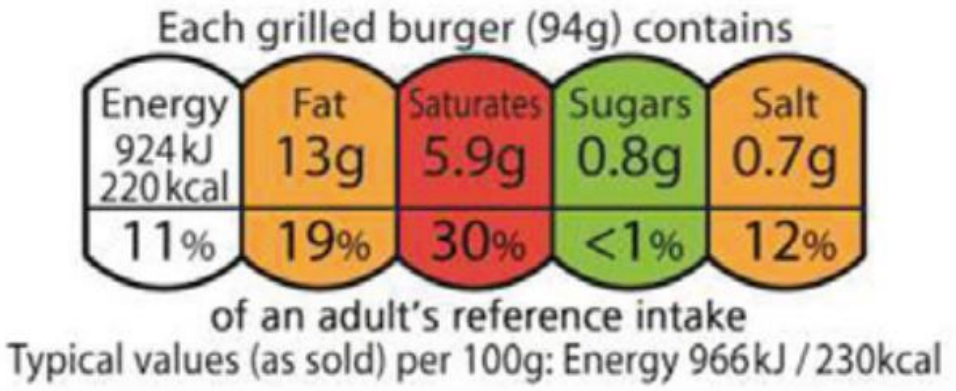
The nutrition information can be provided per 100g/ml only; per 100g/ml and per portion; or on a per portion basis only (applies only in the case of energy + 4). Where information is provided per portion only for the four nutrients (energy + 4), the absolute value for energy must be provided per 100g/ml in addition to per portion.

Percentage reference intakes (%RIs) can be given on a per 100g/ml and/or per portion basis.

Where %RI information is provided on a per 100g/ml basis, the statement 'Reference intake of an average adult (8400kJ/2000kcal)' is required.

The "Traffic light" system, combined with RI% use is described in detail below:

- The colour coded (red, amber and green) labels on the front of pack show at a glance if the specific food has high, medium or low amounts of fat, saturated fat, sugars and salt.
- Red colour on the front of the pack, does not mean that consumers should not or cannot eat it, but that they should try to keep an eye on how often they choose these foods, or how much of them they may eat. A diet with fewer reds can help consumers achieve a healthier diet.
- Amber colour, means that these foods are neither high nor low for that nutrient. Foods with ambers help consumers balance their diet.
- Green colour means that the food is low in a specific nutrient that a consumer may wish to avoid overconsuming to improve his diet. The more green colours, the healthier the choice, but consumers don't have to eat only green colour coded foods: including a few ambers and reds can be part of a balanced diet and will help them get all the beneficial nutrients they need.
- Many of the foods with front of pack labelling in the shops have a mixture of red, amber and green. So, when a consumer chooses between similar products, it is better to prefer more greens and ambers, and fewer reds.
- While colours provide information at a glance, the % RI information gives a little more detail about how much of an average adult's daily intake limit of each nutrient is in a portion and helps consumers put it in the context of a healthy balanced diet. For example, 50% RI of salt means that the serving contains half of an average adult's maximum daily intake for salt and so consumers should try to choose options lower in salt for the rest of the day.
- The %RIs also helps for more accurate comparisons between equal portions of products. A consumer can use the detailed RI information to help him choose between products that have the same colour per 100g/ml or the same portion size.



FSA, which has designed and developed this label design, provides a detailed manual to food manufacturers to instruct them on the correct colours (red, amber, green) and RI%. The table below shows a summary of how the colours are calculated for each nutrient.

Table 2: Criteria for 100g of food (whether or not it is sold by volume)

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

Note: portion size criteria apply to portions/serving sizes greater than 100g

The "Traffic Light" system is a multicolour, Evaluative (interpretative) and complex FOP label scheme.

Consumer attention. As concerns specific FOP schemes, studies show good attention-grabbing potential of Nutri-Score, Multiple Traffic Lights, and warning labels [Ares et al., 2018; Vidal et al., 2013]. A study from Mejean et all, 2013. Shows that the Multiple Traffic Lights fared best in terms of self-reported liking, acceptance, and attractiveness, although Simple Traffic Light and Green Tick (and to some extent the PNNS logo) also scored well on liking and several dimensions of attractiveness.

Consumer understanding. A study from Talati et all, 2016c, reported that two evaluative FOP schemes Multiple Traffic Lights and Health Star Rating were preferred over the reductive Daily Intake Guide. The two main considerations were trust and ease of interpretation. The FOP schemes were also more likely to be considered in the product evaluation than health claims (this was especially true of the Health Star Rating and Multiple-Traffic-Lights labels). Of the two evaluative FOP schemes, participants preferred the one with the summary indicator (namely the Health Star Rating). The Multiple Traffic Lights helped respondents better distinguish more and less healthful products. Labelling schemes without reference point information (e.g. nutrition table) were found less easy to interpret when no comparison product was available. (Van Herpen, Hieke & van Trijp, 2014). The Multiple Traffic Lights had some impact on willingness to pay (specifically for foods at either end of the healthfulness spectrum), while the Daily Intake Guide had no impact on this variable (Talati et all 2017a).

An important criticism against TL is that it may disadvantage certain categories of foodstuffs, for example dairy products, where it is difficult to avoid red colour for saturated fat, despite the presence of many beneficial nutrients (Wartella et al., 2011).

In addition, TL system doesn't provide guidance on the consumption frequency of a foodstuff in the overall diet. Indeed, Chandon and Wansink (2006) note that consumers who choose healthy foods according to nutrition labels, are willing to compensate the choice with a sort of indulgence, producing a negative impact in terms of dietary intake.

[2] Evolved Nutrition Label (ENL)

ENL label was a voluntary initiative that aimed to provide simple, easy-to-understand and meaningful nutrition labelling information on foods and beverages. While it was built on the existing EU-wide Reference Intake scheme, ENL added colors to the numbers that reflect the nutrient content per actual portion consumed.

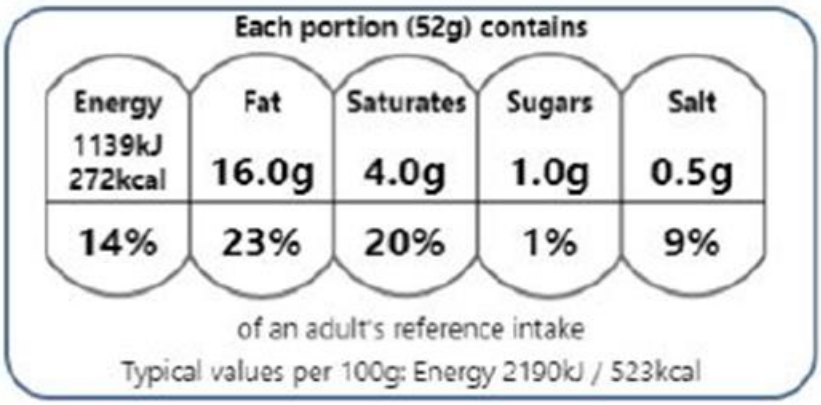
The Evolved Nutrition Label (ENL) scheme was initiated in 2017 by a group of multinational food companies. The approach builds on the Reference Intakes label and adds colours similar to the UK MTL scheme. However, the ENL uses less than 100 g as the reference base for assigning the colours amber and red for products considered to be consumed in small portions (e.g. sweet spreads, cookies).

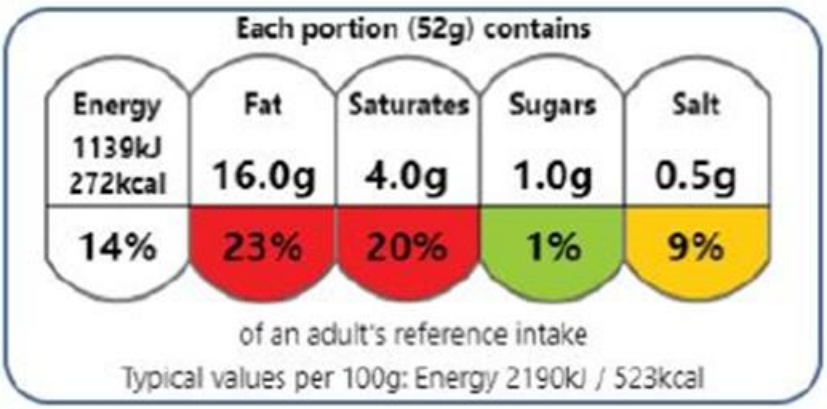
For the green colour, a per 100 g basis is applied to align with the 'low in' nutrition claims as per Regulation (EC) No 1924/2006.

In November 2018, companies involved communicated their decision to suspend/cease label trials for food.

The main features of this FOP label scheme are:

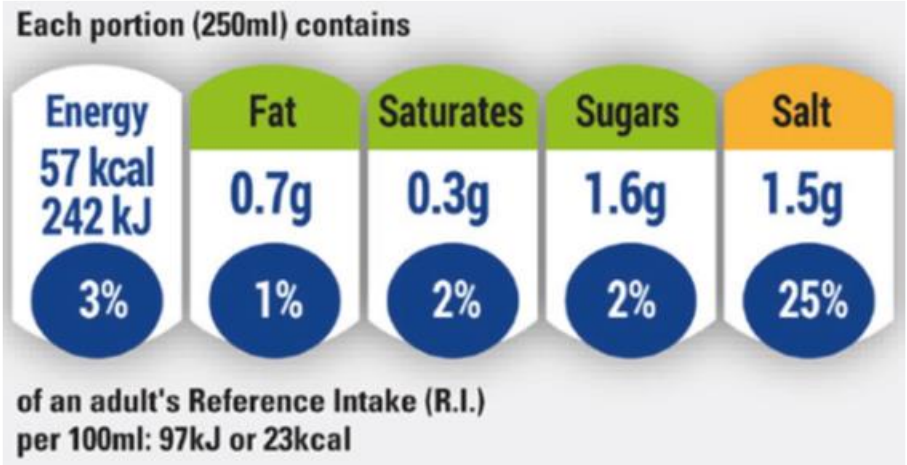
- Nutrition information (energy plus 4 nutrients) as percentage of daily reference intake, similar to the UK MTL scheme.
- Portion as reference base for both numerical information and colour coding (except for green colour where 100 g/ml is used as basis), plus energy per 100 g or 100 ml.
- ENL was a is a multicolour, Evaluative (interpretative) and complex FOP label scheme.
- How to interpret the Evolved Nutrition Label?
- For example: food label below shows that each portion of 52g will provide a consumer with 16.0g of fat and 4.0g of saturated fat, which are 23% and 20% of reference intake for fat and saturated fat respectively.
- The red colour shows that the food item is high in fat as well as saturated fat.
- This food product consists of 0.5g of salt which is equivalent to 9% of reference intake for salt.
- The amber colour shows that the item contains a medium amount of salt.
- Green colour means a food is low in a particular nutrient/ingredient. For example, this product is low in sugar content, as it only consists of 1.0g with reference intake of 1%.





According to the latest updates, the ENL companies acknowledge that the lack of EU-defined portion sizes has led to insufficient understanding and support of the proposed scheme. Along with this and the absence of legally defined portion sizes, the companies have decided to suspend ENL trails for food.

There have been concerns over ENL's efficacy and potentially misleading color codes. One example is how Nutella, a high-fat, high-sugar product, had no "red light" under the proposed ENL logo.



Consumer understanding

A study from Egnell et all, 2018b., reported that portions selected in the ENL condition were smaller than in the control group only for cheese, and significantly larger than the control for sweet spreads. There was no difference between ENL and control for sweet biscuits and across all products. The authors conclude that per portion FOP nutrition information might not help consumers choose healthier portion sizes and that instead 100 grams would be a more suitable reference base.

2.2.3. Overall rating schemes

[1] Nutri-Score

The Nutri-Score scheme developed under the aegis of the French Ministry of Health and implemented in France in 2017 and in Belgium in 2019. In November 2019, a total of 236 brands had adopted it and >35% of the market share in France. Nutri-Score adoption was also announced by Spain, Germany, the Netherlands, and Luxembourg.

Spread of Nutri-Score across Europe

According to research conducted in May 2018, 15% of the products in the French market were labelled with Nutri-Score. Large international brands and retail chains are also adopting the label. In June 2019, Nestlé announced its support for Nutri-Score as the chosen labelling system for food products in Europe. One month earlier, Albert Heijn, the leading retail chain in the Netherlands, announced adoption of Nutri-Score following a successful introduction at its Belgian brand Delhaize, the first retailer to adopt this food label in Belgium. Aldi Suisse also announced plans to introduce the Nutri-Score labelling system in 2020.

Description. Nutri Score displays five letters (A, B, C, D, and E), which correspond to a nutritional rating of the food from best to worst. A is coloured in dark green, B in light green, C in light orange, D in orange, and E in dark orange. The letter corresponding to the rating of the food is made larger than the four remaining letters. The general algorithm to calculate the score considers a food's content of energy, sugars, saturated fat, sodium, fruit, vegetables, legumes and nuts, fibre, and protein. Three updated algorithms apply for cheeses, beverages, and added oils/fats to improve alignment with French dietary recommendations for these food groups.

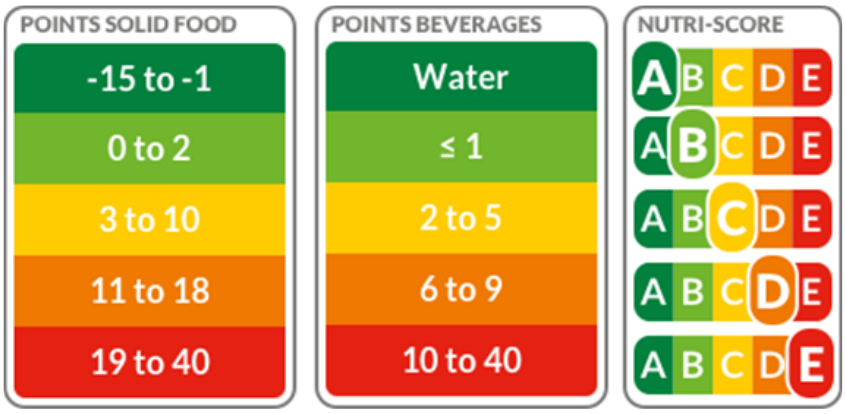
Nutri-Score is entirely transparent. There is even an implementation in Excel if someone wants to test it. Products receive positive points for fiber, proteins, fruits, nut and olive or rapeseed oil. They receive negative points for sugar, energy, saturated fat and salt. All values are expressed per 100 grams (or 100 milliliters) of product. The score goes from -15 to 40 and is further summarized on a 5-step scale that goes from A (green) to E (red). It is shown in front of the packaging, unlike the list of ingredients, which remains in the back.

- There is a graphic scale that divides the nutritional score into 5 classes (expressed by a color and a letter), based on the food's content of energy, sugars, saturated fat, sodium, fruit, vegetables, and nuts, fibre, and protein.
- Algorithm based on UK Food Standards Agency (FSA) Nutrient Profiling system; minor modifications to FSA score algorithm for cheese, added fats, and beverages have implemented in order to improve consistency between Nutri-Score classification and French nutritional recommendations.

This is how the Nutri-Score is calculated

The algorithm gives points for each element in the nutrition table (per 100 g or ml), that means bad nutrients (energy, sugars, saturated fatty acids, salt) as well as good nutrients (proteins, fibre, percentage of fruit, vegetables, nuts, rapeseed oil, walnut oil and olive oil).

We then subtract the positive points from the negative ones and convert the result to the Nutri-Score table.



Points	Energy (kJ)	Sugar (g)	Saturated fatty acids (g)	Sodium (mg)
0	≤ 335	≤ 4,5	≤ 1	≤ 90
1	> 335	> 4,5	> 1	> 90
2	> 670	> 9	> 2	> 180
3	>1005	> 13,5	> 3	> 270
4	> 1340	> 18	> 4	> 360
5	> 1675	> 22,5	> 5	> 450
6	> 2010	> 27	> 6	> 540
7	> 2345	> 31	> 7	> 630
8	> 2680	> 36	> 8	> 720
9	> 3015	> 40	> 9	> 810
10	> 3350	> 45	> 10	> 900
TOTAL	1 point	0 points	0 points	7 points

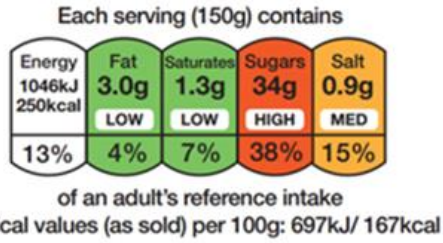
Points	Fruit, vegetables (%)	Fibers (g)	Proteins (g)
0	≤ 40	≤ 0,9	≤ 1,6
1	> 40	> 0,9	> 1,6
2	> 60	> 1,9	> 3,2
3	-	> 2,8	> 4,8
4	-	> 3,7	> 6,4
5	> 80	> 4,7	> 8,0
TOTAL	0 points	5 points	5 points

7 − 10 = -3 = A

Change in Nutri-Score calculation. To be more in line with current nutritional recommendations and based on scientific research by the French High Council for Public Health, the national public health agency Santé Publique France (initiator of the Nutri-Score) has partly adapted the method for calculating the Nutri-Score. The amounts of rapeseed oil, walnut oil and olive oil in a product are now taken into account in the calculation of the Nutri-Score. This means that the Nutri-Score can improve for certain products like olive oil and products that are high in oil (such as hummus and cold sauces).

Nutri-Score is a composite label, in which all nutritional information about a product is brought together in one simple score (A through E, with A being the healthiest option). The Traffic Light label, on the other hand, does not show a compound score but instead separate assessments for main ingredients (fat, saturated fat, sugar and salt) plus the food's energy content.

Compare Nutri-Score and Multiple Traffic Light labels. Source: Rabobank, 2018 (images Consumentenbond, 2018)



+ **Intuitive**: Easy to see and understand
+ **Simple**: A single score limits the amount of knowledge needed for interpretation
+ **Balanced**: Includes positive elements in the calculation, such as fibres or micro-nutrients
- **A black box**: Calculation not visible, so trust is required, and some consumers consider it an oversimplification

+ **Complete**: A lot of information in one label
+ **Transparent**: Contents fully visible
- **Difficult:** Small print, content requires knowledge
- **Complex**: No single, clear outcome
- **Redundant**: Most of the information is already on the back of the packaging

The French food manufacturers that initially supported Nutri-Score (Danone, McCain, and Fleury Michon) are producers of (semi-)fresh and relatively healthy products. In fact, Danone released a statement after the introduction of the Nutri-Score in Belgium, boasting that 70% of its products already have an A or B score, and the remainder a C on average. To these companies, supporting the Nutri-Score is thus relatively risk-free: no big red E's anywhere.

Source: Rabobank, 2018

One of the leading contenders, however, threatens to confuse consumers more than they already are, while dealing a body blow to some of the foods that are the cornerstones of the Mediterranean diet. The Nutri-Score system works by assessing the nutritional benefits and drawbacks of a product, then grades it with a color-coded rating from A to E. The catch? The system evaluates every item on a rigid

scale of 100g or 100ml and uses an oversimplified algorithm which does not account for certain ingredients which comprise an intrinsically healthy part of an overall diet, such as vitamins, minerals, and essential fatty acids.

The ample limitations of Nutri-Score's design mean that it unfairly maligns Mediterranean diet staples including olive oil, hard cheeses, and regional meats—even giving them lower marks than chocolate cereal or soft drinks. Understandably, this has prompted strong criticism from certain sectors, including Spanish olive oil producers, who dubbed Nutri-Score "incomplete, misleading, and as such, false."

Nutri-Score does not reflect an accurate snapshot of a food's healthful qualities because the algorithm classifies food on the basis of a standardized quantity, such as 100 grams or 100 milliliters.

In the case of a food item such as extra virgin olive oil, which receives a C-grade from Nutri-Score due to its fat content, the aforementioned quantities do not reflect realistic consumption levels.

No one will ever eat 100 milliliters of olive oil while dining. Maybe some could eat 100 grams of oats, but in just a small spoon of extra virgin olive oil we find polyphenols and many other compounds which are essential for our own health.

The calculation algorithm has given too much weight to fats, this in combination with the fact that the calculation base is 100gr makes it possible to eliminate any good nutrients of a food if it has a large amount of fat and is consumed in small amounts. In this case the food even if it has a lot of good nutrients will rank in bad grades.

In a Nutri-Score environment, we could even find sugar-free carbonated drinks that rank better than olive oil because, once again, not all of their contents are considered, since they are packaged in containers far larger than the standard 100-milliliter measurement used by Nutri-Score.

Some experts believe that if a product is labelled with the warning colours and low scores employed by Nutri-Score, consumers will simply select foods labelled with an A or B grade and ignore the contents of the item.

The consumer will not buy a product labelled C, simply because it is labelled that way. Consumers are not used to making a comparison between foods of the same category, but their attention is mainly captured by the absolute score.

Perhaps more worryingly, a food company could improve the score of its products by replacing sugar with artificial sweeteners. Diet sodas have a B score, for instance, while most natural fruit juices have a C score. Other molecules, like antioxidants, are likewise disregarded by Nutri-Score.

Consumer understanding

The GDA label was rated as the most attractive and liked label (yet it was rated not easy to identify and understand). The Nutri-Score label, in turn, was rated as the easiest to identify and to understand rapidly. (Ducrot et all, 2015).

The Nutri-Score was the most preferred FOP scheme, followed by Multiple Traffic Lights and the SENS scheme. Conversely, the modified Reference Intakes yielded the highest number of responses on negative dimensions of perception (complexity and time processing). (Julia et all, 2017).

According to Engell et all, 2018c, all five FOP schemes that examined in their study, improved the number of correct responses over the 'no label' condition. Improvements were most pronounced for the Nutri-Score, followed by Multiple Traffic Lights, then Health Star Rating and warning signs with almost equal effect, and finally the Reference Intakes. When analyses included only participants reporting to have seen the label during the survey, the Nutri-Score and the warning sign resulted in the highest level of improvement compared to the Reference Intakes label. Trends were similar for individual product categories and all products together.

All labels were found to be effective in allowing consumers to identify more healthful products compared to a 'no label' situation. The Nutri-Score performed best (Odds Ratio (OR) 12.61), followed by Multiple Traffic Lights (OR 8.71), GDA (OR 7.74), and the Green Tick (OR 2.36). These findings did not vary across socio-demographic characteristics. (Ducrot et all, 2015b).

[2] Health Star Rating

The HSR system was implemented in Australia and New Zealand in June 2014, and is jointly funded by Australian, state and territory and New Zealand governments.

The Health Star Rating (HSR) is a voluntary front-of-pack labelling system that rates the overall nutritional profile of packaged food and assigns it a rating from ½ a star to 5 stars. It provides a quick, easy, standard way to compare similar package ed foods. The more stars, the healthier the choice.

The number of stars is determined using a calculator designed to assess positive and risk nutrients in food (The Health Star Rating Calculator). The algorithm that drives the calculator was developed in consultation with Food Standards Australia New Zealand and other technical and nutrition experts.

Food manufacturers and retailers are responsible for the correct and accurate use of the Health Star Rating system. This includes correctly calculating the Health Star Rating, accurately displaying nutrient information, ensuring consistency of information between the Health Star Rating and the Nutrition Information Panel, and complying with all relevant legislation and regulations.

Under the system, packaged products are given a rating based on their nutritional profile, according to a strict calculation (called the Health Star Rating Calculator). The calculation takes into account components that are linked to increased risk of developing chronic diseases, as well as beneficial components.

Ratings are based on:

Total energy (kilojoules)

- Saturated fat, sodium (salt) and sugar content. Consuming too much of these risk nutrients is linked to being overweight and obese, some cancers, heart disease and type 2 diabetes.
- Fibre, protein, fruit, vegetable, nut and legume content. Increasing consumption of these nutrients and ingredients is good for your health.

The Health Star Rating system does not take into consideration other real, claimed, or potential health effects of particular ingredients, products or processing methods. Neither does it consider other important nutrients. If a consumer aims to, for example, avoid certain preservatives, consume organically grown products, or increase the intake of certain nutrients, he should review the ingredients list, Nutrition Information Panel and/or seek guidance from other reliable sources.

The star ratings for all products are calculated based on a consistent measure of either 100g or 100mL of a product. This means that the star ratings of similar products can be compared at a glance.

On 19 December 2019, the Australia and New Zealand Ministerial Forum on Food Regulation (Forum) published a response to the recommendations of the HSR Five-Year Review.

Monitoring data strongly indicates that, despite some challenges in measuring the public health

impact of the HSR System, the System is generally well used, recognised, reliable and has the potential to be a successful public health intervention by assisting consumers to make healthier choices when purchasing packaged foods and beverages.

Of particular note is recent Australian survey data which shows that of consumers purchasing a product displaying the HSR in the last three months, almost two thirds stated that the HSR influenced their decision and one third were influenced to purchase a product with more stars. This equates to 23% of all consumers being influenced by the HSR to change their purchasing behaviour to purchase a product with more stars. This is in line with other research that found more than three in five Australians who bought a product displaying the HSR reported that it had influenced their product choice. One in two purchased a product they would not normally buy due to the presence of the HSR and close to nine in 10 of these shoppers said they would continue to buy the product.

[3] Healthy Eating System

Taking into account the results of a market research for understanding the consumer perception on the effectiveness of FOP systems in orientation for the healthiest food options, in 2011, the Sanitarium Health & Wellbeing company (which is one of the most trusted food brands in Australia and New Zealand), in cooperation with Public Health Association from Australia had proposed a new FOP labelling system, based on the strengths of the existing systems (Traffic Light and GDA), which was intended to overcome their limitations.

The new model, called "Healthy Eating System" (HES), uses colour coding and incorporates the following elements:

- undesirable nutrients, associated with the increased risk of chronic diseases (saturated fat, sugars, salt)
- parameters with positive impact on health: whole ingredients and dietary fibre, in grams per 100 g of foodstuff (marked by two distinct icons, such as "fruit, vegetables, whole grains, nuts and legumes" and "fibre").

Thus, by including parameters related to whole ingredients and fibre, the system highlights also the items with positive impact on health, answering at the most important criticisms of the "Traffic Light" system. In addition, the "Healthy Eating System" includes a general dietary advice for consumers, which may result in: "eat often", "eat occasionally" or "eat sparingly". These recommendations are based on nutritional criteria imposed by FSANZ (Food Standards Australia & New Zealand).

The Health Eating System is a coloured, almost complex, Evaluative (interpretative) FOP label scheme.

2.2.4. Endorsement schemes ('positive logos')

[1] Keyhole

The Keyhole label was launched in Sweden in 1989. The label began as a local initiative in Swedish Västerbotten in the mid-1980s to improve public health. Today, the Keyhole serves as a Nordic health label for food, and for which Sweden, Denmark, Norway and Iceland jointly establish the criteria. The Keyhole has been used in Denmark and Norway since 2009, and in Iceland since 2013. It is an independent label for healthier food choices.

The Keyhole is the label of the Swedish Food Agency. It is based on the latest research on Nordic dietary habits and on what the consumer needs to eat in order improve his health (the Nordic Nutrition Recommendations, NNR). The Keyhole focuses on five areas:

- less salt
- less sugar
- more fibre
- more wholegrains
- healthier fat

The above areas have proven to be of great importance for health.

Established and recognised. Almost everyone in Sweden (96%) recognises the Keyhole. Studies show that consumers consider it positive that the Keyhole is administered by an independent authority, the Swedish Food Agency. It makes the label more trustworthy.

The Keyhole has been used in Sweden for over 30 years and is established on the Swedish market. Nordic consumers see it as a strength that the label is used on products in several markets in the Nordic region.

Products that can be labelled with the Keyhole. The Keyhole is divided into different product groups. Different foods contain different types of nutrition and in different quantities. Therefore, the requirements differ in terms of what the foods in different groups must contain in order to be labelled with the Keyhole. Cereals are compared to cereals, sausages to sausages, etc. Keyhole-labelled products in a particular product group are thus a healthier option within that particular group.

The keyhole comprises 11 main groups and 32 subgroups:

- Vegetables, fruits, berries and nuts
- Flour, grains and rice
- Porridge, bread and pasta
- Milk, fermented products, plant-based alternatives, etc.
- Cheese and corresponding plant-based products
- Fats and oils

- Fish and shellfish and derivative products
- Meat, cold cuts, sausages, etc.
- Plant-based products
- Ready meals
- Dressings and sauces

It is an endorsement scheme ('positive logo') based on threshold levels for energy and various nutrients depending on product category. Foods labelled with the Keyhole contain less sugars and salt, more fibre and wholegrain and healthier or less fat than food products of the same type not carrying the symbol. Some food categories are not permitted to carry the logo (e.g. sweet and savoury snacks). The reference base typically is 100 g or 100 ml.

Awareness of nutrition and health claims during shopping appeared limited, and claims were generally not used on a conscious level in the purchasing decision process. The Keyhole label was among the best-understood claims; it was most often used in the evaluation of products, but only by few consumers. No substantial connection was found between consumers' understanding of the Keyhole scheme and choice of Keyhole-labelled products (Aachmann et al., 2013). According to the study of Wang et all, 2016, Keyhole labelling increased perceived healthfulness of snacks relative to % Daily Value and plain labels.

[2] Choices Logo

In 2014 the National Food Committee of Thailand therefore advised to implement a front-of-pack label. The Healthier Choices Nutrition Symbol was then developed with the agreement of food companies, academics, and the public sector.

On 31 August 2016, the Minister of Public Health of Thailand has given the official kick-off for a front-of-pack logo programme that identified healthier food options per food category. Choices International welcomed this initiative as an important element of the public health actions in Thailand.

Overweight, obesity and related non-communicable diseases as well as a rapid increase in the number of elderly causes growing nutritional problems in Thailand that have affected at least 30% of the population. This can cause huge financial burdens for the country's health care system.

The logo is owned by the Thai Food & Drug Administration, which assigned the Nutrition Promotion Foundation of the Mahidol University, to manage the symbol's use.

Intended to cover all food products, the logo now can be applied to ready-to-eat meals, beverages, dairy products, sauces, instant noodles, and snacks. Other food groups will follow. From the start already 56 products are certified, and 15 food companies have joined the programme.

The "Healthier Choice" nutritional logo has been developed in Thailand, based on nutritional facts of each food product, however indicating as a simple and easily noticeable logo of a particular food product. It is aimed to create consumers' awareness on nutrition qualities of their foods products.

Consumers understanding. Multiple Traffic Lights allowed consumers to better identify more healthful foods, compared to a simple 'healthy choice' tick, a monochrome and a traffic lights coded

GDA label, and a 'no label' condition (Gorton et all, 2009).

According to the study of Borgmeier & Westenhoefer (2009), no significant differences emerged between the various FOP labels in terms of energy and nutrient intakes. The authors note, though, that different labelling schemes may work differently depending on the food group. They observed that Traffic Lights yielded the most correct choices for dairy products, whereas the healthy choice tick performed best on breakfast cereals. This latter result may have been influenced by the different number of products available for comparison: 8 dairy products vs. 3 breakfast cereals.

[3] Finnish Heart Symbol

The Heart Symbol tells the consumer at a glance that the product marked with this symbol is a better choice in its product group regarding fat (quantity and quality) and salt. In some product groups, also sugar and fibre contents are considered. This FOP label is based upon the grouping of all foods in main groups that include:

- milk and dairy products
- oils and fats
- fish
- meat
- meat products
- bread and cereal products
- convenience foods
- spices and seasoning sauces
- vegetables, fruits and berries

Criteria. The criteria for the symbol are based on the Finnish nutrition recommendations. In all, the criteria are defined for nine main food groups that may further be divided into subgroups. The Heart Symbol tells the consumer at a glance that the product marked with this symbol is a better choice in its product group regarding fat (quantity and quality) and sodium. In some product groups, also sugar and fibre contents are considered.

Example of criteria for oils and fats

Fat spreads	Hard fat max 30 % of the total fat Salt max 1,00 g/100 g
Vegetable oils	Hard fat max 20 % of the total fat
Liquid oils	Hard fat max 20 % of the total fat Salt max 1,00 g/100 g
Salad dressings, mayonnaise, hamburger- and sandwich dressings	Fat min 25 g/100g Hard fat max 20 % of the total fat Salt max 1,00 g/100 g Sugars max 15 g/100g
Vegetable pastes and spreads (e.g. hummus)	Hard fat max 20 % of the total fat Salt max 0,38 g /100 g Sugars max 10 g/ 100 g

The criteria are regularly updated, if needed, by the Heart Symbol expert group, which includes seven professionals in nutrition and medicine appointed by the organisations in charge, i.e. FHA and FDA.

In Finland, the need for nutrition labels to help consumers to make healthier food choices was emphasized in the consensus statement for promoting Finnish heart health in the late 1990's. The system including a front-of-pack logo, a Heart Symbol, for Finnish consumers was developed and launched jointly by Finnish Heart Association (FHA) and Finnish Diabetes Association (FDA) in the year 2000.

The development was heavily based on work of Finnish experts in the field of nutrition and medicine,

and there was active collaboration with Finnish Food Safety Authority and other relevant authorities that has continued since then. As the Finnish nutrition recommendations were updated in 2005, the Heart Symbol was included in guidelines for consumers. Based on European regulations (EC N° 1924/2006) on nutrition and health claims made on foods, the Heart Symbol has been notified to be the only symbol on the Finnish market to be regarded as a nutritional claim.

It is an endorsement scheme ('positive logo') based on threshold levels for energy and various nutrients depending on product category. The logo identifies options with a better nutrient profile in a given category regarding fat (quantity and quality) and salt; in some product groups, also sugar and fibre contents are considered. Reference base is 100 g.

Consumers understanding. Men and women with the highest education were best aware of the Heart symbol and more likely to use products bearing such a label in the early 2000s. The educational differences diminished or disappeared during the study period. (Lahti-Koski et all, 2012).

Whilst nutrient intakes could be improved by using products labelled with the Keyhole or the Finnish Heart Symbol, the most substantial effect was achieved with 'best-in-class' products. However, it is unlikely that such products would achieve 'best-in-class' status for all nutrients concomitantly. As an example, choosing the lowest salt level product in a given category might not result in an optimal saturated fat reduction for that category (Ahlin, 2015).

[4] Pick the Tick

The "Pick of the Tick" FOP label was developed and implemented in Australia and New Zealand, but it is no longer in use. The Pick the Tick programme of the National Heart Foundation of New Zealand aimed to provide a framework for cooperation with the food industry to improve nutrition labelling and to develop a healthy food supply. Food manufacturers, whose products meet defined nutritional criteria, were able to display the "Pick the Tick" logo on food labels. The logo was used by 59% of shoppers in assisting them make healthy food choices. Food companies were encouraged to reformulate product composition if they fail to meet criteria and develop new products to specifically meet the Pick the Tick criteria.

It was an endorsement scheme ('positive logo') based on threshold levels for energy and various

nutrients depending on product category. The Reference base was100 g. Manufacturers who wanted to feature the Tick on their products had to formulate (or re-formulate) products to fit the Tick criteria for that particular product category.

Tick products often had lower levels of energy, fat, saturated and trans fat, sugar and salt than other products in the same category, and they're often higher in calcium and fibre. So it could be a useful way to quickly choose products with a healthier profile. There was a cost to manufacturers to be in the Tick programme, and not all took part. It is also important to notice that a Tick did't mean that a consumer could eat unlimited amounts. Some food categories that included Tick products will always be 'sometimes' foods.

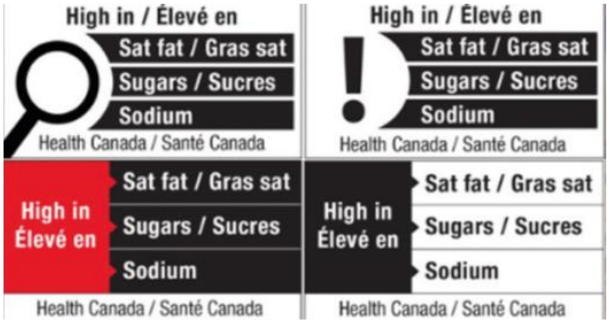
[5] Smart Choices

The "Smart Choices" FOP label was developed in USA and it is no longer in use. It was an endorsement scheme ('positive logo') based on threshold levels for energy and various nutrients and food components depending on product category. Adaptations concerning one or more of several nutrients (total fat, saturated fat, trans fat, cholesterol, sodium, added sugars) for various categories. Portion was the reference base.

Consumers understanding. The Multiple Traffic Lights helped respondents better distinguish more and less healthful products. Labelling schemes without reference point information (e.g. nutrition table) were found less easy to interpret when no comparison product was available, and the Smart Choices logo could reduce consumers' ability to compare categories, leading to a potential misinterpretation of product healthfulness. None of the labels affected food preferences. (van Herpen, Hieke & van Trijp, 2014). According to Roberto et all, 2012, 64% of the products carrying the industry-developed 'Smart Choices' label did not meet the nutrient profile model criteria for a healthful product.

2.2.5. Warning signs

In some countries, such as Chile or Peru, warning signs FOP labels have been developed, in order to discourage consumers from consuming certain foods frequently because they contain large amounts of specific ingredients. Since June 2017, all processed food and drink products sold in Peru must carry a "stop sign" warning label if they are high in salt, sugar and saturated fat, or if they contain trans-fat.



2.3. Comparative Analysis and Synthesis

The following tables illustrate the comparative analysis of the FOP labels systems. The comparison has been made per group of FOP labels, in order for the results to be really comparable, since the groups of labels have significant differences between them, both in terms of the amount of information they contain and the evaluation or not of the food.

The comparison was made based on the following criteria:

1. Whether the system was implemented inside or outside the EU.
2. The number of countries that have implemented the system.
3. If it is still active or its application has stopped.
4. If it is an evaluative or reductive system.
5. If it is based on the application of a Directive.
6. The complexity of the information it provides to the consumer.
7. And finally a very basic factor that is the reference base of the label: 100g food, or portion.

2.3.1. Reference Intakes and similar schemes

FOP label name	Inside / Outside EU	Number of Countries	Active	Evaluative or Reductive	Directive oriented	Simple - Complex	Reference
Facts-Up-Front	Outside	1	YES	Reductive	Semi-Directive	Complex	Portion
Guideline Daily Amounts (GDA), Reference Intakes	Outside	1	YES	Reductive	Semi-Directive	Complex	100g
Daily Intake Guide	Outside	2	YES	Reductive	Directive	Complex	Portion
Nutriform Battery	Inside	1	YES	Reductive	Directive	Complex	Portion

2.3.2. Colour-coded nutrient-based schemes

FOP label name	Inside / Outside EU	Number of Countries	Active	Evaluative or Reductive	Directive oriented	Simple - Complex	Reference
UK Multiple Traffic Lights (MTL)	Outside	1	YES	Evaluative	Semi Directive	complex	100g
Evolved Nutrition Label (ENL)	Inside	Some EU countries	No	Evaluative	Semi Directive	complex	Portion

2.3.3. Overall rating schemes

FOP label name	Inside EU Outside EU	Number of Countries	Active	Evaluative or Reductive	Directive oriented	Simple - Complex	Reference
Nutri-Score	Inside	5	YES	Evaluative	Semi Directive	Simple	100g
Health Star Rating	Outside	2	YES	Evaluative	Semi Directive	Complex	100g
Healthy Eating System	Outside	2	YES	Evaluative	Semi Directive	Complex	100g

2.3.4. Endorsement schemes ('positive logos')

FOP label name	Inside / Outside EU	Number of Countries	Active	Evaluative or Reductive	Directive oriented	Simple - Complex
Keyhole	Inside	3	YES	Evaluative	Semi Directive	Simple
Choices Logo	Outside	1	YES	Evaluative	Semi Directive	Simple
Finnish Heart	Inside	1	YES	Evaluative	Semi Directive	Simple
Pick the Tick	Outside	2	YES	Evaluative	Semi Directive	Simple
Smart Choices	Outside	1	NO	Evaluative	Semi Directive	Simple

2.4. Discussion

Grunert and Wills (2007) note that consumers are more interested in some nutrition information like calories and fat, followed by salt and sugar; this means that information on nutrition principles do not all have the same weight in the choice of foodstuff.

Balcome et al. (2010) and Hieke and Wilczynski (2012) note that consumers have a hierarchy of importance in terms of perception of the various colours: the consumers placed greater emphasis on a change in a product's nutrient characteristic from red to amber compared with a change from amber to green.

The results from an UK market study on food labels, realized by European Food Information Council (EUFIC) in 2008, shows that only one in four consumers (25%) looks for nutrition information on foodstuffs' package in supermarkets and UK consumers spend around 25 seconds in adopting a purchase decision of foodstuff. This study revealed a high awareness of consumers for both GDA and TL labelling systems. In terms of understanding, the GDA concept is good (with 89% of consumers correctly defining a guideline daily amount as a maximum rather than a target to reach). The understanding of the "Traffic Light" system seems to be characterized by some exaggeration of the meaning of the colours (73% of consumers thought that red colour indicated avoidance rather than the occasionally consumption) and a lack of understanding that the system is applied per 100g (EUFIC, 2008). Moreover, participants liked and appreciated the use of traffic-light colours (yet showed some confusion, especially in relation to the colour amber and when faced with healthfulness assessments based on mixed traffic lights). (De la Cruz Gongora et all, 2017).

In this context, it is essential for any front-of-pack label to be widely adopted (and preferably widely supported), as well as certified and controlled by reputable third parties. Ideally, the algorithms behind the calculation of the label should be scientifically sound and adopted internationally without any differences between countries, retailers, or brands. Acceptance of such a label by consumers requires trust – only then it will nudge consumers towards different products.

All in all, reaching a wide audience is important in order to encourage healthier food choices. Rabobank believes the Nutri-Score is potentially best suited for this purpose. To interpret the Traffic Light label, a consumer still needs a considerable amount of knowledge, whereas Nutri-Score is more intuitive. And those who want to know more can always turn to the back of the pack.

Despite this, the diet is actually at risk from a system purportedly designed to help boost public health by allowing consumers to make more informed choices. With numerous studies indicating that consumers find current nutritional labelling confusing and a barrier to eating a healthy diet, it is not surprising that the EU wants to implement a harmonized front-of-pack (FOP) labelling system by 2022.

There is no such thing as a bad food or a good food, there is no such thing as a food that can be eaten and one that cannot. We need to focus on the quantities as related to the overall food and nutritional daily intake. It is misleading to believe that to combat the obesity epidemic, for instance, we should remove fat or sugars. We need to focus on awareness and education and not on simplification. We cannot hope to win over obesity by misleading people, we need the exact opposite.

It is also suggested that it would be good for the FOP labels to be colourful, because this attracts the attention of the consumers. The very simplified FOPs are easy to read by consumers, but because they are "black box" it is not clear to the consumer how the ratings of each food are obtained.

The best practice is to "train" consumers to read and understand food ingredients labels, that will lead to more informed choices on which food to consume and in what quantity. By understanding the ingredients' data and especially the %RI, consumers will be able to evaluate not only whether they will buy a food or not but also in what quantity they will include it in their daily diet.

Traffic light FOP label is very easy for the consumers, but red (much more) and orange (less) "prevent" them more from consuming a food that falls into this category, than "encourage" them with green

colours to consume foods belonging to the relevant category. There are no "bad foods", there are "bad amounts" of eating a specific food product.

Maybe in correspondence with the "enjoy responsibly" in the drinks, there should be an indication "consume the right amount" on the food packages. By this logic, a FOP could be adopted that would be "traffic light" regarding the daily consuming amount. This can be calculated depending on the level the specific food product "charges" the "battery" regarding the "ideal", or "maximum" or "minimum" amount that a person should consume of a component during the day.

For example, the FOP label should prevent the consumer from eating the largest percentage of his daily fat from a snack and to "determine" the maximum amount to be consumed from that particular snack per day, without generally making it a "forbidden" food.

The average daily intake needs of ingredients are for the "average adult". But in reality, there is no "average consumer". Children have other needs; the elderly have other needs. Maybe it would be good to also refer on the FOP label the up or down %RI deviations for large groups of consumers (children, the elderly, etc).

It is, therefore, advisable to separate the ingredients that are harmful to health when a daily intake is exceeded (such as fat, or sugar) from the ingredients that have beneficial effects to the body and which, in a balanced diet, should be consumed at least a minimum amount.

That is, if the FOP will be in a form of "battery" to warn when the battery is "overcharged" for "bad" components and "urge" to "charge" the battery more when it refers to "useful" components.

3. Ecolabel Schemes for Food Products

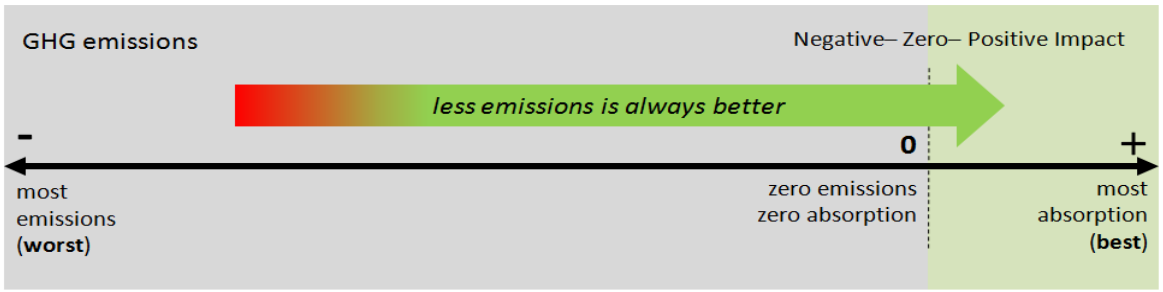
3.1. Introduction

3.1.1. Distinction of nutritional to environmental and other ethical labels

Environmental concerns, and in particular those that concern the emission of greenhouse gases (GHG) that cause human-made climate change, differ from concerns related to the nutritional content of foods in normative and practical ways. In the literature, the environmental aspects of food production fall under the broad category of "ethical" attributes. This category includes several end-product and production process-related aspects that are highly important both in ethical terms and in public perception. Notable examples are altruistic concerns about labour conditions and animal welfare, as represented by well-established labelling initiatives in the market. In our times, **environmental issues,** and particularly the emission of GHG, stand out among other ethical attributes because of the overarching catastrophic potential of manmade climate change and the urgency to combat it. Although all ethical concerns and all environmental impact categories are undoubtedly highly important, the **severe and overarching** catastrophic effects of manmade climate change foreseen at all levels (social, economic and also on the biodiversity and on broader ecosystems), as well as the **urgency** stemming from the limited time available to be effective in mitigating it, **dictate** that the maximal functionality for combating climate change should be the **top priority** of any EU-issued environmental label. This is particularly important in the case of food production, which amounts to about a quarter of total manmade GHG emissions globally.

A significant practical difference that distinguishes the labelling approach recommended for GHG emissions from that recommended for nutritional content *is inherent in the nature of these attributes: the ideal content in most food ingredients is often a matter of optimal dose and also varies with the physiology of the consumer. In contrast,* ***the GHG emissions attribute exhibits a clear normative direction of the "less is better" type:*** *it is always good to emit less GHG, It is better to emit none, and it is even best to be most negative in emissions, i.e. to absorb as much GHC as possible (e.g. carbon offsetting).*

Therefore**,** in other words, other things being equal, i.e. analytically speaking, it is always good to emit less GHGs. It is better to emit none, and it is even best to be most negative in emissions, which means to absorb as much GHCs as possible (e.g. carbon offsetting). Besides its effect on the recommended labelling approach, this clear normative direction makes it also meaningful to describe products using normatively charged terms, such as "better" and "worse" products, with regards to their GHG emissions.



x Axis: *normative direction of GHG emissions from worst (-) to best (+) as assessed by the scientific consensus.*

Considering the above, the focus of the second part of this report is on environmental labelling. Its special focus is the functionality of any considered labelling approaches for mitigating climate change.

3.1.2. Environmental labels in the broader ethical labelling and policy context

The fundamental functionality of environmental labels is to enable concerned consumers to identify the environmental impact of products. Then, concerned consumers can inform their subjective purchase choices also on the environmental performance of products, alongside their preferences for other attributes. The **value added** for suppliers of environmentally superior products includes increased product demand, and also improved reputation, higher workforce morale and ability to attract high quality employees (Carroll and Shabana 2010).[3] The opposite effects can be expected by poor performers (worst polluters). These effects incentivize environmental improvements among supply actors when the environmental performance of products is publicly communicated. Accordingly, environmental labelling is a tool that incentivizes the evolution of production towards improved environmental impact. However, **these effects also explain the resistance** of a considerable part of the supply sector to compulsory labelling initiatives that inform consumers and society on the performance of worst performers.

Environmental labelling has a unique place within the broader public policy toolbox used to direct the real economy towards more socially desirable outcomes. Its value rests on that it effectuates positive change driven by "ethical consumerism", which is a market force stemming from the revealed ethical preferences of consumers. Namely, the driving force for improvement is not the coercive power of the state, as is the general case with state regulations, but the informed choices made freely by consumers in the market. Seen in this light, environmental labelling is a tool that empowers the public and improves market democratization. Importantly, it also bears a second characteristic that makes it **invaluable as a complementary policy instrument**: It has the potential to incentivize market improvement above and beyond the limits accessible to state regulation, when the latter is constrained from being sufficiently effective.[4]

The **importance of environmental labelling for enabling systemic environmental consumerism** cannot be overstated. Ethical consumerism invariably involves unobservable "*credence*" attributes. Because of their nature, credence attributes are difficult or impossible to verify by consumers, even after using a product.[5] Therefore, it is ***not possible*** for consumers to make market choices that reveal their ethical preferences ***unless*** appropriately informed on relevant product performance. This includes environmental and climate change-related preferences. The most readily available information is communicated to consumers through Front-of-Pack (FoP) product labels. In the last decades, a large variety of FoP labelling schemes developed to inform consumers on the ethical aspects of their consumption. These include credible and institutionalized initiatives, alongside a considerable variety of less credible ones that are often perceived as tools for "greenwashing" suppliers' environmental performance.[6]

Nevertheless, contrary to initial optimism that ethical consumerism could provide the motivation for socially responsible production, and contrary to consistent survey results showing increased ethical concern among the international public, ethical consumption remained a market niche. So far, it failed to develop into a market force of a magnitude sufficient to incentivize the much needed substantial and widespread improvements in the environmental, societal, and animal welfare performance of the

[3] Carroll, Archie B., and Kareem M. Shabana. 2010. "The Business Case for Corporate Social Responsibility: A Review of Concepts, Research and Practice." International Journal of Management Reviews, 85–105. doi:10.1111/j.1468-2370.2009.00275.x.

[4] The regulatory power of state-like entities is restricted by public policy constraints. These constraints may refer to international trade law and also to all kinds of political, practical and normative factors that may prevent aggressive regulatory measures such as, for instance, Qualified Market Access (QMA) measures against imports to prevent environmental or social dumping. (Holmes et al. 2008). Holmes, Peter, Jim Rollo, Alan Winters, Kamala Dawar, and James Mathis. 2008. "QUALIFIED MARKET ACCESS". European Commission (DG Trade). http://dare.uva.nl/document/137295.

[5] Darby, Michael R., and Edi Karni. 1973. "Free Competition and the Optimal Amount of Fraud." Journal of Law and Economics 16 (1): 67–88.

[6] Gruère, Guillaume. 2013. "A Characterisation of Environmental Labelling and Information Schemes." *OECD Environment Working Papers*, no. No. 62 (October). https://doi.org/10.1787/5k3z11hpdgq2-en.

supply sector, and more generally in the socially responsible conduct of businesses (Vogel 2006; Smith 2007; Higgins 2010; Carroll and Shabana 2010).[7]

3.1.3. Disrupting the pattern: A more ambitious approach to label functionality

Restating the above in economic terms, environmental attributes suffer from information asymmetry, which excludes them **by default** from the process of consumer-driven market optimization. Information asymmetry is thoroughly studied by prominent economists, including Nobel laureates George Akerlof, Michael Spence, and Joseph E. Stiglitz, and further contributing authors. Its effects include moral hazard, adverse selection, rent seeking, and the free-rider problem. These lead to the proliferation of negative environmental externalities and cause the market for environmentally friendly production to fail. Valuable are also the contributions of Nobel laureate Herbert Simon and further contributing authors on "bounded rationality". These explain that cognitive constraints in gathering, and processing information are among the factors that prevent individuals from making rational choices. Such theoretical grounds underline more recent insights favouring information that is easily accessible and easy to understand and compare, as provided by intuitive front-of-pack (FoP) color-coded labelling schemes. From this perspective, environmental labelling generally aims to cure information asymmetry on the environmental impact of products, so as to arrive at a distribution of environmental impacts that matches the one that would be achieved should the environmental impact of production was observable.

The ambitiousness of different labelling approaches in achieving this goal differs. A large majority of available labels provides information that is functional for the goal of raising a part of production to a higher environmental standard. In the best cases, the said environmental standard is sustainable. Such labelling initiatives succeeded to bring some improvement through product certification and also through positive spill over effects to the broader production sector. However, decades of experience shows that **the progress achieved is insufficient** as compared (a) to the magnitude of the environmental challenges and of the scale of changes needed to address them and (b) to the potential available as indicated by consistent experimental and survey results showing increased ethical concern among the international public.[8]

A more ambitious labelling goal is to establish for environmental attributes the full effect that normal market forces have on the evolution of observable (non-credence) attributes that belong to the more-is-better type (similarly as the less-is-better type, the more-is-better attribute type has a clear normative direction). Namely, to engage the market in a consumer-driven and open-ended race for continuous improvement. Consider for instance the evolution of the processor capacity attribute in electronic devices. No standard ever rules a computer processor as fast enough. Contrarily and similarly to other such attributes, there is an implicit consumer acceptability threshold that is dynamic and market defined. This consumer acceptability threshold is set by the evolution of processor speed among marketed products, and it is continuously raised by suppliers that compete in investing, inventing and incorporating innovative technology to improve the speed of their products. In this continuous race for improvement, **novel improvements by competitors effectively downgrade in the eyes of consumers the performance of products that until that moment were regarded as**

[7] Vogel, David. 2006. *The Market for Virtue: The Potential And Limits of Corporate Social Responsibility*. Washington D.C.: Brookings Institution Press.
Smith, N. C. 2007. "Consumers as Drivers of Corporate Responsibility." In *The Oxford Handbook of Corporate Social Responsibility*, edited by Andrew Crane, Abagail McWilliams, Dirk Matten, Jeremy Moon, and Donald Siegel. Oxford: Oxford University Press.
Higgins, Colin. 2010. "Is a Responsive Business Also a Responsible Business?" *Journal of Business Systems, Governance and Ethics*, Special Issue, 5 (3): 23–32.
Carroll and Shabana 2010, as cited previously.

[8] This discrepancy between the statements and the market choices of consumers is termed the "words-deeds gap". Although consumer research accepts that social factors such as the "social desirability bias" in consumer statements explain a part of it, the size of the gap remains largely unexplained.

satisfactory. This inherently 'free-market' dynamics is spontaneously at work in markets for observable attributes: Products and firms must continuously innovate and improve, or else they will become obsolete and will exit the market. **Although this dynamic is business as usual in markets for observables, it sounds alien in markets for ethical attributes that are ruled by the dominant labelling paradigm.**

On this basis, and while acknowledging that any environmental labelling approach ought to also respect other important requirements that fall outside the focus of this report (referring to its validity and credibility), the present report focuses especially on the functionality of available labelling approaches for incentivizing systemic, continuous, and open-ended environmental improvement of the food supply sector.

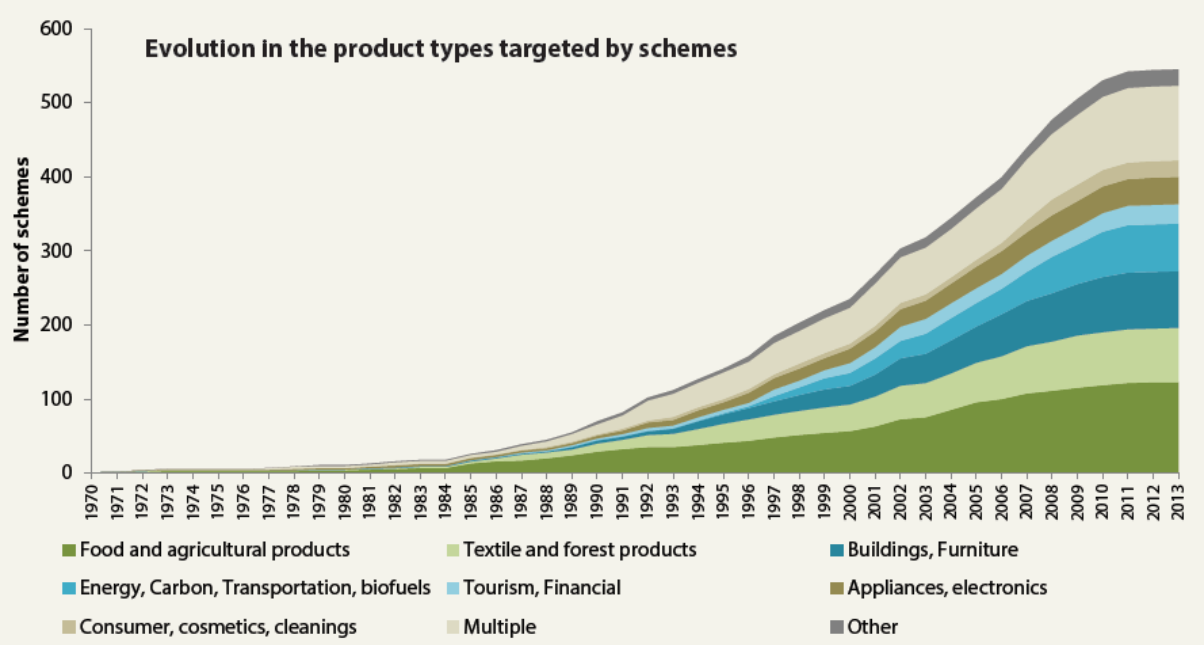
Its goal is to contribute to the success of a potential environmental labelling initiative by the European Union **to disrupt the established paradigm of negative environmental effects** in production, in general and also particularly in relation to climate change.

3.2. Landscape of eco-labelling schemes for food products in Europe and beyond

3.2.1. Background

Most broadly perceived, the plurality of approaches used to provide consumers with information on the environmental performance of products includes press releases and publicized commitments to codes of ethical conduct and social responsibility reports. Among these, FoP product labelling provides such information in its most readily accessible form. Its credibility depends on the type of the issuing authority and of supporting partners. Governing authorities and major not-for-profit organizations are regarded as the most credible among them.

Globally, the number of environmental labels increased rapidly between 1990 and 2010, at which point the market started to stabilize (Fig. 2). The EU experienced a similar increase in the number of labels that refer to at least one environmental aspect. Although the present number of environmental labels in the EU is not officially tracked,[9] indicatively, in 2015 the Ecolabel Index[10] listed 455 eco-labels globally, 231 of which had some presence in Europe. At least 85 of these eco-labels apply partially or exclusively to food products.



(source: OECD[11])

[9] https://ec.europa.eu/environment/gpp/eco_labels.htm
[10] http://www.ecolabelindex.com
[11] OECD. 2016. "Environmental Labelling and Information Schemes - Policy Perspectives -Synthesis Report." https://www.oecd.org/env/policy-persectives-environmental-labelling-and-information-schemes.pdf.

3.2.2. The ISO Classification of Environmental claim Types

To structure the environmental claims landscape, the International Standards Organisation (ISO) developed standards for three types of environmental claims. These also cover claims communicated to consumers through FoP product labels.[12] According to ISO's prominent classification, environmental product labels are classified into three categories, termed as *ISO Type I, II* and *III*. These cover a variety of label types and are described as follows.[13]

*"**Type I** (ISO 14024) claims are based on criteria set by a third party and are multi-issue, being based on the product's life cycle impacts. The awarding body may be either a governmental organisation or a private non-commercial entity. Examples include the EC Eco-label, Nordic Swan and German Blue Angel*

***Type II** (ISO 14021) claims are based on self-declarations by manufacturers or retailers. There are numerous examples of such claims eg 'made from x% recycled material'*

***Type III** (ISO/TR 14025) claims consist of quantified product information based on life cycle impacts. These impacts are presented in a form that facilitates comparison between products e.g. a set of parameters. However, there is no comparing or weighting against other products inherent within the claim. An example which has similarities with Type III claims is Volvo's product profile for its S80 passenger vehicle."*

However, the classification put forward by ISO is not exhaustive:

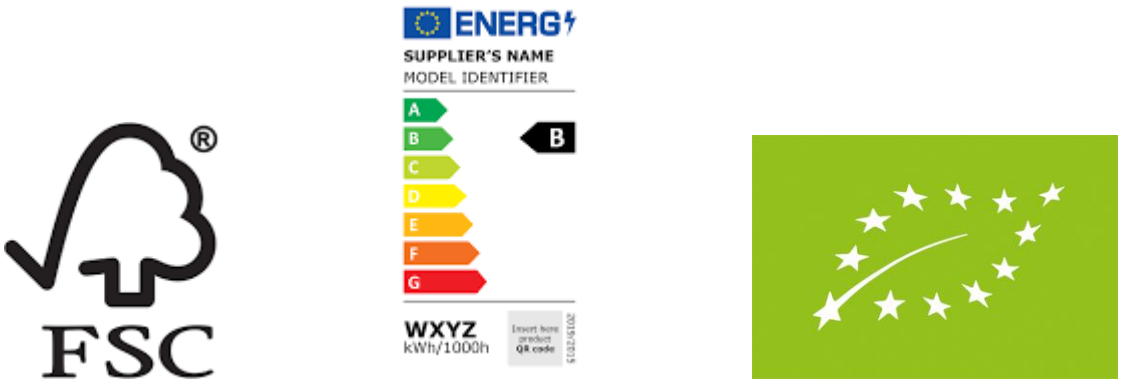
"The ISO claim types form a method of standardising 3 forms of environmental product claims out of the multiplicity which currently exist. Other notable and established forms of product environmental information not covered by the ISO standards include:

- *Single issue labelling schemes with third party certification such as the private Forest Stewardship Council (FSC) and organic food labels do not fall within any of the three ISO claim types but are partially covered by ISO 14020 - General Guidelines for Environmental Claims and Declarations.*
- *Mandatory labels can take any form and may be multi or single issue. Current examples include the EC Energy Label and EC CO2 emissions label for passenger cars.*
- *User information which provides instructions on how to use or dispose of the product in a manner which reduces its environmental impact comes in a variety of forms. This form of information is extremely important for those products whose main impact is concerned with its method of operation and/or disposal. An example is the 'Wash Right' code for detergents, developed by the European Detergent Manufacturers Association (AISE) (1).*
- *Other forms of product environmental information currently used in business-to-business communications such as quantified Environmental Product Declarations not based on LCA and supplier questionnaires also fall outside the ISO claim types. "*

[12] "Environmental claims made in regard to products may take the form of statements, symbols or graphics on product or package labels, or in product literature, technical bulletins, advertising, publicity, telemarketing, as well as digital or electronic media, such as the Internet."

[13] Environmental Resources Management DD Environment. 2000. "Study on Different Types of Environmental Labelling (ISO Type II and II\ I Labels): Proposal for an Environmental Labelling Strategy." European Commission. https://ec.europa.eu/environment/ecolabel/about_ecolabel/reports/erm.pdf.

Notably, among the available labels not covered by the above types are widely established labels such as the private **Forest Stewardship Council (FSC)** and the **organic food labels**, as well as all **mandatory labels,** either multi or single issue. The latter include the EC Energy Label and the EC CO2 emissions label for passenger cars.



Label examples not covered by the ISO Types (I, II, III) classification

Labels not covered by the Type I, II, III classification are partially covered by *ISO 14020 - General Guidelines for Environmental Claims and Declarations:*

1. *Environmental labels and declarations shall be accurate, verifiable, relevant, and not misleading.*
2. *Procedures and requirements for environmental labels and declarations shall not be prepared, adopted, or applied with a view to, or with the effect of, creating unnecessary obstacles to international trade.*
3. *Environmental labels and declarations shall be based on scientific methodology that is sufficiently thorough and comprehensive to support the claim and that produces results that are accurate and reproducible.*
4. *Information concerning the procedure, methodology, and any criteria used to support environmental labels and declarations shall be available and provided upon request to all interested parties.*
5. *The development of environmental labels and declarations shall take into consideration all relevant aspects of the life cycle of the product.*
6. *Environmental labels and declarations shall not inhibit innovation that maintains, or has the potential to improve, environmental performance.*
7. *Any administrative requirements of information demands related to environmental labels and declarations shall be limited to those necessary to establish conformance with applicable criteria and standards of the labels and declarations.*
8. *The process of developing environmental labels and declarations should include an open, participatory consultation with interested parties. Reasonable efforts should be made to achieve a consensus throughout the process.*
9. *Information on the environmental aspects of products and services relevant to an environmental label or declaration shall be available to purchasers and potential purchasers from the party making the environmental label or declaration.*

3.2.3. A critical typology of eco-label formats according to their functionality

A more relevant categorization of available eco-labels is according to technical traits that considerably affect their functionality for consumer-based market optimization. In the non-exhaustive account of available types that follows, two basic label types are identified in the form of Endorsement and the Quantitative labels. Comparative, dynamic, and color-coded variations of these basic types are also noted as functionally important.

Endorsement labels

The endorsement label type refers to single grade "seals of approval" given by certification bodies to products that comply with specified (sets of) criteria. These criteria constitute the Standard of the particular certification scheme denoted by each label. Endorsement labels are generally voluntary.[14] They can be multi or single-issue. Depending on additional requirements they can fall within the ISO Types I or II, or else they fall outside the ISO classification.

Essentially, the only information conveyed to consumers by this label format is that product performance meets the threshold levels set by the certification standard for the environmental aspects concerned. **The credible** third party-issued and monitored certifications generally signify products with **superior** environmental performance. Such products have lower negative environmental impacts (or higher positive ones) than the average of their product substitutes- namely, than functionally or competitively similar products.

Accordingly, endorsement labelling has been used since long to inform consumers, so as to reward the most ethical products in the market. In this function it has served a useful purpose. It increased consumer awareness about the environmental performance of products and motivated products near certification levels to make further improvements to achieve certification, which create positive secondary spill over effects within the production sector.

However, the **limitations** of endorsement labelling with regards to its functionality for consumer-driven environmental optimization of the market, include:[15]

1. **Endorsement labels do not make possible comparisons** of the environmental performance of specific products, because voluntary endorsement labelling allows the actual performance of specific products to remain unknown. Any certified product may actually perform well above or barely at the level required by the certification threshold. Consequently, two labelled products cannot be readily compared. This is true also for products having different labels because comparison of different labels requires consumers to research and analyse the precise requirements of each certification.

   In addition, although credible endorsement labels generally indicate products that perform better than the market average, in principle any particular unlabelled product can actually be worse, equal or better than a labelled one.

   Consequently, although certification by credible authorities is an indication of improved-than-average environmental performance, ample room remains available for inferior non-credibly labelled and/or unlabelled products to invest on sophisticated promotion campaigns, using selective reporting or subtle normatively-charged colours and symbols associated with environmental friendliness, as well as non-credible certifications used for greenwashing, in

[14] Mandatory endorsement labelling amounts to the command-and-control measure of Qualified Market Access (QMA), in which case products must fulfil certification criteria to be allowed in the market. This is a rather strict form of state intervention that bans products not qualified for certification, subject to strict conditions by the WTO (Holmes, 2008, as cited previously). Accordingly, it falls outside the scope of ethical consumerism and therefore of this report.

[15] This list of limitations is not exclusive. Notably, the basic endorsement label form also fails to incentivize products to improve further above certification requirements. See section "Variations: comparative and dynamic (vs. static) labels".

order to communicate misleadingly positive environmental profiles to concerned consumers.

As a result, endorsement labels do not protect best performers from what Akerlof (1970)[16] termed as "dishonest sellers", who prefer to distort the market by investing in disguising -or, to use Milton Friedman's[17] term: in "cloaking"- their poor environmental performance rather than in improving it, to their own interest but to the cost of best performers, of consumer preferences, and -at least in the case of climate change- of society at large.

2. **Endorsement labels are most functional for revealing the preferences of strongly concerned and also of sufficiently high-income consumers** who can afford to pay systematically for the price premiums usually attached to labelled products. However, this consumer segment is a limited fraction of the market, leading the market for ethical products to remain a niche (albeit, with notable exceptions). On the contrary, stated and also revealed consumer preferences suggest that a majority of consumers are **moderately (yet-positively) concerned.** Such consumers are optimizers and not perfectionists with regards to ethical issues. They are much more likely to act in the market in order to avoid or to punish the worst performers (namely the most polluting products, or those that mistreat labour and animals), than to pay the financial, convenience, etc. costs required to reward the best ones.[18]

   Therefore, by not removing information asymmetry from poor performers, endorsement labelling does not provide crucial information that motivates and enables the systematic revealing of environmental preferences by moderately-yet-positively concerned consumers. Namely, endorsement labelling is not fit to capture the environmental preferences of a consumer segment that by reasonable assumption includes most of society.[19]

[16] Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." The Quarterly Journal of Economics 84 (3): 488–500

[17] Friedman, Milton. 1970. "The Social Responsibility of Business Is to Increase Its Profits." The New York Times Magazine, September 13. http://dx.doi.org/10.1007/978-3-540-70818-6_14

[18] The notable exception of egg consumption, in which ethical consumerism stopped being a niche and became mainstream, illustrates precisely the point that consumers are non-perfectionist optimizers. In 2016, the Noble Foods website (www.noblefoods.co.uk) reported that an increase in welfare conscious consumers within a ten-year timespan resulted in a 51.4% volume market share for Free Range eggs (which are middle-range with regards to animal welfare), while the previously dominant Cage eggs (worst animal welfare) fell to 44%. At the same time, the combined share of Barn, Organic and Speciality eggs (top animal welfare egg options) represented a mere 4.9% of the total UK eggs market. Similarly in The Netherlands, from 2006 till 2012 the share of free-range systems increased from 34% to 63%, whereas cage systems decreased from 47% to 18%. During this period, the combined share of organic and outdoor systems increased from 2% to 5%. Between 2012 and 2014 the market stabilized.
Other market evidence on the motivational power of information on poor ethical performances include costs paid by firms to communicate socially responsible brand images, security costs paid to prevent whistle-blower access to factory farms and sweatshops, costs paid to lobby policymakers against information disclosure and costs paid for ethical improvements when unethical practices are revealed.

[19] See also section "Effects on market allocation". For a more detailed discussion see Chapter 2 in: Michalopoulos, Tassos. 2016. "The Citizen Goes Shopping: A Framework for the Assessment and Optimization of Production from the Perspective of Society." Wageningen: Wageningen University. http://edepot.wur.nl/393482.

Notwithstanding their functionality drawbacks for market optimization, the overwhelming majority of available eco-labels adopt the above basic endorsement format.

Endorsement product labels of the basic type

Quantitative labels

The basic format of the quantitative label type refers to labels that indicate the actual performance of products for the indicators concerned. This information can be communicated using either a continuous (numerical scale) or a categorical (multi-grade, multi-level) format.[20] Accordingly, they can fall within the Type III ISO category, provided they also fulfil all other requirements concerned.

Unlike endorsement labels, the pure quantitative label form is purely descriptive and does not communicate a judgment on the superiority of product performance. However, **quantitative labels enable consumers** to make such judgments themselves, by providing them with the information needed to compare the performance of products for the issue(s) concerned.[21]

Quantitative labels **can be either compulsory or voluntary, with profound effects on their functionality for market optimization:**

- Obvious reputation effects on poor performers make the **voluntary** variants attractive only to average and above-average suppliers. Therefore, voluntary comparative labels do not typically inform consumers on poor product performance. This allows information on poor performers to remain asymmetric, to the effect that product performance comparisons that involve low-performing products are not enabled for consumers. Accordingly, similarly as described in the case of endorsement labels, poor performers can still exploit information asymmetry through sophisticated promotion campaigns to communicate misleadingly positive environmental profiles to concerned consumers. Consequently, the voluntary variation still allows for market distortion to the cost of average and best performers, and its functionality for consumer-based market optimization is therefore limited.
- In contrast, **compulsory** quantitative labels enable consumers to compare environmental performances across the entire range of available products, by informing them on the performance of all available product substitutes. They thus permit no market segment to suffer from information asymmetry for the indicators concerned, thereby limiting the space available for misleading promotion practices by low performers.

[20] Although this label type is generally referred to as "quantitative", the categorical format in particular enables labelling also for qualitative indicators besides quantitative ones. Such is the case of the multi-grade numerical marking on eggs used to denote the rearing system for laying hens, originally set from 0 (organic) to 5 (cage).

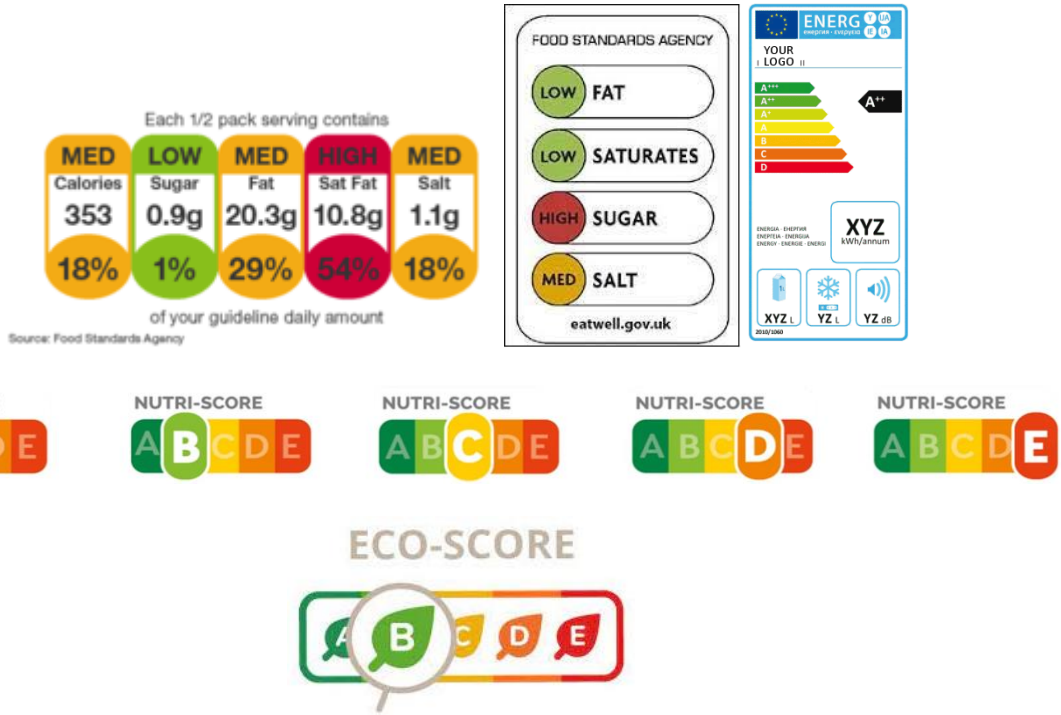
[21] For this reason, quantitative labels are sometimes inaccurately termed as "comparative". Comparative labels are discussed in the forthcoming section "Variations: comparative and dynamic labels".

Voluntary quantitative product label of the basic type

Variation: color-coded quantitative labels

A significant variation to the purely descriptive quantitative label form is the color-coded type. This variation typically concerns quantitative labels that employ a three to five-grade colour scale to signify different levels of product performance. Appropriately, the colour scale is normatively charged in an intuitive way to reduce the cognitive costs of product comparisons: Tones of green and of red are reserved, respectively, for highest and lowest quality products, while tones of amber usually denote middle-range products (average).

The several high-profile and successful applications of this label type include the Traffic Light nutritional label of the Food Standards Authority (FSA) in the UK, the French Nutriscore and the EU Energy label for electrical appliances. As depicted in the figure, these color-coded variants can contain little to no further information than the ranking of the product on the colour scale, while the display of the full colour scale itself is not always deemed necessary.



Color-coded categorical quantitative labels

Among these, in particular the success of the long-tested single-issue EU Energy label to drive the market towards more energy-efficient products has been remarkable. This makes the EU Energy label a familiar and useful case to study and learn from, concerning the functionality of this label type for market optimization.

Variations: comparative and dynamic (vs. static) labels

Comparative labelling is a variation of endorsement and quantitative labelling. Comparative labels of the endorsement type distinguish products of particularly superior performance, using as a reference the performance of competing products in the market. They are usually excellence-based. Excellence-based comparative endorsement labels are awarded only to products in the top range of performance. For instance, only to the X top performing products in each product category, where X can either be an absolute number or a percentage (e.g. top 10 products or top 10% of products).

Comparative quantitative labels come in different variations. One variation informs on how well a product performs as compared to the market norm (e.g. X % better than the market average). Another variation simply scores products on a scale set by the performances of competing products in the market.[22] This second subtype provides consumers with information about how well products perform within the observed range of competitive products' performances.

Accordingly, comparative labelling introduces an element of judgement that alters the functionality of the basic labelling formats.[23] Depending on the specifics of particular labelling applications, comparative labelling can improve label functionality insofar it contains information on how well a product performs as compared to its competitors. Thereby, it reduces the cognitive costs to survey and process information for consumers with relevant preferences.

Comparative labels (left: endorsement type; centre and right: quantitative type)

[22] For example, the US EnergyGuide label scores products on a "Cost Range of Similar Models" https://www.energy.gov/energysaver/appliances-and-electronics/shopping-appliances

[23] Particularly for endorsement labels, the comparative type *does not* inform that a product achieves a certain performance level (for instance, that it is sustainable or carbon neutral). Instead, it informs that a product is among the best alternatives for the indicator(s) concerned. This has been described as 'a lesser evil approach', because all product alternatives could be unsustainable. However, this criticism is inaccurate in case of products with positive overall impacts.

A notable functional feature of comparative labels is their dynamic character. Labels can be described as dynamic depending on the kind of values used to define their threshold levels. The said "threshold levels" of endorsement labels are the performance levels that must be achieved to grant a product label as set in the certification Standard. In the case of quantitative labels, the term refers to the performance levels that define the outer boundaries and the borders of the grades used to score products. Similarly, the threshold levels of labels that inform as compared to the market norm are the market standard or the market average. In all these cases, labels can be characterized as dynamic when their threshold levels are defined as a function of dynamic parameters, instead of being defined in absolute terms (i.e. instead of having fixed values).

As it follows, comparative labels are inherently dynamic because their threshold levels depend on market data on the performance of competing products. Accordingly, their labelling requirements **are bound by design to evolve automatically together with the evolution of the performance of competing products.**

In comparison, the threshold levels of static (i.e. non-dynamic) labels have absolute values that remain in principle fixed, unless and until updated. Obviously, static labels function as dynamic ones if updated in the same way. However, since static labels are not market based, the update of their threshold values requires some form of legitimization. This legitimization is generally achieved through a participatory approach to the update process.[24] In practice, this means that the **timing**, **content** and even the **direction**[25] of the evolution of static labels is generally subject to expert and stakeholder opinion. Consequently, as exemplified also by the case of the EU Energy label that is discussed in a following section, whether labels are dynamic, or static affects their updates and therefore their functionality in a decisive way.

> *What these results ultimately show is that the emergence of any quality standard is not the predetermined reflection of a rational or scientific process, but instead a negotiated process open to social dynamics within and external to the supply chain*
>
> *(Bush and Oosterveer 2015)*
>
> *The introduction of a label or a modification of its standard creates gainers and losers, who stand on opposite sides with respect to implementing such regulatory intervention. […] The relative political power of such groups, as well as the benefits at stake, will most likely shape the type of regulation finally observed.*
>
> *(Bonroy and Constantatos 2015)*

Another notable functionality limitation of non-dynamic labels is that they fail to incentivize further improvement once product performance has reached top grade (in case of categorical quantitative labels) or certification grade (in case of endorsement labels). To overcome this limitation in order to preserve the motivational power of the label for further improvement, requires successive upward revisions of the scoring scale or of the certification standard.

---

[24] Bush, Simon, and Peter Oosterveer. 2015. "Vertically Differentiating Environmental Standards: The Case of the Marine Stewardship Council." Sustainability 7 (2): 1861–83. doi:10.3390/su7021861.
See also, Bonroy, Olivier, and Christos Constantatos. 2015. "On the Economics of Labels: How Their Introduction Affects the Functioning of Markets and the Welfare of All Participants." American Journal of Agricultural Economics 97 (1): 239–59. doi:10.1093/ajae/aau088.

[25] Third-party certification Standards, when revised have reason to become both stricter and more relaxed. Stricter standards create incentive for further improvement to those labelled, while more relaxed standards increase the pool of products to which the certification appeals. Stakeholder opinion can favor either stricter (so as to differentiate best performers) and more relaxed standards (to reduce certification costs).

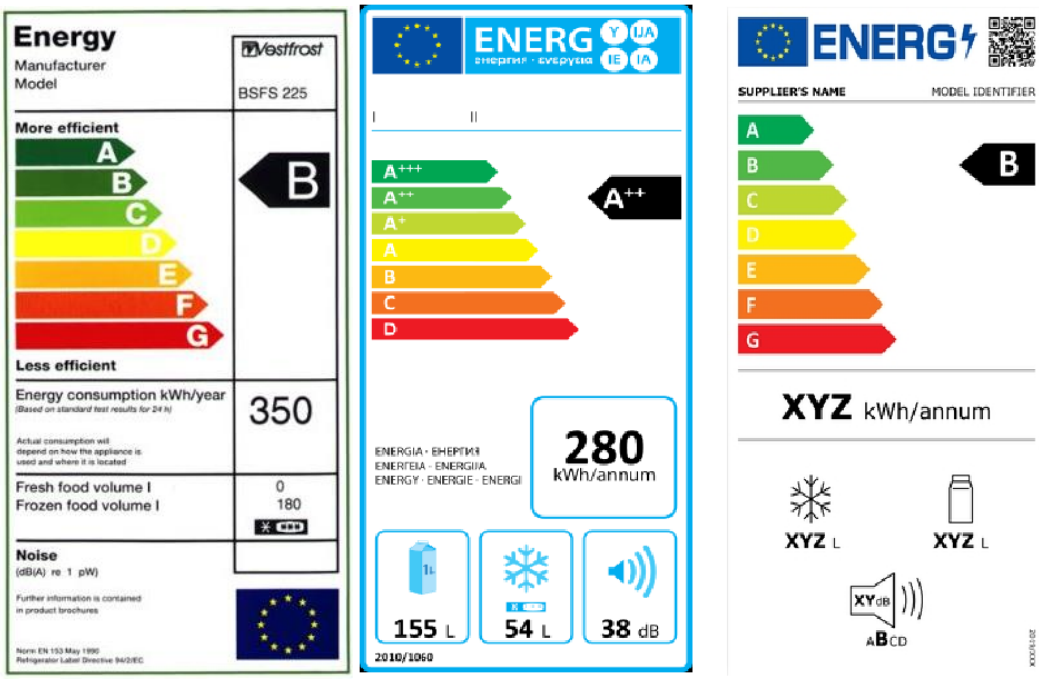
The functionality limitations of non-dynamic labelling are discussed in the following section, using the EU Energy label as a case study. The functionality of dynamic labels, and in particular of the quantitative categorical type, is discussed in the context of the labelling approach proposed at the section "Synthesis" of this report.

3.2.4. Case Studies

The EU Energy label

The original EU Energy label was described as "an undeniable market transformation success" (Egan and Waide 2005, p. 813). After its introduction in 1995, the unprecedented influence it exerted on consumers and producers sufficed to incentivize **market evolution that accumulated at the top energy efficiency grade around 90%** of refrigerators, dishwashers and washing machines by 2010 (European Commission 2010). In an EU-wide Eurobarometer survey in 2019,[26] 93% of consumers confirmed that they recognised the label and 79% confirmed that it had influenced their decision on what product to buy.

With regards to its basic format, the EU Energy label is quantitative, non-dynamic and uses a categorical format. The energy efficiency of electrical appliances is communicated using a traffic light color-coded and seven-grade scale, while raw data on key performance indicators is also provided. In its two revisions pictograms have gradually replaced text. **The impressive functionality of the original label** for stirring the market towards more efficient appliances has been well-documented. Among other possible factors, this functionality can be attributed to the compulsory character of the label, which forced information on all products per product category (including on the worst performers) to be disclosed, and also to the communication of this information using an intuitive color-coded scale. These greatly reduced the cognitive costs of making comparisons between all available substitutable products for consumers.



The evolution of the EU Energy label (from left to right: 1st version, 1995-2011. 2nd version, 2011-2021, 3rd version, 2021-present)

[26] https://ec.europa.eu/energy/en/data-analysis/eurobarometer-energy

Learning from the update process of the EU Energy label

The update process of the EU Energy label was less successful. The label relies on fixed (i.e. non dynamic) performance thresholds that use absolute values to define the borders of its grades. Consequently, the EU Energy label suffers from the common limitation of all such labels, which is that they do not incentivize products to improve further above the boundary of their top grade. To overcome this limitation requires successive upward revisions of the scale or of the certification standard, in order to preserve the power of the label to incentivize technological innovation and further improvement.

> *"The original energy label has been very successful, saving an average household in Europe several hundred euros per year and motivating companies to invest into research and development. Until the end of February, over 90% of products were labelled either A+, A++ or A+++. The new system will be clearer for consumers and ensure that businesses continue to innovate and offer even more efficient products. This also helps us to reduce our greenhouse gas emissions."*
>
> *Commissioner for Energy, Kadri Simson, 2021*
>
> *https://ec.europa.eu/commission/presscorner/detail/en/IP_21_818*

The participatory update process resulted in the 1st label revision between 2010 and 2012 (effective for refrigerators from 30/11/2011). Industrial **stakeholder resistance** to the intuitive update approach was documented,[27] preventing the update to focus simply on the threshold values that define the borders of grades on the scoring scale. The compromise achieved was counterintuitive: The original upper grades were preserved unaffected by the update, and enhanced class A grades (A^{+}, A^{++}, A^{+++}) were added atop the existing grade A. As a result, the 2nd version of the EU label displayed an inflated A class, which contained around 90% of refrigerators, dishwashers and washing machines already by 2010.[28]

> *Two solutions were suggested to overcome this problem. One option, favored by consumer and environmental organizations (ANEC/BEUC, 2008), was to maintain the existing seven-point scale ranging from A to G, but tighten the criteria on a regular basis, so that every year only the most efficient products would be A-labeled. A product which would be placed at the top of the scale in one year would then be reclassified into a lower efficiency class in another year. This option would require the inclusion of a date on the label indicating how long the label would be valid (BUND/DUH, 2009). The other option, backed by industry associations (CECED, 2007), was to extend the scale by means of introducing new categories 'beyond A'. The energy efficiency class of one particular appliance would remain unchanged over time so that no updated sticker needed to be attached on the appliances in the store (ECEEE, 2009).*
>
> *(Heinzle and Wüstenhagen. 2012)*

This update approach preserved middle and low efficiency products inside or close to class A, to the effect that in 2018 all refrigerators were labelled either A+, A++ or A+++, reducing their motivation

[27] Heinzle, Stefanie Lena, and Rolf Wüstenhagen. 2012. "Dynamic Adjustment of Eco-Labelling Schemes and Consumer Choice - the Revision of the EU Energy Label as a Missed Opportunity?" *Business Strategy and the Environment* 21 (1): 60–70. https://doi.org/10.1002/bse.722.
Ölander, Folke, and John Thøgersen. 2014. "Informing Versus Nudging in Environmental Policy." *Journal of Consumer Policy* 37 (3): 341–56. https://doi.org/10.1007/s10603-014-9256-2.

[28] European Commission. 2010. "Questions & Answers: New Energy Labels for Televisions, Refrigerators, Dishwashers and Washing Machines." MEMO /10/451. Brussels: European Commission. https://ec.europa.eu/commission/presscorner/detail/de/MEMO_10_451.

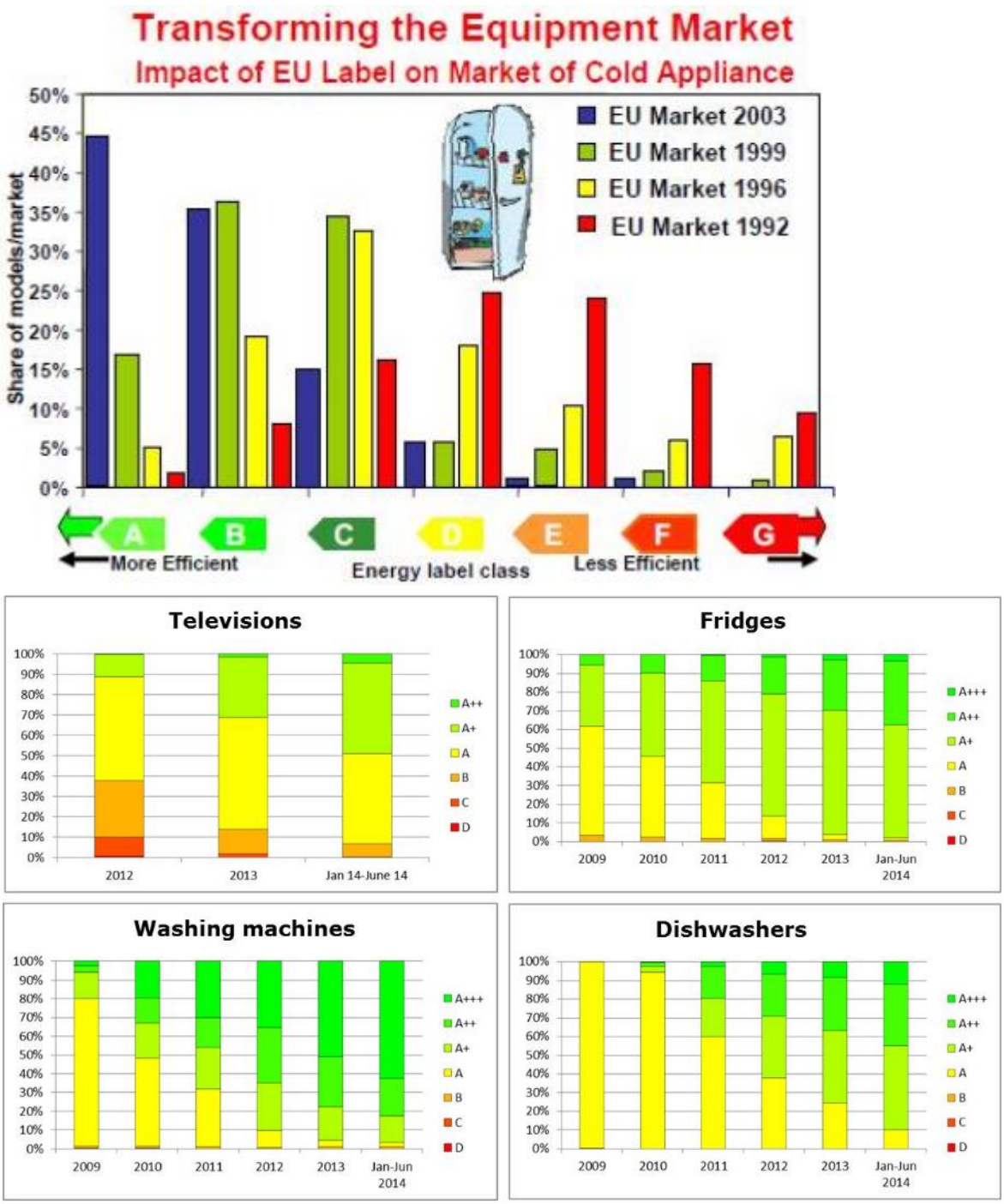
for further improvement. At that point the lowest label classes (D to A) contained no products. Nevertheless, it was possible to include void classes because **the EU Energy label is not comparative**. Preserving these classes in the labelling scheme **misled** consumers to believe that A+ class refrigerators were a comparatively superior product option, whereas in fact they were the least efficient option available.

> *[...] without most of the general public aware, the "worst" refrigerator that you can find in a shop currently is not G, as the energy label would suggest, but the "green" A+.*
>
> *[...] The rest of the scale on the energy label remains there, but with no real information value. Even though you find the B (or D) class on the label, you cannot buy a B refrigerator any more.*
>
> *EU energy label: a bumpy way towards good policy.*
>
> *The Bennett Institute for Public Policy, Cambridge, 2018.*
> *(https://www.bennettinstitute.cam.ac.uk/blog/eu-energy-label-bumpy-way-towards-good-policy)*

Consequently, the functionality of the label to drive positive change declined.[29] In practice, **for at least three years**, between 2018 and 2021, during which period consumers who bought A+ refrigerators trusting their EU Energy label class spent their incomes on products of the lowest available energy efficiency, the effectiveness of the label was reduced **against the good interests of concerned consumers and also of energy efficient suppliers**. Learning from this experience, to its credit the European Commission restored the original A to G energy efficiency scale, effective from 2021 for the first set of products.

Nevertheless, the functionality limitation identified above applies more generally to non-dynamic labels that base the legitimacy of their updates on complex and time-consuming negotiation processes with suppliers and other stakeholders. As shown in the relevant *Figure*, also the 1st version of the EU Energy label practically succeeded in almost exhausting its power to incentivize improvement within seven years after its introduction. This means that for the remaining period until its update, the label had little functionality for improving the energy efficiency of appliances.

[29] "Introducing the new label with its additional categories (A+, A++, A+++) weakened the efficacy of the label, resulting in lower consumer awareness of energy efficiency as an important attribute." Heinzle and Wüstenhagen (2012), as previously cited.



Energy efficiency evolution of selected appliances in selected EU markets after introducing the 1st and 2nd version of the EU Energy label

Up: Transformation of the EU Market for refrigerators and freezers 1992-2003

Down: Transformation of the market for selected appliances (2009-2014, for Belgium)

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52015SC0143&from=ET

Notwithstanding these shortcomings, **the new EU Energy label** remains non-comparative and non-dynamic, and still uses fixed consultation-based threshold values as borders of the different product classes.

Consequently:

(a) as product efficiency improves, new updates will be necessary to preserve the EU Energy label's power to incentivize technological innovation and further improvement.

(b) the interests of stakeholders wishing to minimize costs will be **a force opposing the effective timing and content** of the next update.

(c) the effectiveness of the new EU Energy label to incentivize positive change in the energy efficiency of electrical appliances based on consumer demand, will be short-lived as compared to its life-span (i.e. as compared to the time required until its next update).

> *The new categories for the rescaled label were agreed after a rigorous and fully transparent consultation process, with the close involvement of stakeholders and Member States at all stages, scrutiny by the Council and the European Parliament and with sufficient involvement of and notice provided to manufacturers.*
>
> *European Commission - Press release, Brussels, 1 March 2021*
> *https://ec.europa.eu/commission/presscorner/detail/en/IP_21_818*

The Eco-Score labels

The ongoing discussions for the development of a EU-wide mandatory nutritional FoP label, revived discussions on the potential introduction of a mandatory label that informs on the environmental impact of food products. The Eco-Score label, variations of which are already applied by major retailers in EU,[30] is currently viewed as a promising candidate.

The Eco-Score is a qualitative categorical type of label. It is not comparative and not dynamic. It scores products on a five-grade scale that is communicated using a normatively charged intuitive color-coded based on the Traffic Light System. Several label designs are put forward, including ones that depict the entire scale on which specific products are scored, and also more minimalistic variations that focus mainly on the color-coded of the particular grade attributed to a specific product.

The Eco-Score label in different label designs

The method used to calculate the Eco-Score of products is based on the Nutri-Score approach as applied to environmental concerns. It relies heavily on LCAs using available product category data[31] with bonus/malus for additional environmental indicators and criteria not covered by the LCAs available. These criteria include product certification with prominent sustainability labels of the endorsement type. The LCAs used are framed by the ISO 14044 standard and comply with the

[30] "Eco-Score's European expansion: Lidl and Colruyt adopt environmental footprint labelling" https://www.foodnavigator.com/Article/2021/04/09/Eco-Score-s-European-expansion-Lidl-and-Colruyt-adopt-environmental-footprint-labelling (Accessed 09/04/2021)

[31] Available Eco-Score applications rely heavily on the AGRIBALYSE® database for LCAs. (https://docs.score-environnemental.com/english/presentation) (https://doc.agribalyse.fr/documentation-en/)

Product Environmental Footprint (PEF) method proposed by the European Commission as a common way of measuring the environmental performance of products.[32] The PEF details the recommended Life Cycle Assessment (LCA)-based method to quantify the environmental impacts of products, including the flows of material/energy and also the emissions and waste streams associated with products (goods or services) throughout their life cycle.

The **positive functionality traits** of the Eco-Score label for consumer-based market optimization include its intuitive categorical color-coded format. This enables quick comparisons between products of different categories by consumers. A compulsory application of the Eco-score can enable environmental comparisons across the entire range of available products, creating also an incentive for the improvement of most polluting ones.

However, the Eco-Score label also includes a number of features that limit its functionality for the consumer-driven environmental improvement of the food supply sector. **Less functional features include**:

Neither comparative nor dynamic

The Eco-Score method calculates the score of each product separately and then assigns it to one of the predefined grades used, based on the score it achieves. The performance thresholds that define the borders of these grades use absolute fixed values, which are not a function of the performance of competing products. Consequently, the Eco-Score label is bound to have similar functionality limitations as the EU Energy label.[33] Specifically:

a. Significant cognitive costs for consumers remain. Consumers have to compare available products by themselves in order to confirm whether a particular product is above average, average, or below average in its category. A simple glance on the label does not suffice to retrieve that information, simply because all available products could be in the same range of classes, or even have the same class (e.g. all competing products could be in classes A and B, or in class A, or in class B , etc).

b. Particularly when lower grades are void and all substitutable products accumulate at top grades, the label can mislead consumers into believing that they purchase environmentally superior product alternatives when in fact they do not (see EU Energy label)

c. Even if the above (a and b) are not true during the first period after label introduction, then it will become true after a while. Provided that the label works to incentivize improvement, market evolution will incentivize the accumulation of products at the top grades. Then and until the next update of its scale, the label will stop being functional for product comparisons and market optimization.

d. Assuming a similarly complex participatory update process as the one used for the EU Energy label, the effectiveness of the Eco-Score label to incentivize positive environmental change based on consumer demand will be short-lived as compared to its life-span (i.e. as compared to the time from its introduction until its update). This effect cannot be afforded in case of pressing environmental issues.

e. Even for the time-span during which it will remain motivational, the Eco-Score label will suffer from the common drawback of all static labels, which is that they do not incentivize improvement above top grade.

Moreover, additional functionality limitations specific to the Eco-Score approach, as currently applied, include:

[32] https://eplca.jrc.ec.europa.eu//EnvironmentalFootprint.html,
[33] See section "*Case study: The EU Energy label*", above, for more details.

Insufficient product differentiation

At present, Eco-Score applications rely on LCA data from the invaluable AGRIBALYSE® database. These data distinguish the environmental performance of foods only at the level of product categories averages (example: cow's milk yogurts, hamburgers, etc.). Therefore, the differences in the environmental performance of competing product substitutes within the same category is not identified (e.g. two yogurts from competing brands or an organic hamburger versus a conventional hamburger). Consequently, a EU-wide application of the Eco-Score label based on this approach would lack the functionality to shift consumption towards the most environmentally-friendly products within each food category.

The Eco-Score method addresses this issue **partially**, by awarding additional bonus/malus points **selectively** to certain sustainability-related certifications. Moreover, it incorporates additional indicators that include impacts on endangered species, the type of materials used, the environmental standards at the country of origin, etc. These differentiate products within categories to some extent.

However, the good function of any EU-wide application of comparative environmental label requires an **improved,** and as thorough as practically possible,[34] **data-based** product differentiation of available product substitutes within all product categories concerned:

a. By relying heavily on **selected** certification schemes to differentiate product performance within categories (regardless how appealing these certifications might be), the Eco-Score introduces bias in favour of the particular production systems. This bias limits Eco-Score's functionality for rewarding /incentivizing performers that invest in environmentally positive **technological innovation** when this is not covered by the selected certification systems.

   In the case of GHG emissions, the development of such technological innovation is of crucial importance. It concerns a range of much-sought technological breakthroughs in the fields of carbon emissions minimization, capture and offsetting with tangible positive impact on the minimization of GHG emissions.[35]

b. Labels that differentiate products only at the level of product categories can be functional for shifting consumption between substitutable product categories. However, the environmental performance ranges of different product categories can partially overlap: The best products of a category that is on average environmentally inferior, can be better than the worst products of a category that is on average environmentally superior. Consequently, the incomplete differentiation of product performance within food categories can direct consumers towards environmentally worse choices.

   Besides, shifting consumption between product categories is an ambitious long-term task that requires high levels of consumer concern and commitment. Therefore, it is questionable whether a public policy that focuses on shifting consumption between product categories can bring the short term results required to combat climate change. Namely, a realistic course of action requires, for instance, promoting less polluting meat options to moderately concerned meat-eaters, alongside the promotion of reducing or quitting meat consumption.

[34] To minimize costs, these assessments could still largely rely on available data about national and sectoral averages, while also welcoming inputs for downward corrections by civil society actors, and for upwards corrections by individual suppliers than provide relevant data. Nevertheless, additional research might also be required.

35 Cultured meat is another example of technological innovation whose improved impact on the environment (and also on animal welfare) is not captured by this approach. This technology is a form of cellular agriculture and refers to meat produced by in vitro cell cultures of animal cells (as opposed to meat obtained by slaughtering animals). From an environmental and also from an animal welfare perspective, synthetic meat can be a superior alternative for meat consumers.

Edelman, P.D., D.C. McFarland, V.A. Mironov, and J.G. Matheny. 2005. "Commentary: In Vitro-Cultured Meat Production System." Tissue Engineering 11 (5–6): 659–62. https://doi.org/10.1089/ten.2005.11.659.

Frequently Asked Questions:

7. Why do all the products in the same category (example: cow's milk yogurts) have the same environmental impact ratings?

The database contains a single set of environmental indicators for each of the 2,500 food products. The data do not allow comparison of food products of the same category (e.g. two yogurts from competing brands or an organic hamburger versus a conventional hamburger). These data therefore reflect an "average" reality which allows the comparison of food products of different categories, but not the comparison of identical products from different production methods. To carry out this fine comparison work, it is necessary to collect additional data and adjust the LCA parameters in the software. Refer to the user guide and the methodological report for more information.

The AGRIBALYSE® database.

https://doc.agribalyse.fr/documentation-en/frequently-asked-questions-faq/frequently-asked-questions#7-why-do-all-the-products-in-the-same-category-example-cows-milk-yogurts-have-the-same-environmental-impact-ratings

Aggregated

The Eco-Score is a multi-issue label that aggregates several environmental impact categories to provide an overall environmental indicator in the form of a grade score for each product. Namely, it does not inform on the non-aggregated scores that products achieve for each environmental issue. Such aggregate environmental indicators have significant practical advantages because they allow for the quick overall comparison of different product options and are also easy to communicate to the public. This makes aggregate environmental indicators understandably desirable in public policy.

However, the meaningfulness and effectiveness of aggregating different environmental impact categories is **controversial** at best. Relevant objections focus on the "weighing" of the relative importance (or contribution) of different impact categories, which is an essential step of the aggregation process. Weighting is not a scientific process. Instead, it relies on subjective or inter-subjective value judgments that effectively define the trade-offs (compensations) allowed[36] between product performance for different impact categories, notwithstanding that these categories are in fact incommensurable.[37] Accordingly, the requirements and guidelines set at the **ISO 14044:2006**[38] Standard, which frames Life Cycle Assessment for environmental management, **forbids it explicitly**: "weighting [...] **shall not be used in LCA** studies intended to be used in comparative assertions **intended to be disclosed to the public**." (Sec. 4.4.5)

"Because weighting is not a scientific process, it is vital that the weighting methodology is clearly explained and documented. Although weighting is widely used in LCAs, the weighting stage is the least developed of the impact assessment steps and also is the one most likely to be challenged for integrity. [...] Several issues exist that make weighting a challenge. The first issue is subjectivity. According to ISO 14042 [:2000], any judgment of preferability is a subjective judgment regarding the relative importance of one impact category over another. Additionally, these value judgments may change with location or time of year. [...] The second issue is derived from the first: how should users

[36] Or partially allowed, in case of "non-compensatory" approaches (i.e. fuzzy aggregation methods, outranking matrices)

[37] Weighing the relative importance (or contribution) of different impact categories, cannot be fully avoided in the process of aggregating different impact categories. To simply use averages implicitly implies that the impact categories have equal weights; which is as an arbitrary assumption as any other.

[38] Note that the ISO 14044:2006 Standard was last reviewed and confirmed in 2016. Therefore this version remains current. https://www.iso.org/standard/38498.html

fairly and consistently make decisions based on environmental preferability, given the subjective nature of weighting? Developing a truly objective (or universally agreeable) set of weights or weighting methods is not feasible." (SAIC 2006).[39]

"ISO 14044:2006 generally advises against weighting, stating that 'weighting, shall not be used in LCA studies intended to be used in comparative assertions intended to be disclosed to the public.' This advice is often ignored, resulting in comparisons that can reflect a high degree of subjectivity due to weighting" (Trusty 2010).[40]

"If desired the LCA study can be concluded with a single figure, or environmental index, in which each environmental problem is weighted in terms of its importance. This figure or index allows an easy and direct comparison of different products or options. The weights used are of course subjective." (De Haes and Van Rooijen 2005).[41]

"Different interpretations of the weights assigned to each of the principles by different people preclude a definitive ethical judgment" (Mepham et al. 2006). [42]

Nevertheless, and while acknowledging the issue, a Technical Report[43] by the Joint Research Centre (JRC), the European Commission's science and knowledge service, does arrive at a set of weights for calculating the aggregate Environmental Footprint (EF) of products using LCA. The proposed weights are based on a thorough and well-documented comparative assessment and combination of existing weighting approaches. The weights proposed refer to the 13 environmental impact categories used by the EF and also by the AGRIBALYSE® database, and consequently by the Eco-Score label.

Any weighting scheme is not mainly natural science based but inherently involves value choices that will depend on policy, cultural and other preferences, and value systems. No "consensus" on weighting seems to be achievable. This situation does not apply only to weighting in a LCA or EF context but seems inevitable for many multicriteria approaches. However, weighting is seen as essential to further aggregate information with the objective to provide better support in complex decision situations. (p.6)

The recommended weighting set, robustness factors and final weighting factors excluding toxicity-related impact categories. (p.5)

[39] SAIC. 2006. "Life Cycle Assessment: Principles And Practice". U.S. Environmental Protection Agency. http://www.epa.gov/NRMRL/lcaccess/pdfs/600r06060.pdf

[40] Trusty, Wayne. 2010. "An Overview of Life Cycle Assessments: Part One." International Code Council Building Safety Journal 8 (5). http://bsj.iccsafe.org/2010Oct/features/lca.html.

[41] De Haes, Helias A. Udo, and Martijn Van Rooijen. 2005. "Life Cycle Approaches The Road from Analysis to Practice". United Nations Environment Programme UNEP/SETAC Life Cycle Initiative. http://www.unep.fr/shared/publications/pdf/DTIx0594xPA-Road.pdf.

[42] Mepham, Ben, Matthias Kaiser, Erik Thorstensen, Sandy Tomkins, and Kate Millar. 2006. Ethical Matrix: Manual. The Hague: LEI.

[43] Sala S., Cerutti A.K., Pant R., Development of a weighting approach for the Environmental Footprint, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-68042-7, EUR 28562, doi 10.2760/945290

	Aggregated weighting set	Robustness factors	Intermediate Coefficients	Final weighting factors (incl. robustness)
	(A)	(B)	C=A*B	C scaled to 100
Climate change	15.75	0.87	13.65	**22.19**
Ozone depletion	6.92	0.60	4.15	**6.75**
Particulate matter	6.77	0.87	5.87	**9.54**
Ionizing radiation, human health	7.07	0.47	3.30	**5.37**
Photochemical ozone formation, human health	5.88	0.53	3.14	**5.10**
Acidification	6.13	0.67	4.08	**6.64**
Eutrophication, terrestrial	3.61	0.67	2.40	**3.91**
Eutrophication, freshwater	3.88	0.47	1.81	**2.95**
Eutrophication, marine	3.59	0.53	1.92	**3.12**
Land use	11.10	0.47	5.18	**8.42**
Water use	11.89	0.47	5.55	**9.03**
Resource use, minerals and metals	8.28	0.60	4.97	**8.08**
Resource use, fossils	9.14	0.60	5.48	**8.92**

Development of a weighting approach for the Environmental Footprint. [44], JRC Technical Reports[45],

Consequently, the aggregation of impact categories by the Eco-Score label results in the following functionality limitations:

1. The aggregated Eco-Score approach permits trade-offs between different impact categories. These allow suppliers to **focus primarily on the most cost-effective improvements**, instead of focusing on the most important or urgent ones, or on those most preferred by consumers.
2. The aggregated approach **reduces the effectiveness of the Eco-Score label for reducing GHG emissions**. While all environmental impact categories are undoubtedly highly important, the **severe and overarching** catastrophic effects of human-made climate change foreseen at all levels (social, economic, and also on the biodiversity and on broader ecosystems), as well as the **urgency** stemming from the limited time available to be effective in mitigating it, **dictate** that the maximal functionality for combating climate change should be the **top priority** of any EU-issued environmental label. This is not in line with a relative weight of 22,19% for the Climate Change impact category.

Seen in the light of this priority, it is **counterintuitive** to read that the above JRC Report **counts the success** of the single-issue carbon footprint approach **as a reason against** single issue labelling, and in favour of an aggregated multi-issue approach.

> *Weighting might also help in achieving greater accountability. Optimization towards 15 indicators is much more difficult than optimization towards 1 variable and this was seen as one of the reasons for success of single* [sic] *existing single issue approaches like carbon footprint.* (p.11). [*Development of a weighting approach for the Environmental Footprint. [46], JRC Technical Reports[47],]*

To be precise, it is strictly speaking inaccurate to state in the above JRC report (p.11) that "*Weighting might also help in achieving greater accountability*": weighing (i.e. aggregation) enables **lower accountability for individual issues**, by spreading overall accountability over a greater variety of issues.

[44] Sala S., Cerutti A.K., Pant R., Development of a weighting approach for the Environmental Footprint, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-68042-7, EUR 28562, doi 10.2760/945290

45 https://ec.europa.eu/environment/eussd/smgp/documents/2018_JRC_Weighting_EF.pdf

[46] Sala S., Cerutti A.K., Pant R., Development of a weighting approach for the Environmental Footprint, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-68042-7, EUR 28562, doi 10.2760/945290

47 https://ec.europa.eu/environment/eussd/smgp/documents/2018_JRC_Weighting_EF.pdf

3.3. Comparative Analysis and Synthesis

This section first summarizes the comparative typology of functionality-related features of available labels. Following that, the effects of the two main label types on the allocation function of the market are analysed. Next, a novel Dynamic Categorical Comparative labelling approach is presented. Selected functionality aspects of the Dynamic Categorical Comparative labelling approach and of Eco Score labelling are compared.

3.3.1. Comparative typology of functionality-related features

The table below summarizes the functionality features of the most relevant FoP label types available, as critically discussed at section 2.2. The "DCCL" entry of the last column summarizes the features of the Dynamic Categorical Comparative labelling approach introduced at section 2.3, as a synthesis of the most effective functionality features.

Table . Comparative analysis of functionality features in relevant FoP labels available.

Label type / Features	EU Ecolabel	Carbon Neutral	Blue Angel	Carbon Footprint Lower CO2	Carbon Footprint	US Energy Guide	EU Energy label	Eco- Score v.1	Eco- Score v.2	CDCC
voluntary vs. compulsory	voluntary	voluntary	voluntary	voluntary	voluntary	compuls.	compuls.	voluntary	compuls	compuls
endorsement vs. quantitative	endors.	endors.	endors.	quant.	quant.	quant.	quant.	quant.	quant.	quant.
provides information on poor performers	no	no	no	no	no	yes	yes	no	yes	yes
comparative, dynamic	no	no	yes	yes	no	yes.	no	no	no	yes
categorical color-coded	n/a	n/a	n/a	no	no	no	yes	yes	yes	yes
multi issue/ aggregated vs. single issue / non aggregated	multi issue	single issue	multi issue	single issue	single issue	single issue	single issue	aggr. (either possible)	aggr. (either possible)	non aggr. (either possible)

DCCL: Dynamic Categorical Comparative labelling, as introduced at section "Synthesis", below

3.3.2. The effect of endorsement vs. quantitative labels on market allocation

A significant difference in the functionality of available label types concerns their effect on the allocation function of the market. Specifically, on the distribution of environmental impacts that they incentivize among suppliers subject to relevant consumer preferences. This effect depends mainly on the number of grades a label uses. Therefore, the main distinction in this respect is among labels of the two main types: endorsement vs. quantitative ones.

The qualitative model of Figure 1a illustrates this effect by using the simple case of one single-grade endorsement label operating in the market. As observed by Van Tongeren et al. (2009), "With [single-grade] labelling there are two prices clearing the market, since the label makes it possible to identify the two qualities by segmenting the market".[48] Assuming only domestic production (no imports), the said "two qualities" correspond to the domestic production (a) of conventional products that must comply only with applicable environmental legislation, and (b) of certified production that must also comply with certification requirements. Correspondingly, 'T_L' represents the legally allowed threshold

[48] Van Tongeren, F., J. Beghin, and S. Marette. 2009. "A Cost-Benefit Framework for the Assessment of Non-Tariff Measures in Agro-Food Trade". 21. OECD Food, Agriculture and Fisheries Working Papers. OECD Publishing. http://dx.doi.org/10.1787/220613725148

for the environmental attribute at stake as set in domestic and applicable international law, while 'T_C' represents the threshold value required for the same attribute by the certification scheme behind the label assumed. Other qualities besides these two (i.e. any other performance level for any unobservable attribute that is not represented by a labelling grade) will not be identifiable by consumers and therefore it will not yield a new price to incentivize producers to achieve that performance level. As observed in reviewed literature (Tlusty 2012),[49] single-grade labels "translocate" the distribution of producer practices by creating a zone of "pull" before the certification threshold and a sharp decline after: Suppliers performing level near below the certification requirements are incentivized to improve and get certified, while suppliers far below these requirements have less incentives to improve because certification "is likely beyond their technical or financial means". Moreover, the shape of the distribution curve above the threshold is usually[50] concave because certified suppliers lack incentives to improve beyond threshold (Tlusty 2012), while all suppliers are subjected to a continuous rational incentive to minimize costs. Figure 1a summarizes these observations and also applies the same reasoning to the legal threshold 'T_L' to illustrate the overall distribution of environmental impacts incentivized among suppliers. To pronounce the mismatch between the distribution incentivized and the actual[51] environmental preferences of consumers, normal[52] bell-shaped curve is included.

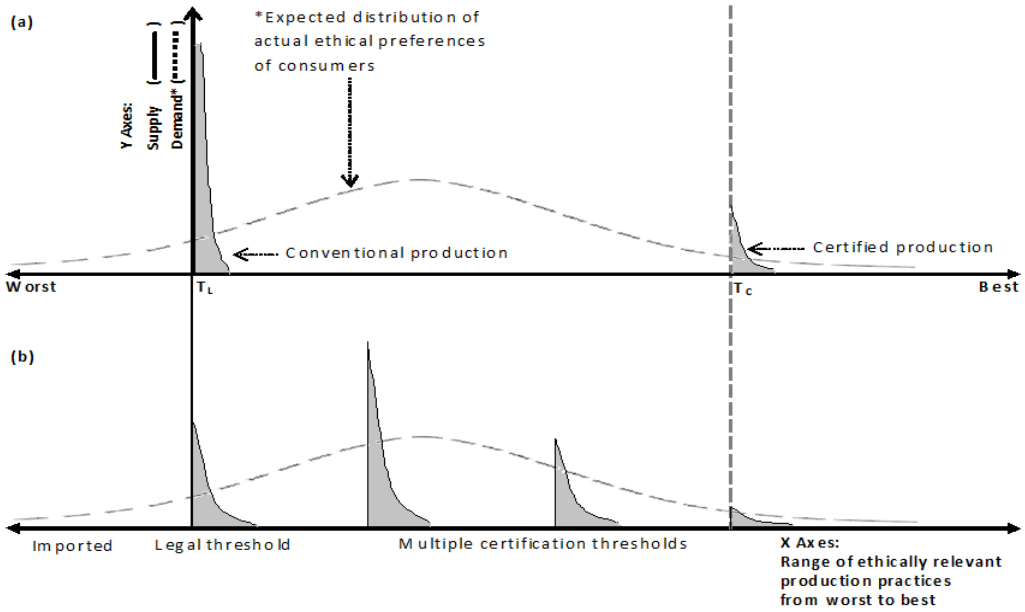


Figure 1: 'T_L', 'T_C': Threshold values that represent, respectively, the legal and certification requirements for an ethically relevant production practice. (a) The distribution of ethically relevant production practices (supply) that is incentivized by a single-grade label, as compared to the expected distribution of actual ethical preferences of consumers (demand). (b) Modification of the above in the case of multi-grade schemes. (Michalopoulos 2016)

[49] Tlusty, Michael F. 2012. "Environmental Improvement of Seafood through Certification and Ecolabelling: Theory and Analysis." Fish and Fisheries 13 (1): 1–13. doi:10.1111/j.1467-2979.2011.00404.x.

[50] Except in rare cases where environmental improvement is positively correlated to economic efficiency.

[51] The term "actual preferences" refers to the revealed environmental preferences of consumers (demand) under conditions of no information asymmetry about the environmental performance of available products. An estimate of this indicator can rely on research data and indirect market evidence suggesting substantial (but not extreme) consumer concern about environmental issues.

[52] The precise shape of the distribution of consumer preferences does not affect the validity of the argument. To invalidate the argument requires a distribution moved or skewed towards T_L to the extent that it would be sufficiently approximated by the shaded area of Figure 1a. This is unlikely because it would suggest widespread satisfaction with the present state of the environment.

Accordingly, the figure above schematically illustrates that the distribution of environmental practices (supply) that is incentivized by a single-grade label is sub-optimal insofar it deviates substantially from the actual preferences of consumers (demand), **causing the market for environmentally friendly products to fail.** An important observation is that the consumer segment not served by endorsement labels consists of moderately concerned consumers (preferences above T_L but below T_C). This is particularly concerning because, by reasonable assumption, the moderately concerned consumer segment includes most of society.

What is Market Failure?

Market failure is the economic situation defined by an inefficient distribution of goods and services in the free market. In market failure, the individual incentives for rational behavior do not lead to rational outcomes for the group.

In other words, each individual makes the correct decision for him or herself, but those prove to be the wrong decisions for the group. In traditional microeconomics, this can sometimes be shown as a steady-state disequilibrium in which the quantity supplied does not equal the quantity demanded.

Market Failure Definition, INVESTOPEDIA

https://www.investopedia.com/terms/m/marketfailure.asp

Market failures are often associated with public goods, time-inconsistent preferences, information asymmetries, non-competitive markets, principal–agent problems, or externalities.

Market Failure. Wikipedia, https://en.wikipedia.org/wiki/Market_failure

In comparison, Figure 1b schematically illustrates that a multi-grade label can incentivize the market to arrive at a distribution of supplier practices that approximates more closely the distribution of preferences among consumers. This is true for any possible shape of the demand distribution.

The simplified model of Figure 1 differs from reality in at least two important respects:

In reality, a large number of single-grade schemes with similar standards often coexist in the same markets. Consequently, one can be inclined to expect that when taken together, a plurality of single-grade schemes, or even a set of "tiered" labels,[53] could spontaneously work synergistically to function as a multi-grade one. However, and besides other requirements (Tlusty 2012, as cited previously):

- The synthesis of multiple single-grade labels into a composite multi-grade one requires to compare these single-grade labels with regards to their performance for a common dimension of consumer concern (e.g. GHG emissions). Consumers should be informed on this ranking for each environmental concern, in order to become enabled to make comparisons between differently labelled products. Unless this requirement is met, **a large number of single-grade schemes does not function as a multi-grade one**.
- Furthermore, even if the above condition would be satisfied, any synthesis of single-grade voluntary labels into a multi-grade scheme will be limited to high performance grades. This is because the communication of information on unfavourably low product performance is unlikely in a voluntary certification context. Therefore, the worst performers will not be identified for consumers willing to reveal preferences on their production. Therefore, the synthesis of **a large number of voluntary single-grade schemes does not function as a multi-grade mandatory one**.

[53] For instance, the system of "tiered" single-threshold certifications that developed around the Marine Stewardship Council (MSC) certification (Bush and Oosterveer 2015, as previously cited).

In reality imports exist. This means that moderately concerned consumers (whose preferences are above T_L but below T_C) unknowingly also encounter in the market products that perform even below the minimal legal standards T_L set for domestic production. This happens because state regulation cannot usually prevent **environmental dumping,** namely market access to products (or parts of) from locations with lesser environmental standards, a segment of which uses polluting technologies to minimize costs (Holmes et al. 2008, as cited previously)**.** Consequently, moderately concerned consumers must gather and process by themselves environmental information on unlabeled products (or on non-credibly labelled products) in order to reveal preferences above T_L but below T_C. This process, as well as discerning what part of the confusing plurality of available information is honest, imposes added cognitive costs that restrict the rationality of moderately concerned consumer choices. Expectedly, they also lead much of the moderately concerned consumer segment to quit the task of revealing environmental preferences altogether.

In comparison, compulsory multi-grade schemes tackle this issue by enabling consumers to identify such products at the lowest label grade. By enabling moderately concerned consumers to reveal preferences against environmental dumping, multi-grade labels also provide some **protection to law-abiding domestic production from unfair competition by environmental polluters** that exploit information asymmetry to externalize the environmental costs of their production. Worst polluters become differentiated from law-abiding conventional domestic production on the basis of their environmental impact.

Overall, the above analysis and also on the critical typology presented at section 2.2, indicates that **compulsory quantitative labels are by far more functional than voluntary endorsement ones** for the purpose of incentivizing systemic consumer-driven positive change in the market.

3.3.3. Synthesis: Compulsory Dynamic Categorical Color-coded labelling

Nevertheless, as already discussed, the functionality of different types of quantitative labels also varies. To accommodate the identified limitations of quantitative labels, this report proposes for consideration of a novel labelling approach as a synthesis of the established label types previously discussed. The proposed label is of the quantitative type. Its main features include that it is Compulsory, Dynamic, Categorical, and Color-coded (**CDCC**). As described in the following sections, it is designed to incentivize (a) **objective**, **market-based** and **consumer-driven**, and also (b) **continuous** and **open-ended improvement** of the environmental performance of production.[54] Figure 2 depicts a non-stylized presentation of the basic CDCC label concept, using as an example a 5-grade and single-issue label variation.

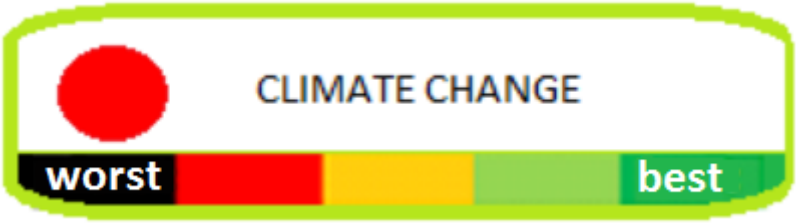


Figure 2: Non-stylized presentation of the basic CDCC label concept (5-grade and single-issue variation)

[54] The presented labelling approach was initially described within the context of research conducted at the Business Economics and Philosophy Groups of Wageningen University, The Netherlands. For more details see Michalopoulos, 2016, Ch.2, as previously cited.

Working Principle of virtuous cycle functionality

The functionality of the CDCC label is described along its implementation phases: (1) label introduction, (2) market evolution and (3) label update. These are followed by successive repetitions of phases 2 and 3. For reasons of simplicity, Figure 3 uses a 3-grade label format to illustrate the working principle of the CDCC label functionality. Next, the main steps of the entire process are outlined at Box 1.

Phase 1: Label Introduction. At phase 1 the label is introduced per category of products. The scale on which products are scored is comparative. Similarly, to the US Energy Guide label that scores products on a "Cost Range of Similar Models",[55] the CDCC label scores products on a scale that is dynamic and market defined. Accordingly, the outer boundaries of the scoring scale are defined impartially by the best and worst performers in each product category for the indicator concerned. For instance, by the performances of the domestically available substitutable products that have e.g. the highest and lowest GHG emissions, respectively, within a product category.[56] Similarly to the US Energy Guide and EU Energy labels, the CDCC label is compulsory in order to incentivize improvement also among the most polluting products, (which are generally unlikely to apply the label voluntarily,) by removing information asymmetry on their performance.

Similarly, to the EU Energy and Eco-Score labels, the CDCC is categorical. In line with the rationale of the Eco-Score label a middle grade is included to avoid dichotomous thinking, and a normatively charged color-coded based on the Traffic Light System is used to minimize cognitive costs for product comparisons.

The threshold values that define the boundaries between grades are calculated by dividing the distance between the best and worst observed performance (i.e. entire scoring range) in even parts, the number of which equals the number of grades decided. Therefore, the threshold values that define the borders between grades are also impartial, market-based and therefore dynamic. Consequently, the grade or class assigned to products does not depend only on product performance (as is the case for non-comparative labels like the EU Energy and Eco-Score). The grade of a product depends also on the performances of the best and worst competing products, and it can change as these performances change, even if a product's own performance remains unchanged. This becomes evident at phase 3. Accordingly, the upper part of Figure 3 illustrates the scoring of available products on a scale consisting of three equal grades according to their environmental impact.

Phase 2: Market Evolution. At phase 2 the market is offered time to evolve. During that period, product grade updates are possible if sufficient information is provided by suppliers. New products whose performance exceeds the upper boundary of the scale are graded as top grade, and *vice versa*. However, the inner thresholds that define the boundaries between different grades on the scoring scale are not updated.

During phase 2, the CDCC label incentivizes the improvement of product performance across all grades. The incentives applied to worst performers are straightforward and well-documented in the literature. As mentioned, they include reduced consumer demand, lower employee morale, lower ability to attract high quality employees and also divestment from ethical funds. The incentives applied to average and best performers become evident at phase 3. Accordingly, the middle part of Figure 3 illustrates the evolution of environmental impacts of products that the CDCC label incentivizes during phase 2.

[55] https://www.energy.gov/energysaver/appliances-and-electronics/shopping-appliances

[56] This straightforward approach serves the purpose to illustrate the functionality of the CDCC label; real-life applications mightrequire adaptations to maximize label functionality. See section "Fine-tuning and enhancements"

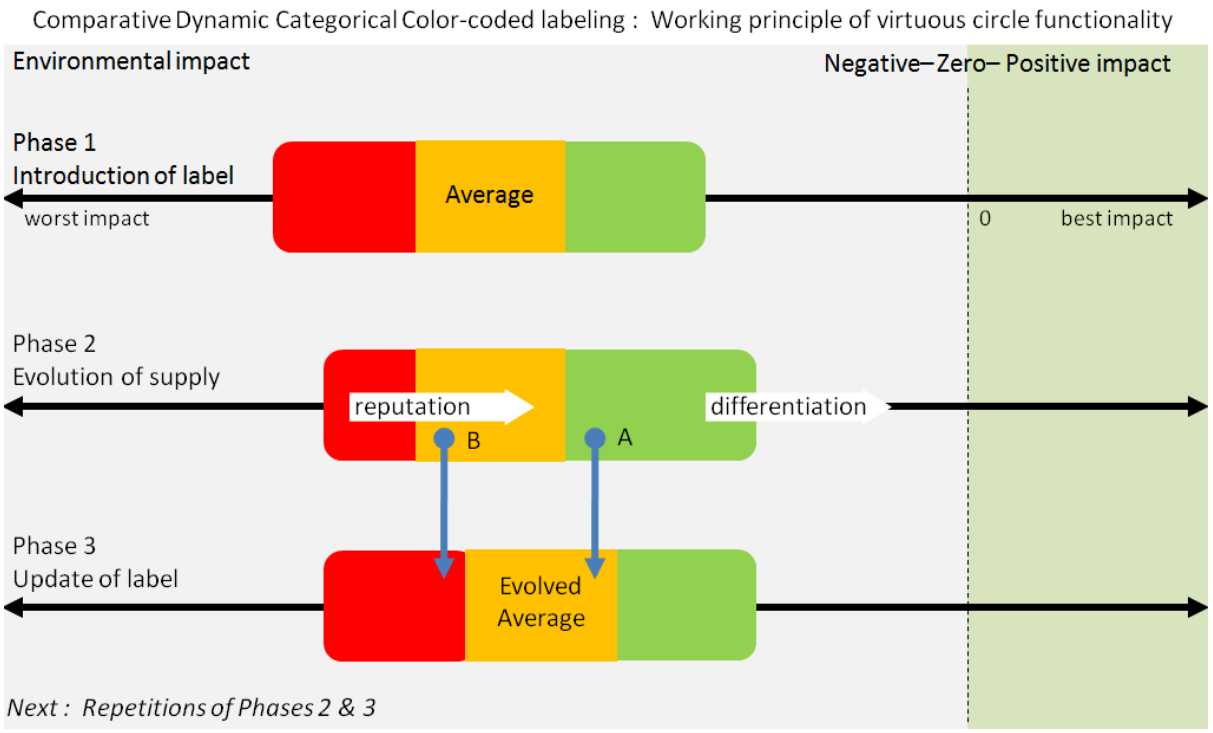


Figure 3. Schematic illustration of the positive translocation effect on the observed range of environmental impacts of competing products that is incentivized by the introduction of CDCC labelling. The sizes of arrows and the translocation of outer boundaries are qualitative: they do not indicate the relative strength of incentives or the expected relative change of upper and lower boundaries. Axis x: Environmental impact from worst to best. Coloured segments of Axis x: The observed range of environmental impacts of marketed products that compete within a product category. Segment of Axis x on the right-hand side of zero Impact: area of environmental restoration. Adapted from Michalopoulos (2016)

Phase 3: Label update. At phase 3 the entire scoring scale on which products are graded is updated. These updates are required to account for the entry and exit of products, as well for changes in the environmental impact of others. Similarly, to the regular update of the US Energy Guide label and to avoid the functionality limitations discussed in the case of the EU Energy label, frequent updates of the CDCC label at predefined time intervals are required to maintain its functionality to motivate continuous improvement.

The update process follows the same methodology as described at phase 1 (Label Introduction). Accordingly, the lower part of Figure 3 illustrates the scoring scale as updated in phase 3, to account for the evolution of the market incentivized during phase 2.

As indicated in Figure 3, changes in the performance of best and worst products during phase 2 move the outer boundaries of the scoring scale. Consequently, the inner boundaries between different grades also move because they are a function of outer boundaries. Therefore, any positive changes in the performance of best and worst performers that are incentivized by the label cause the entire scale to move towards the positive direction, Crucially, this can affect the grading of products that perform close-above grade boundaries. Figure 3 illustrates this effect with the help of hypothetical products A and B. Although the performance of product B did not deteriorate (it remained unchanged), product B was nevertheless downgraded because the grade boundary moved towards the positive direction. Consequently, product B lost its safe middle-range grading because it failed to improve during phase 2. After update, it ranks among the most polluting product substitutes available, and it will encounter relevant effects as mentioned. Product A is downgraded from the green category to the average one, for the same reason. Therefore, its appeal to green consumers diminishes. Notably, **the case of product A illustrates the rational incentive for the pioneers** of positive environmental performance to improve above top grade: By improving further they raise the environmental friendliness threshold, therefore forcing competitors that fail to sufficiently improve out of their target market. Accordingly,

the cases of products A and B illustrate that the incentives provided by CDCC labelling are not limited to best and worst performers: The incentives provided permeate the entire range of labelled products. Through this mechanism the proposed CDCC labelling system **incentivizes the positive translocation of the entire range** of environmental impacts observed in the market.

Beyond phase 3: Successive repetitions. Phase 3 is followed by successive repetitions of phases 2 and 3. These repetitions proceed as described above. Each of these repetitions translocate the entire range of available products towards the environmentally positive direction. Taken together, the successive translocations achieved in each repetition amount to a continuous race for environmental improvement. Importantly, this race is also open-ended towards the positive direction because it does not have a predefined upper limit. The consumer based CDCC labelling system does not target a predefined sustainability goal (e.g. zero impact). Therefore, as driven by consumer demand and enabled by environmental innovation, the market race for environmental improvement that is incentivized by the CDCC label is free to proceed beyond zero impact and into the territory of positive impact. Namely, into the territory of environmental restoration.

System stabilization endpoint

The system is expected to stabilize when supply meets demand, at which point the marginal utility of further improvement will become zero. The parameters that affect this endpoint include the cost of innovation and the satisfaction consumers derive from further environmental improvement. Innovation costs can be positively affected by the economies of scale that will be created by generalizing the race for environmental improvement across the supply sector. Consumer satisfaction from further environmental improvement will reduce as the state of the environment improves and will become minimal when that state becomes satisfactory in public perception. Given the presently negatively perceived state of the environment, reaching consumer satisfaction requires positive impact (restoration) and not merely zero impact (preservation). When satisfactory environmental improvement is achieved, consumer concern subsides, and the label effectively loses its incentivizing power. Until then, the incentive on products to further improve remains continuous and open-ended.

An outline of the main implementation steps of the CDCC label, as described above, is provided at Box 1.

Phase 1 Introduction

1. Decide the categories of substitutable products
2. Assess performances of substitutable products within a category for the indicator concerned based on research, economically rational assumptions, available data, submitted data and national or sector averages → The scoring range is determined by best and worst performers
3. Decide on the number of grades → Grade borders are calculated by dividing the entire scoring range in a number of even parts that equals the number of grades
4. Score the grade of products on the graded scoring scale
5. Introduce FoP sticker labels

Phase 2 Evolution

6. Only update product grades on the basis of new evidence as needed. Accept information submitted by stakeholders (suppliers, competitors, civil society actors, certification schemes, etc). The threshold values that define the borders between product grades are not updated.

Phase 3 Update

7. Calculate the updated scoring scale as described at phase 1.
8. Score the grade of products on the updated scale.
9. Update product labels as needed

After phase 3

10. Successive repetitions of phases 2 and 3 at sufficiently frequent time intervals

Endpoint

11. Satisfaction with the state of the environment

Box 1. Outline of the implementation steps of the CDCC label

To achieve the above-described functionality of the CDCC labelling approach in real life applications, the fine-tuning of several technical implementation aspects is required to maximize label functionality.

These include:

The number of grades. The optimal number of grades must balance the minimization of cognitive costs for product comparisons (which favors a low number of grades) against the need to lower the costs of upgrading one's production to a higher environmental grade, so as to support label functionality for continuous improvement (which favors a high number of grades). The number of grades should facilitate the movement of products between grades without compromising the simplicity of the label. The smaller the number of grades, the larger their difference in performance. Therefore, a smaller number of grades requires larger investments from suppliers to upgrade their products.[57] However, an excessively large number of grades increases label complexity and potentially reduces consumer ability to understand and compare product scores at-a-glance.

The colour of grades. A TLS-based color-coded is the dominant candidate because it is normatively charged, intuitive, and also consumers have been already trained to recognize it. However, in the light of the severity of environmental issues and urgency to bring substantial improvement, further enhancements of the normativity conveyed by the color-coded can be welcome. In this respect, an interesting variation can be the introduction of a lowest-level Black grade,[58] which, as discussed, will be populated mainly by environmental dumpers, i.e. by products that perform below the legal environmental requirements for domestic production.

Scoring scale calculation method. To simply define the scoring scale as the distance between the best and worst performing products observed, makes the pace of the continuous improvement race that is incentivised by the CDCC labelling system (i.e. the translocation of the scoring scale) to be heavily dependent on performance improvements by worst performers.[59] The disclosure of poor environmental performances generally provides strong incentives for improvement (e.g. lower profit margin, smaller market share, lower employee morale, lower quality employees, ethical divesting, reputation damage and consumer boycotts to parent companies). Nevertheless, it is possible that certain marketed products will target the environmentally unconcerned consumer segment. These products, especially when not linked to major parent companies, might have little reason to improve above the legal threshold that is applicable to the location of their production. In addition, suppliers might be incentivized to tactically preserve in the market certain low-performing products for the purpose to minimize the translocation of the scoring scale, in order to protect favourable grades of other products. Countermeasures against such possible effects include to publicly disclose the parent companies of worst performers, to maximize the normativity of the label's color-coded (e.g. black products), to take into account only the performances of viable products during scoring scale updates, as well as to to employ an alternative method for the calculation of scoring scale updates that captures the positive translocation of the overall market without being vulnerable to worst performer tactics. Affirmative action by socially responsible retailers inclined to exclude from their shops the environmentally worst (black) products can also be effective in reducing the viability of such products. This non-exclusive list indicates the rich range of countermeasures available to restore the full pace of improvement incentivized by the CDCC labelling system if required.

Depending on their feasibility, the measures available to further boost the functionality of the CDCC

[57] see Tlusty 2012 (as previously cited) for the "zone of pull" of certification thresholds.

[58] See e.g. Vanclay et al. (2010)

Vanclay, Jerome K., John Shortiss, Scott Aulsebrook, Angus M. Gillespie, Ben C. Howell, Rhoda Johanni, Michael J. Maher, Kelly M. Mitchell, Mark D. Stewart, and Jim Yates. 2010. "Customer Response to Carbon Labelling of Groceries." Journal of Consumer Policy 34 (1): 153–60. doi:10.1007/s10603-010-9140-7.

[59] The worst performers cannot reverse the positive direction of the incentivized race because that would require worsening their environmental performance. However, the performance of worst performers has a lower threshold defined by legal requirements at the location of their production. Nevertheless, the worst performers can minimize the pace of the incentivized race for improvement, by not improving their product performance.

labelling approach also includes targeted complementary regulatory policies. These can include inspections, taxation and qualified market access measures specifically targeted to lowest-grade products, as well as benefits provided to best performers.

This latter point also indicates that the proposed CDCC labelling system should be perceived as part of a 'smart mix' of measures meant to boost the pace of environmental improvement of the supply sector. The purpose of the proposed information system is not to replace but to complement state regulation. Specifically, its purpose is to use normal market forces to incentivize the improvement of the levels of environmental externalities of production, above and while relying on the safety limits set by environmental legislation.

Label format

Figure 4 presents two non-stylized standalone variations that cover the basic functionality requirements of the CDCC label concept. In this example, 5-grade scoring scales are used. The left-hand side of Figure 4 depicts a single-issue variation devoted to Climate Change (emission of GHG). In contrast, the right-hand side of the figure depicts a multi-issue label variation, which focuses on three main environmental impact categories. In line with the crucial importance of mitigating climate change, and as discussed above,[60] the impact category that concerns climate change remains non-aggregated. On the contrary, and in order to accommodate the request for aggregate indices in public policy, the remaining two environmental issues depicted on the label refer to aggregated impact categories. However, an overall aggregation of those issues is not attempted. A lower level of aggregation is preferred as more meaningful (impact on the destruction of natural habitats, and the production of plastic and toxic waste).

Figure 4. Non-stylized standalone 5-grade variations of the basic CDCC label concept. Left: single issue. Right: multi-issue.

Useful observations on these minimalistic CDCC label examples include:

(a) the single-issue variation (left) illustrates the minimal information to be communicated for the CDCC label to be functional. This required amount of information is small and composed of three basic elements: the colour of the grade achieved, the title of the impact category, and the color-coded scale used. Therefore, its space requirement on product packaging is limited. This is true not only for standalone applications as the one presented in this example, but also for applications that include the CDCC labelling system into a host label (e.g. by adding an extra row for climate change an a prospect Nutri-Score label)

(b) the multi-issue variation exemplifies how multiple issues can be covered by the same label.

[60] see discussion on the aggregation of different environmental impact categories at section "The Eco-Score label"

3.3.4. A comparison of Eco-Score vs. CDCC label functionality

Based on the above, Table 2 presents a comparative functionality assessment of the Eco-Score and the CDCC labelling approaches. The presented comparison focuses on their ability to incentivize a market-based virtuous cycle for environmental indicators, and particularly for manmade GHG emissions that cause climate change. To level the field of the comparison, a compulsory version of the Eco-Score label that differentiates product performance within categories is assumed.

Label type Functionality	Eco-Score *(Compulsory version with product differentiation within categories is assumed)*	CDCC
Market-based	no	yes
Continuous incentive for improvement across entire range of substitutable products	no	yes
Incentivizes improvement above top grade	no	yes
Explicit focus on climate change	no	yes
Incentivizes virtuous cycle	**no**	**yes**

Table 2. Comparison of selected Eco-Score and CDCC functionality features

As summarized at Table 2, the functionalities of the Eco-Score and CDCC labelling systems differ in important respects. Contrary to the Eco-Score approach,

(a) the scoring scale of the proposed CDCC labelling system and also its update process are entirely market-based, which provides objectivity and legitimacy to product grading and to its update process. Its legitimacy relies on objective market data and does not require stakeholder consent.

(b) the incentive for improvement that is provided by the proposed CDCC labelling system to competing products is continuous and covers the entire range of substitutable products. Products are not only incentivized to improve merely in order to achieve their preferred label grade. They are also incentivized to improve continuously in order to remain in their preferred label grade.

(c) the proposed CDCC labelling system provides best-performing products with incentive to innovate and improve further even after reaching top grade. The rational (economic) incentive for further innovation provided to pioneers is continuous and open-ended.

(d) the proposed CDCC labelling system provides explicit focus to the crucial issue of climate change, by recommending the non-aggregation of the particular impact category into a composite environmental score.

Conclusively, (e), the proposed CDCC labelling system provides the much-sought functionality that it incentivizes a race to the top, or virtuous cycle, for the environmental issues at stake. This functionality is not provided by the Eco-Score approach as presently applied, even if a compulsory version that differentiates product performances also within categories is assumed.

3.4. Discussion

The proposed CDCC labelling system synthesizes good practices from a range of presently available labelling approaches. Its most relevant features and functionalities are summarized at Table 3.

	Functionality	**Feature**
1	It incentivizes a distribution of environmental practices among suppliers (supply) that matches the distribution of environmental preferences among consumers (demand)	it is multi-grade and compulsory; therefore it removes information asymmetry on the comparative environmental performance of all substitutable products
2	It minimizes cognitive costs to consumers for product comparisons	it is simple, quantitative, categorical, comparative and colour coded, therefore supports at-a-glance product comparisons
3	It engages moderately concerned consumers to ethical consumerism	it is multi-grade, compulsory and normatively color-coded, therefore worst performers are identified and can be avoided
4	It minimizes the space available for misleading greenwashing practices by poor performers (by all performers) to the cost of superior ones	it is comparative and compulsory, therefore worst performers (all performers) are identified
5	It minimizes the space available for stalling practices by poor-performing stakeholders during label updates	it is objectively updated, and its legitimacy is market-based
6	It minimizes the periods of reduced label functionality because of lagging behind market evolution	frequent updates are recommended
7	It protects the market from bias against environmentally positive technological innovation	data-based differentiation of product substitutes within product categories is recommended
8	It protects label functionality and credibility from the misleading inclusion of void lower grades	it is comparative; therefore, it rules out void lower grades. Frequent updates are recommended
9	It provides incentive for continuous improvement that permeates the entire range of category products	its scoring scale is dynamic; therefore, products are incentivized to improve not only for achieving their preferred label grade, but also for remaining in their preferred label grade
10	It is inherently inclined by design towards the environmentally positive translocation of the entire range of marketed product substitutes	the upper environmental performance threshold used is dynamic and set by the evolving frontiers of technological innovation, while the lower performance threshold is rigid and set in law

	Functionality	Feature
11	It provides incentive for improvement above top grade	by improving further, the best performing pioneers effectively downgrade their competitors that fail to also improve (i.e. innovators force competitors out of their targeted market)
12	It provides explicit focus and maximal functionality towards combating climate change	the non-aggregation of the GHG emissions impact category is recommended
13	It incentivizes competition for environmental improvement that is open-ended and extends to positive impacts	the scoring scale is defined by best and worst performers and no upper performance threshold is anticipated
14	It engages the supply sector in a consumer-driven and market-based race for continuous environmental improvement, wherein suppliers ultimately compete not merely for minimizing additional destruction of the environment, but in terms of environmental restoration.	as a result of all the above.

Table 3. Selected features and functionalities of the CDCC labelling system

As described in previous sections and summarized at Table 3, the proposed CDCC labelling system is a combination of features already met in available labelling approaches. This said, the novelty of the CDCC labelling system rests in the selection of features it includes, and also in the way those features are combined.

The disruptive character of the CDCC labelling system

The particular synthesis of features that constitute the CDCC labelling system, is designed to support the working principle of a continuous and open-ended race for environmental improvement among the supply sectors. This is a crucial functionality that is much-sought in our effort to mitigate catastrophic climate change: It engages the supply sector in a consumer-driven and market-based virtuous cycle for environmental improvement. Moreover, the virtuous cycle incentivized is not limited to minimizing additional destruction of the environment. As much needed, it expands into the territory of environmental restoration.

The proposed CDCC labelling approach is designed to use the market as an instrument against its own bad self. That is, to initiate a continuous competition race in which normal market forces, 'the law of supply and demand', **incentivize businesses to compete freely and transparently for environmental reputation** in the market. Towards this goal, it makes consumers aware of the comparative environmental impact of their consumption, so that the purchase of environmentally inferior products will need to happen knowingly, subject also to peer pressure and in light of the climate change effects observed. Synergies in this process can be expected from related markets, such as the labour and finance markets. These can contribute effects that include reduced ability to recruit highly qualified employees, lower employee morale, and disinvestment by ethical funds, which can enhance the motivation of the worst polluters to improve.

Essentially, the proposed CDCC labelling approach is designed to cure information asymmetry regarding the environmental performance of products. Moreover, the comparative approach adopted is designed to effectuate that the market optimization of unobservable environmental impact attributes happens similarly to that of observable product attributes that also have a fixed normative

direction (such as the processing speed of processors in electronic devices). In that case, suppliers are engaged in a continuous race for improvement, where novel improvements by competitors effectively downgrade in the eyes of consumers the performance of products that until that moment were regarded as satisfactory. **This inherently 'free-market' dynamics is spontaneously at work in markets for observable attributes: *Products and firms must continuously innovate and improve, or else they will become obsolete and will be forced to exit the market.*** Although this dynamic is business as usual in markets for observables, it remains alien in markets for ethical attributes that are ruled by the dominant labelling paradigm. **The proposed CDCC labelling approach disrupts the established paradigm of environmental labelling by supporting the same market optimization process for environmental attributes.**

Falling short of satisfying a longstanding economics aspiration, the described approach does not incorporate the costs of negative production externalities into product prices. Instead, environmental externalities affect the reputation of products and suppliers with the aforementioned effects. Nevertheless, the above-described approach to the disclosure of environmental performance addresses a number of information asymmetry effects. 'Moral hazard' by businesses that exercise environmental 'free-riding' is addressed through enabling their identification by consumers, employees, and investors, which trigger effects that incentivize their improvement. Similarly, transparency on the environmental performance of products addresses the 'adverse selection' of environmentally inferior firms to the cost of superior ones based on misleading brand promotion campaigns. Finally, the objectivity of the anticipated market-based approach to label updates addresses 'regulatory capture' effects of information asymmetry: Supply stakeholders cannot extract 'rents' by (co-)determining the thresholds of grades so that their products are assessed more favourably than they would if environmental attributes were observable.

This report is composed in the light of the dire state of multiple environmental indicators and while the press release to the Sixth Assessment Report by the IPCC (9/8/2021) proclaims "Climate change widespread, rapid, and intensifying".[61]

> *From a physical science perspective, limiting human-induced global warming to a specific level requires limiting cumulative CO2 emissions, reaching at least net zero CO2 emissions, along with strong reductions in other greenhouse gas emissions. Strong, rapid and sustained reductions in CH4 emissions would also limit the warming effect resulting from declining aerosol pollution and would improve air quality.*
>
> IPCC Sixth Assessment Report (9/8/21)
>
> https://www.ipcc.ch/report/ar6/wg1, SPM-p.36

In this light, the historic goal of a welcomed EU-wide environmental labelling initiative should be nothing less than to at last unleash the full force of market forces towards rapid, systemic, tangible, widespread and sufficient environmental improvement.

The disruptive functionality offered by the proposed CDCC labelling approach makes it a powerful tool for achieving this goal, as dictated by science, proclaimed in international and global initiatives, and as owed to life on earth as we know it, including the future generations of humanity.

[61] "Climate change widespread, rapid, and intensifying – IPCC", The Intergovernmental Panel on Climate Change, Press Release 9 August 2021
https://www.ipcc.ch/site/assets/uploads/2021/08/IPCC_WGI-AR6-Press-Release_en.pdf,
The full IPCC Sixth Assessment Report is available at: https://www.ipcc.ch/report/ar6/wg1/

4. Main Findings – Recommendations

4.1. Main Findings

A significant practical difference that distinguishes the labelling approach recommended in this report for environmental impact (especially GHG emissions) from that recommended for nutritional content is inherent in the nature of these attributes: The ideal content in most food ingredients is often a matter of optimal dose and also varies with the physiology of the consumer. In contrast, the GHG emissions attribute exhibits a clear normative direction of the "less is better" type: it is always good to emit less GHG. It is better to emit none, and it is even best to be most negative in emissions, i.e. to absorb as much GHC as possible (e.g. through carbon offsetting). Furthermore, it is by now scientifically established and overwhelmingly accepted that the simultaneous deterioration of multiple environmental parameters poses a severe threat, with overarching catastrophic effects of human-induced climate change foreseen at all levels (social, economic, and also on the biodiversity and on broader ecosystems).

These parameters, namely the clear definition of the problem and the clear normative direction of the change needed mark a stark difference between the two issues analysed in this report. In a sense, when dealing with the environmental issue we have the more fortunate situation that the enemy, so to speak, namely greenhouse gasses - GHG, is clear and what needs to be done is beyond doubt as well: to reduce their concentration in the atmosphere as much as possible. This is particularly important in connection to food production, which amounts to about a quarter of total manmade GHG emissions globally. The problem, however, is not as simple when it comes to how to achieve this reduction in the present political and economic context. Despite decades of warnings, by now we have already missed the opportunity to avoid catastrophic effects of climate change and the time available to mitigate them also runs out.[62]

In this regard, the prospect of an EU-wide Front-of-Pack (FoP) environmental labelling initiative, even if it applies only to food products, is certainly a move in the right direction. Product labelling is a tool long used in efforts to remove information asymmetry that prevents ethical consumerism from optimizing the levels of unobservable (credence) product attributes in the market. Such labelling initiatives succeeded to bring some improvement through product certification and also through positive spill over effects to the broader production sector.

However, decades of experience shows that the progress achieved is insufficient as compared (a) to the magnitude of the environmental challenges and of the sheer scale of changes needed to address them, and (b) to the potential for ethical consumerism that remains available as indicated by consistent research results showing increased environmental concern among the international public. In this context, the present report begins with a critical assessment of the functionality of established label types that are typically used to facilitate ethical consumerism. This critical assessment advances the view that the established product labelling paradigm suffers from functionality limitations that do not allow ethical consumerism to develop into a mainstream market force of a magnitude sufficient to incentivize the much needed, substantial, and widespread, improvements in the environmental performance of the supply sector, causing the market for environmentally friendly products to fail. These functionality limitations essentially cancel from the outset any hope that a conventional EU-wide nutrition and environmental label will be the much-needed game-changer that will radically transform the environmental impact of food production.

Front -of -pack (FoP) Nutrition Labelling Schemes

The main conclusions from the analysis of the various FOP labels systems are:

---

[62] "Climate change widespread, rapid, and intensifying – IPCC", The Intergovernmental Panel on Climate Change, Press Release 9 August 2021
https://www.ipcc.ch/site/assets/uploads/2021/08/IPCC_WGI-AR6-Press-Release_en.pdf,
The full IPCC Sixth Assessment Report is available at: https://www.ipcc.ch/report/ar6/wg1/

Regarding **reference intake schemes**, we notice that several attempts have been made (most outside the EU) and all of them remain active until now. These are labels that are relatively difficult for consumers who do not have the relevant knowledge to understand. Consumers need to be "educated" in order to be able to evaluate which food they will choose.

However, these FOP labels do not involve the "subjective judgment" of any algorithm, so they are fair labelling systems as they list food ingredients in detail and allow consumers based on the knowledge and information, they have to choose both which food to consume and in what quantity. Most have the portion as reference (which is a more reliable than 100g) and are also created according to the rules of a Directive.

Colour-coded nutrient-based schemes are more understandable to consumers because of the labelling of ingredients in three colours (red, orange, green). An evaluation of each ingredient is applied separately, and this is useful for the consumer, since for health reasons he may want to avoid consuming a large amount of a particular ingredient. It is also positive that some of them have the portion as a reference.

Studies have shown that the fact that there are three colours does not usually encourage consumers to increase their consumption of "green" coloured food but leads them to avoid products with "red" and "orange" colours.

Overall rating schemes are currently the most popular FOP labels in the countries that apply them. Their main advantage is that they are very easily understood by the consumer, since they give him a unique information, that of the "evaluation" of the food and in also coloured with characteristic traffic light colours. This way of labelling helps a very large number of consumers who are not "trained" to decide on the foods they will consume based on the FOP label.

However, there are two main disadvantages to this type of labelling. The first is a reference to the amount of "100g" for the evaluation of the individual ingredients and then the food as a whole. This is not a "fair" system for foods that are consumed in quantities much less than 100g and contain ingredients that are beneficial to health when consumed in such small quantities. The second, is the fact that the consumer is informed only about the final result of the food's evaluation, without being informed about the rating of the individual ingredients. It is also difficult for a consumer to be informed and understand the algorithm with which the final food is evaluated.

In this way it is possible to challenge the rating system. Consumers are also not "educated" in information and critical thinking about the foods they consume and their eating behaviour in general.

Endorsement schemes ('positive logos') are quite developed in number both inside and outside the EU. They are obviously very popular because they do not require "education" of the consumer and are also based on the evaluation of food by the relevant algorithm.

The specific types of FOP label are not given a "negative" rating, but the food is "rewarded" with the positive logo as long as it meets the nutritional value requirements according to the algorithm of each system.

Ecolabel Schemes for Food Products

The present report proposes for consideration an innovative labelling system as a synthesis of functional features of available labelling approaches. The proposed labelling system is Compulsory Dynamic Categorical and Color-coded (CDCC). Contrary to the established eco-labelling paradigm, the information provided by the proposed labelling system does not merely aim to raise a part of production to a higher environmental standard, which in the best cases is sustainable. Instead, the CDCC label is designed to serve the more ambitious labelling goal to establish for environmental attributes the full effect that normal market forces have on the evolution of observable (non-credence) attributes that belong to the more-is-better type (similarly as the less-is-better type, the more-is-better attribute type has a clear normative direction). Namely, to engage the market in a consumer-driven and open-ended race for continuous improvement. To exemplify, consider the evolution of the processor capacity attribute in electronic devices.

Similarly, to other attributes of the same type, the implicit consumer acceptability threshold for processor capacity is dynamic and market defined. It is set by the evolution of processor speed among marketed products, and it is continuously raised by suppliers that compete in investing, inventing, and incorporating innovative technology to improve the speed of their products. In this continuous race for improvement, novel improvements by competitors effectively downgrade in the eyes of consumers the performance of products that until that moment were regarded as satisfactory. This inherently 'free-market' dynamics is spontaneously at work in markets for observable attributes: Products and firms must continuously innovate and improve, or else they will become obsolete and will exit the market. Although this dynamic is business as usual in markets for observables, it sounds alien in markets for environmental attributes that are ruled by the dominant eco-labelling paradigm.

The proposed CDCC labelling approach disrupts the established environmental labelling paradigm by incentivizing for environmental attributes the same market optimization dynamic that applies to observable ones. Simultaneously, it disrupts the established paradigm of negative environmental externalities in the supply sector: It is designed to engage product substitutes into a continuous, market-based, and consumer-driven race for environmental improvement. The incentivized race is open-ended towards the positive direction. As driven by consumer demand and enabled by environmental innovation, it is free to proceed beyond zero impact and into the territory of positive impact. Namely, into the territory of environmental restoration (e.g. carbon offset).

The proposed CDCC labelling approach is designed to use the market as an instrument against its own bad self. **That is, to initiate a continuous competition race in which normal market forces, 'the law of supply and demand', incentivize businesses to compete freely and transparently for environmental reputation in the market**. Towards this goal, it makes consumers aware of the comparative environmental impact of their consumption, so that the purchase of environmentally inferior products will need to happen knowingly, subject also to peer pressure and in light of the climate change effects observed.

Synergies in this process can be expected from related markets, such as the labour and finance markets. These can contribute effects that include reduced ability to recruit highly qualified employees, lower employee morale, and disinvestment by ethical funds, which can enhance the motivation of the worst polluters to improve.

In the light of widespread, rapid, and intensifying climate change (IPCC 9/08/2021), this report proposes that the historic goal of a welcomed EU-wide environmental labelling initiative should be nothing less than to unleash the full force of market forces to incentivize rapid, systemic, tangible, widespread and sufficient environmental improvement.

The disruptive functionality offered by the proposed CDCC labelling approach makes it a powerful tool for achieving this goal, as dictated by science, proclaimed in international and global initiatives, and as owed to life on earth as we know it, including the future generations of humanity.

4.2. Recommendations

4.2.1. Front -of -pack (FoP) Nutrition Labelling Schemes

There are no "good" and "bad" foods. All foods provide nutrients necessary to the human body. What is often "wrong" is the amount of food consumed. This is because the body is provided with large or excessive amounts of some ingredients with harmful effects on health. It would be good in foods that have a high caloric value or large amounts of specific ingredients to indicate with "emphasis" on the package maximum amount of consumption per meal or per day.

The proposed FOP label system and the methodology of "evaluation" of the food that it will provide should be based on international standards and regulations.

The system should be "appreciable" and use colour gradation, so that it is understandable and "attractive" to as many consumers as possible. It should be noted here that it is very important to "educate" consumers about nutrition so that they can understand even more complex labelling systems and be able to become the same "evaluators" of the food they are going to choose on the shelf. instead of blindly "trusting" the algorithm of an evaluation system.

The system should be "fair" towards food, using as a reference the portion (and not 100gr or 100ml) but also the percentage of supply of individual ingredients in relation to the maximum daily allowable amount.

4.2.2. Ecolabel Schemes for Food Products

Environmental labelling systems should be mandatory. If labelling is optional, only the best products will be labelled and therefore the consumer about will not be informed on the less good environmental products. Thus, if labelling is mandatory the "worst" polluters will be motivated to improve their environmental performance and it will also minimise the possibility of misleading consumers through greenwashing.

Environmental labelling systems should use color-coded multilevel Front-of-Pack marking, which is available in the form of a security sticker (sticker label). The security sticker form (i.e. with notches) facilitates frequent updates without changing the packaging, therefore the visibility is maximised and quick product comparisons without much effort from consumers is facilitated.

Environmental labelling systems should be "Unconcentrated" (at least in terms of greenhouse gas emissions). An overall environmental assessment (score) which covers all the environmental impact categories of each product should not pre presented. It is suggested the product score for the category Climate Change (Greenhouse Gas Emissions - GHG) to be presented separately from the other environmental impacts, in order to maximize the effectiveness of the labelling for this criterion.

Environmental labelling systems should be "Dynamic" and "Comparative" of the available products in each category. The final product rating should not be absolute (as in Eco score) but should correspond to / depend on existing alternatives: in this regard, the least environmentally friendly product in each category should receive the lowest rating (grade) by definition, while the best receives the highest (the other products are distributed intermediate). This maximizes the incentive to improve both the worst and the best products:

(1) Avoid the phenomenon that all available products have e.g. score A (see EU Energy label)

(2) Incentives are created for the best performers to invest in innovation in order to improve even more, moving the whole rating scale to the best score, and thus forcing other products to improve (virtuous circle)

(3) The periodic update of the product evaluation scale and product ratings can be done automatically, impartially, and objectively, based on the evolution of the performance of the best and worst products.

Environmental labelling systems should be updated at regular (frequent) intervals. To avoid the phenomenon observed in the case of the EU Energy label, where scale updates occur every ten years, where as a result the rating scale quickly becomes obsolete and ineffective, updates on environmental labelling systems should be annual.

The differentiation of the environmental impact of competing products during their grading, should be based on scientific data.: Although it is logical that some assumptions may be made for reasons of lack of data and cost reduction, it is important not to see a systemic bias against technologically innovative forms of production that could provide necessary solutions (e.g. carbon capture, synthetic meat, etc.), due to the tight attachment to specific production systems (e.g. organic, biodynamic) as Ecoscore does.