Food Policy 116 (2023) 102435

Contents lists available at ScienceDirect

Food Policy

journal homepage: www.elsevier.com/locate/foodpol

Less is more: Information overload in the labelling of fish and aquaculture products

Francesco Bogliacino [a,*], Rafael Charris [b], Cristiano Codagnone [c], Frans Folkvord [d], George Gaskell [e], Camilo Gómez [f], Giovanni Liva [g], Felipe Montealegre [h]

[a] Università degli studi di Bergamo, Dipartimento di Scienze Economiche, Via dei Caniana n 2, Bergamo, Italy
[b] NYU Grossman School of Medicine, Department of Population Health, Epidemiology Division, United States
[c] Università degli studi di Milano, Italy
[d] Tillburg School of Humanities and Digital Sciences, Tilburg University, The Netherlands
[e] London School of Economics, United Kingdom
[f] CERGE-EI, Czechia
[g] Open Evidence, Italy
[h] Università di Bologna, Italy

ARTICLE INFO

JEL Classification:
C91
D12
Q56

Keywords:
Fish
Label
Information overload
Recall task
Trustworthiness

ABSTRACT

Food labels have been used extensively for informing consumers to make more rational and safer decisions. However, this carries the risk of confusing consumers with multiple claims which may distract from key information such as the country of origin of the product. To inform the European legislation, we have tested labels on fish and aquaculture products in three separate experiments, across several European Member States. The main results showed that mandatory information is better recalled than voluntary information. In addition, consumers perceive, and process differently labels for farmed and caught fish, relying more on quality claims for the former. Nonetheless, in both cases, while they value visual information, they are likely to be confused by voluntary claims including flags. Finally, when additional claims are added step by step, they lead to a decrease in accuracy of recall and comprehension. In sum, less is better, because too much information on food labels lead to cognitive overload and consumer confusion.

1. Introduction

Public policies promoting dietary changes make extensive use of product labelling as a vehicle for information provision (Caswell and Majduzka 1996; Cowburn and Stockley 2005). Specifically, Fish and Aquaculture Products (hereafter FAP) are important for a balanced and healthy diet, reducing the risk of cardiovascular diseases (FAO/WHO, 2011). But with increasing demand for FAP came greater complexity in consumers' preferences (Trondsen et al. 2004a; 2004b; Tomić, Matulić, and Jelić 2016; Thorsdottir et al. 2012; Ankamah-Yeboah et al. 2019). For example, buyers prefer to purchase fish whose origin guarantees quality, lower environmental impact, favorable market conditions for small producers in local communities, and low exposure to sources of health risk (Carlucci et al. 2015). In simple words, there is an increasing demand for more detailed information.

At the same time, the supply side is also important. Sellers may use labels to escape competitive pressure. Labels can be used as weapons to obfuscate information to buyers (Kalaycı 2016; Kalayci and Potters 2011; Crosetto and Gaudeul 2017; Gu and Wenzel 2015; 2020), falsely differentiating their products, and achieving a higher markup (Bertoletti and Etro 2016).

The simultaneous presence of different sources of labelling information generates information overload, leading to more complex decision making for consumers. Information overload occurs when the flow of information hits the boundary of processing capacity of an individual or an organization (O'Reilly and Charles, 1980). Information overload can be particularly challenging when a specific piece of information should be conveyed as a priority. This is not uncommon in FAP origin

* Corresponding author.
E-mail addresses: francesco.bogliacino@unibg.it (F. Bogliacino), rafael.charrisdominguez@nyulangone.org (R. Charris), Cristiano.codagnone@unimi.it (C. Codagnone), f.folkvord@tilburguniversity.edu (F. Folkvord), G.Gaskell@lse.ac.uk (G. Gaskell), Camilo.Gomez@cerge-ei.cz (C. Gómez), gliva@open-evidence.com (G. Liva), andres.montelagre@studio.unibo.it (F. Montealegre).

https://doi.org/10.1016/j.foodpol.2023.102435
Received 11 January 2022; Received in revised form 25 February 2023; Accepted 2 March 2023
Available online 31 March 2023

and other labels. As an example of potential confusion, consider a label simultaneously indicating the port of landing, a national flag, and a nautical catchment area, or even worse, consider a fish product that went through multiple processes in different countries, each location reported on the label, that also includes price, weight, possible quality claims, nutritional information, and various logos.

Do consumers understand, trust, and process origin-related information in FAP labels? This research question is of direct relevance to current European Union regulation which, in addition to mandatory country of origin related information, allows producers to display a wide array of voluntary claims in their product labels. As a result, consumers may be exposed to copious amounts of information that could affect their capacity to make choices that reflect their real preferences.

In this article, we report three studies that address this research question experimentally. In the experiments, conducted in multiple countries and using random samples of the population with experience of recent purchases of FAP, we develop a comprehensive testing protocol to assess labels in terms of recall, trustworthiness, and effectiveness, covering caught and farmed, processed and unprocessed FAP.

The first experiment tests incentivized comprehension and recall of origin-related information after viewing labels. A key methodological aspect is the random assignment of respondents to one of two conditions (i) labels with only mandatory information and (ii) labels with mandatory and voluntary information. Both groups are asked the same origin-related questions, establishing a clean test of information overload. The second experiment adapts a design originally aimed at eliciting normative expectations (Krupka and Weber 2013) into a task aimed at measuring shared beliefs about the trustworthiness of different claims. Respondents are incentivized to match the modal response of the trustworthiness of different claims. Participants in the experiment are allocated randomly to one of two 'time' conditions to test the role of automatic and deliberative cognitive modes in the evaluation of labels (Kahneman 2011; Sanjari, Jahn, and Boztug 2017). Finally, the third experiment tests the comprehension and effectiveness of different labels in a simulated marketplace, designed to maximise ecological validity. Respondents are instructed to complete a set of instructions, where the key information is provided via labels. In this last experiment, labels are controlled as treatments in a between subjects' design.

In brief the results are as follows. Respondents in Study I recall mandatory better than voluntary information, but country of origin claims tends to be understood and recalled less than other claims related to price, weight and other pieces of information. The study points to the risk of information overload. Providing more information in the label decreases accuracy in the response to target questions, by six percent. From Study II, we learn that consumers construe differently the labels for farmed and caught (wild) fish, relying more on quality claims for the former. In both cases, they are likely to be confused by claims including national flags. When judging trustworthiness of labels, there is no effect of time pressure, suggesting no difference between cognitive and deliberative modes of cognition. Finally, Study III documents a monotonic decrease in the accuracy with which the information is processed when additional claims are added step by step.

Our findings contribute to the literature on information overload (Eppler and Mengis 2004), on the optimal design of labels (Bogliacino et al., 2015a, 2015b; Codagnone et al. 2016; Folkvord et al. 2016, 2019, 2020), and of food labels in particular (Verbeke 2008; Verbeke and Ward 2006; Teisl 2003). Since the background of the study was motivated by a policy requirement (Codagnone et al. 2021), this paper contributes to the literature on evidence based policy, particularly behavioural studies (Shafir 2013; Bogliacino et al., 2015a, 2015b; Bogliacino et al., 2016; Codagnone et al. 2014).

This article is organized in seven Sections. After the Introduction, section 2 presents the policy background and theoretical framework and outline a research plan; sections 3–5 present the design, analysis, and results of the three studies; section 6 and 7 include discussion and conclusions. In the Supplementary Online Materials file (SOM), the reader can find the protocols of the experiments.

2. Policy background, theoretical framework, and the research plan

2.1. Policy background of the study

In the European Union, Regulation (EU) No 1169/2011 sets out the requirements for the provision of food information to the consumer and states that information on country of origin or place of provenance should not be misleading. In addition, the specific regulation applying to all FAP except processed ones is No 1379/2013. It requires, within the common market, to properly mark or label the area where the product was caught or farmed (Art. 35.1(c) of Regulation (EU) No 1379/2013, e.g. "caught in the Central Mediterranean, Adriatic Sea", "farmed in Ireland"). This claim is among those considered as "mandatory information".[1]

However, Article 39 of the regulation No 1379/2013 opens the space for "additional voluntary information", which complements but does not replace that which is mandatory. The provision of this additional information must respect other directives that govern the European Single Market, such as the Unfair Commercial Practice Directive. According to this, information should be clear and unambiguous, easily verifiable, and not misleading for the final consumer. The Directive adopts the perspective of the average consumer, who is sufficiently circumspect and pays sufficient attention. However, voluntary information may open the door to general and potentially misleading claims, for example a national flag or a statement such as 'Irish' or 'Product of Italy', offering little actual information and potentially confusing and misleading consumers.

While the regulation is clear, it could be sub-optimal in the provision of accurate information concerning the country of origin, especially when it comes to FAP labelling. But why is origin a valuable and sensitive piece of information? Psychological explanations based on the principle of decision-making driven by heuristic thinking are compatible with a single attribute acquiring a dominant role in consumer choice. Other explanations that account for this fact are the theory of planned behavior (Ajzen 1991; Conner 2001), and salience theory (Bordalo, Gennaioli, and Shleifer 2013).

In this context, country-of-origin becomes a *sufficient statistic*, reducing the degree of uncertainty and choice dimensionality. Since fish is a highly perishable product, country of origin provides valid information in terms of freshness, or the likely exposure to preservation treatments (Birch, Lawley, and Hamblin 2012; Carlucci et al., 2015; Claret et al., 2014). A country may evoke a poor record in terms of safety regulation or environmental protection, or could be associated with a remarkable tradition in terms of food and quality (Juhl and Poulsen 2000; Uchida et al. 2014; Lawley, Birch, and Hamblin 2012). Hansen and Sallis (2011) claim that the intrinsic properties of the fish, being unknown or misunderstood, are often overlooked, in favour of country of origin or other general claims, such as "product of sustainable fishery". Taken together, the findings from the above contributions may be interpreted as if consumers do not interpret fish origin as a piece of information that is processed while updating their beliefs, but rather as a heuristic used to decide on whether to buy and consume the product.

A heuristic is a mental shortcut. Heuristics in judgement and decision-making are a cornerstone of the standard behavioural economics model (Camerer 1995; Tversky and Kahneman 1974; Kahneman and Tversky 1979). The survival of heuristics as mental processes can be justified from an evolutionary point of view (Tversky and Kahneman 1974), as heuristics allow the decision-maker to save on cognitive resources, even though they may introduce cognitive biases and

[1] In the companion report, on which this section is partly built (Codagnone et al. 2021), there are additional details on the legal framework.

systematic errors. For example, in this context, countries of production that are considered as trustworthy, such as Norway or the Netherlands, can create a biased estimation of the quality of the product, while no factual information is provided on quality. Vanhonacker et al. (2011) argue that the combined problem of limited awareness and over-inference out of specific claims is pervasive among European consumers.

This possibility of mistakes is furthered by the intrinsic complexity of the origin-related information, and of the possibility of information overload in labels (Verbeke 2008; Sørensen et al. 2012). In the case of wild fish, the product's provenance is determined only by the catch area and the vessel's flag of a given country. However, the product may undergo further processing (filleting, canning, breading), which adds another origin claim on the label. Regarding the issue of information overload, some of the claims in the label often associate the quality or taste of the product to the origin, without further clarifying whether origin refers to the vessel's country or the processing site.

The fact that consumers value country-of-origin claims, information that is only intuitively processed and not deliberatively appraised, generates an opportunity to extract rents, which profits-driven companies could exploit. As a result, sellers can follow obfuscation strategies, trying to profit on the cognitive mistakes of consumers (Kalaycı 2016; Kalayci and Potters 2011; Crosetto and Gaudeul 2017; Gu and Wenzel 2015; 2020; Bogliacino et al. 2022). According to the Eurobarometer, 25 percent of respondents find that product labelling information is neither clear nor easy to understand. The confusion linked to the ambiguity of information may lead to misinterpretation, and to a potential adverse effect of a decreased level of trust. The functioning of the European Single Market requires a fair competition for the businesses that play by the rules and a significant amount of trust in information provision.

Given the limitations of the current labelling and the ongoing revision of the policy framework, testing the extent of information overload is paramount, to assess what information is better able to generate trustworthiness, and, to identify which claims determines more favourable outcomes in terms of comprehension and capacity to process the information.

2.2. Theoretical background

In the spirit of evidence based policy, the current contribution is trying to address a policy question. The goal of the legislation is to provide correct and effective information. Correct means non deceptive and effective means that it is likely to cause a change in behaviour, allowing consumer to choose according to their preferences. Reaching the second goal requires understanding how the consumers choose in this domain.

Choosing between two food options is partly similar and partly different from choosing between two pairs of shoes or two investment opportunities. In fact, albeit there exists a common process of value-based computation of different attributes of the options, there exists also a homeostatic process of feeding regulation which is specific to food choices (Rangel, 2013). These behavioural controllers compete and may conflict over which option to choose. This neuroscientific perspective suggests that simple dual process models might fail to grasp the complexity of these decisions.

The general philosophy of dual process models is that cognitive determinants of human behaviour are overestimated and providing more information does not necessarily lead to better decision-making outcomes (Jeffery, 2004). Dual-process models (i.e., system 1 versus system 2 as popularised in Daniel Kahneman's 'Thinking Fast and Slow', 2011) assume that an automatic and a reflexive system coexist and operate as the mechanism between information provision and behaviour. In this specific setting, information is food information, and behaviour is consumption behaviour. The two most widely adopted models are the elaboration likelihood model (ELM) (Petty & Cacioppo, 1996; Petty et al., 2005) and the heuristic systematic model (HSM) (Chaiken & Trope, 1999; Eagly & Chaiken, 1993). At the core of both models lies the assumption that under some conditions people carefully process a persuasive message (the systematic process, i.e., they use system 2) and at other times use low-effort mechanisms to respond to a message (the heuristic driven process typical of system 1) (Buijzen, Van Reijmersdal & Owen, 2010).

While system 1 and 2 may capture some of the dynamics induced by the Pavlovian controller and the goal directed controller (Rangel, 2013), they miss the presence of habits and learned association between outcomes and choices. In social psychology, models of this type posit the existence of a third process (the automatic or experiential process) which is characterised by automatic, unconscious reactions to process food information cues (Buijzen, Van Reijmersdal & Owen, 2010). This is a further and more reflexive mechanism than the one put forth by dual process models for system 1, but it is not forward-looking like system 2. A theoretical model that has been developed recently and serves as a guiding framework for understanding the relationship between consumers' responses to food information and subsequent buying intentions is the Reactivity to Embedded Food Cues in Advertising Model (REFCAM) by Folkvord et al (2016). According to the REFCAM, food cues trigger physiological and psychological responses to food (*cue reactivity*), that lead to a change in attitude and perception of the foods (*advertising effect process*), and subsequently causes a reciprocal relationship with consumption behaviour (*incentive-sensitisation process*) (Folkvord et al., 2016). The REFCAM has been assessed and validated in several experimental studies (Folkvord et al, 2016, 2017; Masterson et al., 2018; Spielvogel, et al, 2018). According to Folkvord et al (2016), exposure to cues on food packaging leads to the formation of a certain attitude towards the product (i.e. the *habit*) that is needed to guarantee a certain quality. However, the concept of package design has a multi-dimensional nature, and incorporates many different elements such as shape, logo design, size, colours, illustrations, material, and nutrition information (Underwood, Klein & Burke, 2001).

What are the implications for this study? Different external cues matter for food choice. These cues may activate different circuits and controllers. Stimuli that require the activation of the goal directed controller should compete for attention (Enax et al. 2015; Hare et al., 2011, Hare et al., 2009). Labels may increase the integration of certain attributes in the computation of the value attached to each choice (Rramani et al. 2020; Enax et al. 2016).

Since cues compete for attention, the research shows that the most important are visual cues (with a focus on colour) and textual cues (with a focus on advertising claims) (Ampuero & Vila, 2006; Karnal et al., 2016). Visual package design attributes, such as packaging size, colour usage, shape angularity, and logo design are amongst the packaging features that have been shown to impact product evaluations and subsequent food experiences (van Rompay, Deterink & Fenko, 2016). Colour is often a key feature for food and beverage companies that want to introduce new, healthy-looking quality products (van Rompay, Deterink & Fenko, 2016). Consumers associate warmer colours with tastiness (high arousal), a hedonic attribute, while cool colours may be more readily perceived as natural and healthy (low arousal) (van Rompay, Deterink & Fenko, 2016).

Summarizing this first part of the theoretical discussion, the circuitry of food choice signals the presence of a scarce resource, consumers' attention, on which there is a contest among different stimuli. This informs our design because we aim to assess whether the claim of interest for the policymaker maintains saliency in presence of information overload.

We now move to the second aim, the provision of truthful and transparent information. Although the law clearly requires the information to be true and verifiable, the labelling is in the hands of the companies and the consumer may be unable to distinguish mandatory claims from other voluntary claims (or nutritional ones). This raises an issue of trust, in particular as companies have been changing their labelling policies.

Indeed, the use of textual claims in food products has changed over the years. In the early days of advertising, food manufacturers did not actively promote nutrition and health claims to promote the quality of products. They were concerned of a possible backlash from linking food consumption to health and nutrition benefits without the support of clear scientific evidence (Kim et al., 2009; Klassen et al., 1991). Instead, advertisers sold products by placing a focus on the price, quality, taste, scent, and colour of the product. However, due to a growing interest in the health issues among consumers, companies began promoting their products as nutritious and part of an active lifestyle to boost sales (Fay, 2003).

Health and nutrition benefits have now become a major factor in consumers' food choices (Kim et al., 2009). The rapid increase in the amount of information that advertisers communicate through food packaging and advertisements represents both an advantage and a disadvantage for consumers (Gravel et al, 2012). On the one hand, consumers receive more information about the products they buy. On the other hand, claims could affect consumers' perception of quality and mislead them (van Rompay, Deterink & Fenko, 2016; Silayoi & Speece, 2004). Karnal et al (2016), for example, showed that a product labelled with a health claim was perceived as healthier than a package without such a claim, especially when the claim was presented as the only health information on the packaging. This leads to consumers overgeneralising the quality effect of claims, thinking the product to be generally healthy, although the claim may only mention one specific effect (Garretson & Burton, 2000; Roe, Levy & Derby, 1999; Williams, 2005).

According to Kim et al (2009), advertising claims can be divided into two categories: 1) product information claims, that provide basic information about the taste, quality, or novelty of the product, and 2) nutrition/health claims. More specifically, a nutrition claim is any claim which states, suggests, or implies that a food has specific beneficial nutritional properties due to its energy contribution or nutrient content (European Commission, 2006). This includes claims such as 'no sugar added' and 'fat free'. A health claim is any statement about a relationship between food and health that is based on scientific evidence (European Commission, 2006). This includes claims like 'this food lowers cholesterol' or 'vitamin D is good for your bones'. However, Williams (2005) concluded that consumers are not able to make a clear distinction between nutritional and health claims. The underlying idea of conveying food quality through product packaging features is that consumers will make purchase decisions based on that information, and perhaps even change their buying habits (van Rompay, Deterink & Fenko, 2016). This reciprocal relation between attitude towards products with health claims and behaviour is identified by Folkvord et al (2016) as the incentive sensitisation process, meaning that the attitude towards and perception of certain products becomes a strong predictor of subsequent behaviour over repeated exposures, and that an increased food intake in turn reinforces this attitudinal response. The central process is characterised by food cues being reinforced through classical conditioning (Folkvord et al, 2016). Other literature also confirms that people's attitudes reflect the way they evaluate the world around them, their likes and dislikes, and can be seen as powerful predictors of buying decisions (Ajzen & Fishbein, 2005; Eagly & Chaiken, 1993; Shrum, 2003).

From the assessment of the changing provision of information in labels has changed we conclude that in presence of multiple sources of claims and the use of labels as marketing strategy, it is important to investigate the perception of trustworthiness of the labels.

As a final remark, the evolution of consumer preferences and the increasing complexity of companies' labelling strategies raise further interest in the role of information overload. In fact, within this multifaceted decision-making process consumers preferences for fish appear both complex and partially conflicting. Wild caught fishes have an advantage in terms of freshness, but promote depletion, which is counteracted by farming, which however causes environmental damage (Carlucci et al. 2015; Claret et al. 2014; Reynolds et al. 2014). Whole fish (especially for certain varieties) transmit a signal of quality, but are difficult to prepare and make consumers willing to pay a premium for processed fish (e.g. filleted) that raise a concern in terms of quality along the value chain (Cardoso et al. 2013; Nguyen et al. 2015). These complex trade-offs generate a pressure towards information overload in labels.

2.3. A research plan

Given the existing regulatory framework and the need to revise the existing legislation, this study follows a three-steps approach addressing three corresponding research areas: correct recall of information (understanding), trustworthiness of information (beliefs and values), use of the information for the correct execution of a task (behavioural assessment). For the purpose of this research, we define a "claim" as a piece of information. In other words, labels comprise sets of claims.

As a starting point, the revision of the regulation requires understanding of the information provided by the label. We designed an experiment where we pose questions assessing the correct recall and processing of the different claims of interests. How should we design such a test? The outcome should be an incentivized measure of recall and understanding, to avoid the social desirability bias in the response (Zizzo 2010; Smith 1976). We chose a setting in which participants have restricted time to process the information, as it normally occurs while purchasing groceries. Additionally, the validity of the test depends on the way in which the information is provided: comprehension may change when a single piece of information is administered individually or within a set of other attributes, for reasons of attention, limited processing capacity, presence of distractions etc. For example, the outcome can change if the participant is shown a sentence regarding the country of origin of a caught fish and then immediately asked to answer a related question, or when the consumer is asked to read a full paragraph stating the country of origin among other things and then asked the same question. The claims are introduced into complete labels, as the ones currently in use or newly designed variants. From this study, we and the public authorities can assess whether, on average, a claim is recalled and processed, and how assembling these claims in different formats affects the level of understanding.

The rationale of this first experiment is to accomplish an explicit requirement of the directive. According to the legislation, the information of the label must be understood and must not deceive. Since mandatory claims are simple bits of information, comprehension is measured as a correct recall. Given that labels include multiple and potentially distracting stimuli, this task requires working memory and cognitive control. Of course, recalling labels in a controlled environment may be less ecologically valid than in a real purchase task where people search for information and is exposed to many more sources of information.

The assessment of the labels does not end with recall and cognitive processing. Implicitly, this assumes a neutral source, which is providing the information to accomplish the same goal as the public authority. This is not the case for most of the pieces of information included into labels, which represent statements by companies whose interest is the maximization of profits. In other words, the label should be trustworthy to have meaningful effects on consumers' choices. How do we assess trustworthiness? In standard interactions, we could elicit beliefs about the trustworthiness. But FAP purchase is not the setting to mimic a standard trust game. In FAP purchase, the individual consumer has very limited opportunity to directly assess the informational source and directly discriminate its reliability (Brécard et al. 2009; Menozzi et al. 2020). As a result, besides standard strategic beliefs, the behaviour is likely to reflect conformity to shared beliefs and perceptions (which are higher order beliefs). We designed a task adapted from the literature on natural processing language (Houser and Xiao 2011) and injunctive norms (Krupka and Weber 2013). From this study, the authority can learn whether, on average, adding a claim increases trustworthiness, promoting individual trust (Fehr 2009).

As a last step, we assess the labels behaviourally. Of late, the impact of labels is a concern for policymakers, because sustainability labels are relatively ineffective (Grunert, Hieke, and Wills 2014). When consumers purchase FAP, they are trying to accomplish certain objectives (maximize their utility, satisfice, etc.). We measure to what extent labels help accomplishing such objectives. To achieve greater control, we give the consumers instructions instead of relying on elicited preferences. We designed a marketplace, including types of goods (FAP and others) and varieties within each type. Then we manipulate labels and instructions to identify which labels work best and whether assembling the information in different ways changes the behavioural outcome. From this third study, the Authority can learn whether, on average, a label helps the consumers to achieve their goals.

3. Study I

Study I tests the level of correct recall and understanding of country of origin claims in labels on fish packages, by asking respondents a set of incentivized questions. The labels used are realistic mock-ups (Fig. 1), including additional information unrelated to the country of origin of the fish. Different variations of the labels allow us to identify whether voluntary and mandatory information on country-of-origin are differently recalled, but also whether providing additional information increases or decreases the recall of mandatory claims.

3.1. Experimental design and procedure

In this experiment, the main task consists of answering incentivized questions. Each participant sees a visual stimulus for 15 s and then is asked three questions. Each question has four possible answers, one of which is correct. This task is repeated four times, for a total of 4 visual stimuli and 12 questions. The order of the stimuli and of the questions within each round is randomized. The visual stimuli are labels for FAP products. The labels are: processed fish (tuna), caught unprocessed fish (tuna), farmed processed and farmed unprocessed crustacean (both shrimps). These are the exact instructions for the respondent "you will be asked to memorize the information included in a fish product label. After seeing the label for 15 s, you will be asked to answer three questions about the information presented. There will be four labels and three questions for each of the labels in total. Each one will include four different answers, only one answer is the correct one.".

Fig. 1. An example of label.

Participants are assigned between-subjects (50 % chance) to one of two conditions: *mandatory information only* and *mandatory plus voluntary information*. In the mandatory-information-only treatment, the label has the following baseline attributes: Name; Net weight; Price/kg; Best before; Defrosted; Fishing gear; Production method; Food operator (name and address); Barcode. In this treatment, the following mandatory country of origin attributes are included: the catch area (FAO Area, Sea/Ocean, without map) and the country where it was processed (for caught and processed tuna), the catch area (for caught and unprocessed tuna), where it was farmed and processed (for farmed processed shrimps) and where it was farmed (for farmed unprocessed shrimps). In the mandatory-plus voluntary information the upper part of the labels is identical to the other conditions, but the lower part includes additional information. This additional information includes the following "voluntary origin related" attributes (Vessels' flag state, Port of landing, Quality claim, Environment claim, and Nutritional content), whose content depend on the category of fish. Table 1 summarizes the set of claims across labels.

The set of questions is assembled in the following way. For each round in the mandatory-information-only treatment, the participant answers a question on mandatory country attributes and two additional questions (*distractors*). The former is randomly drawn from a set of questions on mandatory claims and vary by label (e.g. "*Where was the fish farmed?*" could not be asked after a label for caught fish). The two distractors are drawn from the set of questions concerning baseline attributes. For the mandatory-plus-voluntary condition, the participant answers a question on mandatory country attributes, one question randomly drawn among the questions on voluntary attributes, and one distractor (a question on baseline attributes). The set of distractors is the same for all labels and treatments, but the specific questions posed at each round is randomly drawn without replacement. The full set of questions cannot be reported here for reason of space, but is included in the SOM, Section 1, discriminated by label. We will define target questions as the ones over mandatory country attributes and voluntary attributes. We will call distractors the other ones.

The incentives work in the following way. At the end, one question is randomly selected for payment. If the question is answered correctly, the participant receives one euro, zero otherwise. The payment is additional to the participation fee. The choice to pay only one randomly chosen question is done to incentivize responses but controlling for the income effect (by randomly paying one question, the level of income is not changing during the experiment).

The SOM: Section 1 includes the labels, the protocol, and the set of questions.

The survey was administered using an online panel in 8 EU member states: Sweden, Denmark, Croatia, Romania, Germany, Czech Republic, Italy, and Spain. The online panel is administered by a market research company and matches the population in terms of gender, age, region of residence in every European country. In each country, 800 participants, between the ages of 18 to 74 were interviewed. Those who had not bought any FAP in the last 30 days were filtered out. The sample was balanced by gender and age (three age groups) in each country.

The procedure was as follows. The respondent received the

Table 1
FAP labels for Experiment I.

	Mandatory Information Only treatment			
Attributes	Label 1 (Processed – Caught)	Label 2 (Processed – Farmed)	Label 3 (Unprocessed – Caught)	Label 4 (Unprocessed – Farmed)
Name	Yellowfin Tuna – *Thunnus Albacares*	Shrimp – *Caridea*	Yellowfin Tuna – *Thunnus Albacares*	Shrimp – *Caridea*
Price/kg	11€/kg	11€/kg	11€/kg	11€/kg
Weight	1,5kg	1,5kg	1,5kg	1,5kg
Best before	15/08/2020	15/08/2020	15/08/2020	15/08/2020
Defrosted	NO	NO	NO	NO
Fishing gear	Hooks and lines	Trawls	Hooks and lines	Trawls
Production method	Caught	Farmed	Caught	Farmed
Food operator	*Bolton Alimentari SPA Cermenate – Italy +39 031 779,111*	Nordic Seafood Hirtshals, Denmark +45 98 94 15 33	*Bolton Alimentari SPA Cermenate – Italy +39 031 779,111*	Nordic Seafood Hirtshals, Denmark +45 98 94 15 33
Catch area	Northeast Atlantic (FAO Fishing Area 27)		Northeast Atlantic (FAO Fishing Area 27)	
Farmed in		India		India
Processed in	Portugal	Denmark		
Barcode				
	Mandatory plus voluntary information treatment			
Attributes	Label 5 (Processed – Caught)	Label 6 (Processed – Farmed)	Label 7 (Unprocessed – Caught)	Label 8 (Unprocessed – Farmed)
	Same attributes as Label 1	**Same attributes as Label 2**	**Same attributes as Label 3**	**Same attributes as Label 4**
Vessel's flag state	South Africa	South Africa	South Africa	South Africa
Port of landing	Hai Phong Port (Vietnam)	Hai Phong Port (Vietnam)	Hai Phong Port (Vietnam)	Hai Phong Port (Vietnam)
Quality claim	Italian quality	Danish quality	Italian quality	Danish quality
Environmental claim	Dolphin friendly	Dolphin friendly	Dolphin friendly	Dolphin friendly
Nutritional content	20 calories, 1.7 g fat (0.4 saturated fat), 58 mg sodium, 9 g carbs, 0 g fiber, 0 g sugar and 15.7 g protein	20 calories, 1.7 g fat (0.4 saturated fat), 58 mg sodium, 9 g carbs, 0 g fiber, 0 g sugar and 15.7 g protein	20 calories, 1.7 g fat (0.4 saturated fat), 58 mg sodium, 9 g carbs, 0 g fiber, 0 g sugar and 15.7 g protein	20 calories, 1.7 g fat (0.4 saturated fat), 58 mg sodium, 9 g carbs, 0 g fiber, 0 g sugar and 15.7 g protein

invitation, provided informed consent, and was asked the screening question regarding FAP purchase. After these preliminary steps, the incentive system was explained. This was followed by the main experimental tasks, a post-experimental questionnaire including socio-demographic details, and a set of measures on attitudes and perceptions about origin-related information for FAP (answers were reported on a Likert type scale). On average, the survey took around 20 min.

3.2. Hypotheses and analysis

We measure accuracy as a dummy equal to one if the answer is correct, and zero otherwise.

The following three hypotheses can be postulated: (H1) Accuracy in answering a target question is higher than if respondent selects answers at random; (H2a) Accuracy in answering questions over voluntary claims is lower than in answering distractors; (H2b) Accuracy in answering questions over mandatory country of origin claims is lower than in answering distractors; (H3) Accuracy in answering questions over mandatory country claims decreases when additional information is provided.

H1 is equivalent to assuming that consumers increase cognitive effort in the task due to the incentives (Smith, 1976). H2 is based on the observation that country of origin related information requires more processing than price or other more familiar claims (van Rompay, Deterink & Fenko, 2016; Ampuero & Vila, 2006; Karnal et al., 2016). H3 is our central hypothesis of information overload (Verbeke 2008; Sørensen et al. 2012).

H1 can be tested via a simple binomial test. The test consists of assessing that the accuracy in responses is different from 25 %, which would be the outcome in case of random selection of answers. This analysis can be run separately for each question. The second test is a *t*-test comparing the accuracy of target questions (separately for the two groups of origin related claims, mandatory and voluntary) with the accuracy in answering the distractors. The third hypothesis relates to the main treatment effect: a *t*-test establishing that the accuracy in answering mandatory origin related questions is higher when the label only includes mandatory information with respect to mandatory plus voluntary information.

To test H2 and H3, we estimate the Linear Probability Model described in the following Equation.[2]

$$Y_{it} = \alpha + \beta_1 D_{it}^{target} + \beta_2 D_i^{mandatory} + \beta_3 D_{it}^{voluntaryclaim} + \beta_4 D_i^{mandatory} \bullet D_i^{target} + \mathbf{D}_{it}^{question}\boldsymbol{\gamma}_1 + \mathbf{D}_{it}^{fish_{case}}\boldsymbol{\gamma}_2 + \mathbf{D}_{it}^{order}\boldsymbol{\gamma}_3 + \mathbf{X}_{it}\delta + \varepsilon_{it} \quad (1)$$

[2] The reader should notice that we use Linear Probability Model with Limited Dependent Variable everywhere in the article. This is a precise methodological choice, based on the fact that our interest does not lie in prediction, but in causal identification (Angrist and Pischke 2009), for which LPM is superior (Gomila 2021). Additionally, interpretation of the interaction effects in nonlinear model is less transparent and depends necessary on arbitrary choice of the values for the additional controls.

The outcome variable is a dummy for a correct answer (indexed by participant i and round t). The independent variables are a dummy equal to one if the question was a target (D_{it}^{target}), a dummy equal to one for the mandatory only treatment ($D_i^{mandatory}$), a dummy equal to one for a question on voluntary origin related attributes ($D_{it}^{voluntaryclaim}$), and the interaction between the first two indicator variables ($D_{it}^{target} \bullet D_i^{mandatory}$). The regression controls for order fixed effect (23 dummies), and fish type (three dummies), sex, age, education, marital status, household size, income. We now provide a formal derivation using a switching regression (Angrist and Pischke 2009) to clarify the interpretation of each coefficient and their linear combinations.

The model in [1] can be derived as follows. Using the standard notation for counterfactuals (Angrist and Pischke, 2009), we can write the potential outcomes for every subject in every state of the world using three superscripts: whether the subject is in the mandatory experimental condition (m: mandatory) or in the mandatory plus voluntary experimental condition (v: voluntary); whether the question is a distractor question (target or distractor); whether the question is about a voluntary claim (m - mandatory, v – voluntary, 0-none). Accordingly, define three different indicator variables. $D_i^{mandatory}$ has a value of 1 if the participant is assigned to the mandatory condition and 0 if is assigned to the mandatory plus mandatory condition (notice that it varies by subject only). D_{it}^{target} has a value of 1 if the question is a target question (origin-related) and 0 if the question is a distractor. The last one, $D_{it}^{voluntaryclaim}$, has the value of 1 if the question Is about a claim presented in the voluntary information section of the label. The observable outcome for subject i answering question t becomes Y_{it} and can be written as:

$$Y_{it} = D_i^{mandatory}(D_{it}^{target} Y_{it}^{m,target,m} + (1 - D_{it}^{target}) Y_{it}^{m,distractor,0}) + (1 - D_i^{mandatory})(D_{it}^{target}(D_{it}^{voluntaryclaim} Y_{it}^{v,target,v} + (1 - D_{it}^{voluntaryclaim}) Y_{it}^{v,target,m})) + (1 - D_{it}^{target}) Y_{it}^{v,voluntary,0})$$

After some tedious algebra, using the fact than when $D_i^{mandatory} = 1$, $D_{it}^{voluntaryclaim} = 0$ by construction, and redefining $\alpha = E[Y_{it}^{v,distractor,0}]$, $\beta_1 = E[Y_{it}^{v,target,m} - Y_{it}^{v,distractor,m}]$, $\beta_2 = E[Y_{it}^{m,distractor,0} - Y_{it}^{v,distractor,0}]$, $\beta_3 = E[Y_{it}^{v,target,v} - Y_{it}^{v,target,m}]$, $\beta_4 = E[\left(Y_{it}^{m,target,m} - Y_{it}^{m,distractor,0}\right) - \left(Y_{it}^{v,target,m} - Y_{it}^{v,distractor,0}\right)]$, by random assignment of $D_i^{mandatory}, D_{it}^{voluntaryclaim}, D_{it}^{target}$ we can write the following reduced form equation:

$$Y_{it} = \alpha + \beta_1 D_{it}^{target} + \beta_2 D_i^{mandatory} + \beta_3 D_{it}^{voluntaryclaim} + \beta_4 D_i^{mandatory} \bullet D_i^{target} + \varepsilon_{it}$$

Which leads to Eq. (1). In estimating Eq. (1), we explicitly control for order and questions fixed effects and other socio-demographics that are not affected by treatments. It is important to stress that this is a reduced form model, and not a structural model of recall. The coefficients should simply be interpreted as average causal parameters (differences in outcome) and not as structural parameters in the sense of a specific behavioral model.

3.3. Results

In Table 2, we report the individual claims with the average likelihood of correct recall. Whenever a claim is addressed by more than one question, we report the success rate separately by question. We note that accuracy is always greater than 50 %. The results support H1: the level of comprehension achieves an accuracy rate higher than random responses. Notice however, that the level of comprehension for mandatory information (at the top of the table) is higher than for voluntary information: it is 76 % and 65 % for the two questions on catchment areas, 75 % and 63 % for the "farmed in" claim, and 66 % for the "processed in" claim (mandatory) versus 63 % for food operator, 62 % for quality claim, 85 % for the flag, and 68 and 53 % for the port of landing (voluntary).

We now move to the regression results (Table 3). The coefficient $\beta_1 + \beta_3$ identifies whether consumers understand the actual content of the voluntary information currently displayed on FAP labels, as it provides the accuracy rate when responders are asked a voluntary information question with respect to a distractor. The estimated coefficient is −0.086 (t = -7.91, $p < 0.001$), i.e. the likelihood of answering correctly is lower than when the question addresses a nutritional or other claim, by 8.6 percentage points (pp). This supports H2a.

We analyse H2b in two parts: first, we compare comprehension of mandatory claims to distractors in labels with mandatory information only; second, we compare comprehension of mandatory claims to distractors in labels with mandatory plus voluntary information. The first part is provided by the coefficient $\beta_1 + \beta_4$. As seen in Table 3, the estimated coefficient is −0.009, not statistically significant (t = -0.77, p = 0.441). In other words, for mandatory labels (existing legal framework), the evidence for H2a is inconclusive. The second part is provided by coefficient β_1. The estimated coefficient is −0.021 (t = 1.90, p = 0.056). In conclusion, the evidence partly supports the hypothesis that origin-related claims are more difficult to recall and process than other claims in FAP labels.

Table 2
Accuracy in the response by claim and question.

Claim	Question	Probability of Answering Correctly	Binomial Test (p-value)
Catch Area	Where was the fish caught?	0.76	<0.01
	In which of the following regions/countries was the fish caught?	0.65	<0.01
Farmed in	Where was the fish farmed?	0.75	<0.01
	The fish was farmed in which of the following countries	0.63	<0.01
Processed in	Where was the fish processed?	0.66	<0.01
Food operator	Where is the food operator based?	0.63	<0.01
	Choose the country where the food operator comes from:	0.63	<0.01
Quality Claim	What is the country referred in the quality claim?	0.62	<0.01
Vessel's Flag State	Which countr''s flag appears in the vessel flag?	0.85	<0.01
Port of Landing	In which country is the port of landing of the vessel based?	0.68	<0.01
	Where is the port of landing?	0.53	<0.01

Table 3
OLS regression results, Experiment I.

Variables	(1) Question answered correctly
Target	−0.021*
	(0.011)
Mandatory	0.044***
	(0.005)
Voluntary claim	−0.065***
	(0.013)
Target#Mandatory	0.012
	(0.008)
Constant	0.503***
	(0.014)
Observations	76,800
R-squared	0.051
Question Dummy	Yes
Controls	Yes

Note: Robust standard errors in parenthesis. *p < 0.1, **p < 0.05, ***p < 0.01.

Notice that, while we did not formulate an *ex ante* hypothesis, we can compare voluntary claims with mandatory claims. This controls for the possible bias induced by the specificity of origin related questions. The accuracy rate decreases by 6.5 pp (coefficient β_3, t = -5.12, p < 0.001) when the claim is voluntary with respect to a mandatory-one.

Finally, a test of H3 is provided by the coefficient $\beta_2 + \beta_4$. The latter corresponds to the likelihood of correct recall of a mandatory claim in the *mandatory treatment* with respect to the *mandatory plus voluntary treatment*. In support of the *less is more* effect, accuracy is higher by around six pp in the first case (t = 9.68, p < 0.001).

Notice that this goes beyond a simple bandwidth effect on the working memory (Mani et al 2014). In fact, going from mandatory to mandatory plus voluntary information, the accuracy decreases by 6 % for target questions, whereas for distractors, the reduction is only 2.8 % (t = -4.83, p < 0.001). This suggests that the performance is worse than average for country of origin claims, in presence of information overload.

Before moving to the next study, we briefly comment on the characteristics of the sample. The features of the sample may condition the outcome we measured, e.g. more educated people may comprehend more and older people may be less trained to answer questions. If this is the case, lack of representativeness may induce a different distribution between our sample and the population. Since our intention is not descriptive, this might be of limited concern. Nevertheless, different segments of the population may respond differently to the treatment (*heterogeneous treatment effect*). If the policymakers are trying to forecast the overall impact on the population, adjusting for differences in composition becomes relevant.

We benchmark on Eurobarometer (2018), which performed an analysis of FAP consumption. Their question on purchasing experience was slightly different from ours, since their questionnaire asked the frequency of FAP purchase in the previous year, whereas we asked

Table 4
Descriptive Statistics.

Panel A: Sample from Study I

	COUNTRY							
	DE	**IT**	**DK**	**ES**	**SE**	**CZ**	**RO**	**HR**
EDUCATION								
Primary school or less	3.63	6.25	19.50	1.75	8.63	6.25	1.63	0.88
High school	50.25	49.88	29.13	33.38	36.38	62.75	38.50	49.75
Some years of university	10.88	11.50	9.25	14.88	12.50	5.00	5.75	12.50
University degree	26.25	22.75	29.88	38.13	37.63	6.75	41.38	28.88
Post-graduate	9.00	9.63	12.25	11.88	4.88	19.25	12.75	8.00
GENDER								
Male	42.25	41.25	38.38	41.50	39.38	47.00	42.13	38.63
Female	57.75	58.75	61.38	58.38	60.38	53.00	57.75	61.38
MARITAL STATUS								
Single	28.00	30.00	28.88	28.00	34.25	41.38	27.13	25.00
Married or Union	57.50	62.88	63.50	61.38	57.63	45.00	62.75	62.88
Widowed/Divorced	14.50	7.13	7.63	10.63	8.13	13.63	10.13	12.13
SOCIAL CLASS (EUR)								
<20 k	22.00	28.88	15.63	34.00	23.75	57.63	75.25	71.88
From 20 k to 30 k	19.63	27.88	12.25	27.00	16.38	23.38	10.50	15.25
From 30 k to 50 k	29.38	28.88	21.13	26.63	27.75	14.75	6.00	4.88
From 50 k to 150 k	27.63	12.00	44.50	11.25	27.38	3.63	6.13	7.13
More than 150 k	1.38	2.38	6.50	1.13	4.75	0.63	2.13	0.88
CHILDREN AT HOME								
No	60.88	42.13	65.13	42.88	60.75	52.13	49.13	53.75
Yes	39.13	57.88	34.88	57.13	39.25	47.88	50.88	46.25

Note: Sample used for the analysis in Study I. Includes only those that purchased fish at least one in the last 30 days

Panel B. Sample from Eurobarometer

	COUNTRY							
	DE	**IT**	**DK**	**ES**	**SE**	**CZ**	**RO**	**HR**
EDUCATION (How old were you when you stepped out of education?)								
15	18.25	26.85	2.23	32.92	5.39	3.36	11.76	4.78
16–19	44.75	45.30	12.28	29.33	15.67	67.11	55.69	64.09
20	31.60	21.31	73.33	28.09	75.03	25.50	26.27	21.91
still in educ	5.40	6.04	3.47	7.08	3.67	4.03	5.88	8.90
no full time educ	0.00	0.50	8.68	2.58	0.24	0.00	0.39	0.33
GENDER								
Male	50.05	44.91	47.73	46.24	54.39	38.17	44.89	42.46
Female	49.95	55.09	52.27	53.76	45.61	61.83	55.11	57.54
MARITAL STATUS								
Single	16.74	14.63	13.56	18.96	21.38	10.49	11.24	19.19
Married or Union	63.66	75.12	66.21	67.41	59.42	62.50	76.74	69.59
Widowed/Divorced	19.61	10.24	20.22	13.64	19.20	27.01	12.02	11.22
SOCIAL CLASS								
Working class	17.23	12.04	15.29	44.42	16.87	17.91	29.56	18.24
Lower middle class	14.26	15.89	4.51	15.40	8.68	18.82	8.53	11.77
Middle class	55.74	61.87	58.90	36.72	56.11	56.92	57.14	60.86
Upper middle class	12.23	9.70	20.55	3.35	17.60	5.90	3.77	7.79
Higher class	0.53	0.50	0.75	0.11	0.73	0.45	0.99	1.33
CHILDREN AT HOME								
No	83.89	75.44	76.93	72.01	86.40	74.11	64.74	72.77
Yes	16.11	24.56	23.07	27.99	13.60	25.89	35.26	27.23

Note: Eurobarometer (2018). Sample selected for those who bought fish at least one per month during the last year.

whether they purchased fish the month previous to the interview. Since we excluded those that did not, we also selected from Eurobarometer those participants that purchased at least once per month. In evaluating commonalities and differences we should also take into account that the questions are slightly different, and that there is a time lag between the recollection of our data and Eurobarometer's, which may potentially be associated with changes in the pattern of consumption or the standard of living.

With all these caveats in mind, these are the main results of the comparison. Table 4, Panel A report the descriptive statistics from Study I, and Panel B those from Eurobarometer. In Study I, in Germany, Italy, Czech Republic and Hungary, the largest share of participants (very close to or above 50 %) had at most high school, while in the other countries, there are two modes, at high school and university degree. In Eurobarometer, the survey collected information on the age at which the respondents exited from school. The data from Eurobarometer are qualitatively the same: in the same countries the largest share of participants ended school at 16–19, which is high school, whereas in the other countries there is a mode corresponding to tertiary education.

In both samples, most of the participants in all countries are married or in civil union, from half to two thirds of participants. The gender composition is more skewed towards self-declared women in the sample from Study I (in the question we administered, gender was asked with a non-binary response, but the share who chose "Other" is statistically indistinguishable from zero). Most of our Study I sample overestimate the number of respondents with children at home, with respect to Eurobarometer. Data on income distribution are impossible to compare because Eurobarometer asked about perception of social class, while we had information on self-reported income.

Since sampling across studies is performed with the same design and out of the same population, the discussion for Study I extends to Study II and III.

4. Study II

Study II is a labelling design experiment, assessing the trustworthiness of the potential claims. The literature in social science distinguishes between trust and trustworthiness. Trustworthiness is the propensity to return trust. Trust is the willingness to put resources at the disposal of a counterpart (Fehr 2009). The willingness to trust relies on a belief of trustworthiness. In the present setting, buying food requires trust, given the potential risks. To trust, consumers require the labels to be trustworthy.

According to the REFCAM Model by Folkvord et al (2016), visual cues in packaging and labelling triggers a set of reactions over attitudes and perceptions. This reactivity to visual cues reflects shared perceptions of what is accurate and reliable. Since the model states that these perceptions drive purchase, we need to be able to measure perceptions separately from behaviour. Shared perceptions are usually elicited as second order beliefs because this method provides an exogenous benchmark over which to assess accuracy. But literature points at perspective taking as a crucial aspect of shared perception (Wolgast et al. 2020): this justifies the use of second order beliefs.

Our intuition is that this perception of appropriateness has points in common with the literature on social norms (Bicchieri 2016), shared behavioural scripts that decisionmakers follows and expect others to follow in kind. A pioneer method to elicit shared perception uses coordination games (Krupka and Weber 2013). A coordination game is a strategic interaction where agents are rewarded for choosing the same strategy as the counterpart. Extensive evidence shows that human subjects solve coordination game by means of saliency (Mehta, Starmer, and Sugden 1994). The key is posing the question in a way that makes the shared belief immediately focal and avoid other source of focality[3] (Fallucchi and Nosenzo 2020; König-Kersting 2021).[4]

In practice, after posing a question, we pay participants to match the modal answer of the respondents. Matching the modal response leads to a prize while mismatching does not. This becomes a coordination game, where each possible response is a Nash equilibrium. According to the theory of saliency (Mehta, Starmer, and Sugden 1994), if a particular answer is focal (for reasons that may or may not be payoff relevant), it will drive the behaviour of the large majority of respondents. If the question highlights the trustworthiness, the shared perception of trustworthiness should become focal.

There are two treatments, assigned between subjects: a time pressure treatment where the labels are shown for only 10 s and a control, where they are displayed for 60 s. This was designed to assess the extent to which intuitive and deliberative modes of thought produce different outcomes. The main task is repeated twice, once for farmed fish and once for caught fish.

4.1. Experimental design and procedure

Respondents are taken through two hypothetical purchasing scenarios. A FAP is thoroughly described (product type, weight, price per kilogram, scientific name, best-by date, storage condition, fishing gear, environmental information, and nutritional content) using textual information. In a subsequent screen, the same information is displayed in a label, in the style of a real label, but with one blank to be completed adding the country of origin information (Fig. 2).

Respondents are instructed to complete the label by choosing one claim from a set of available claims to fill the blank; the one that would make the label the most trustworthy ("Your objective is to choose the piece of information that you believe will make the most trustworthy label if you want to know where the fish came from, from a set of alternatives that we will show you."). Similarly, they are instructed to select a claim to go into the garbage bin, the one that would make the label the least trustworthy ("After the first selection, you will be asked to pick one of the remaining pieces of information as the one that would have made the label the least trustworthy if you want to know where the fish came from"). The choice set includes the entire list of possible claims, both mandatory and voluntary. Respondents are paid to match the modal response. The screenshot of the task is shown in Fig. 2. The task is repeated twice, once for caught fish and once for farmed fish. For the latter, the textual description and the label displayed to them corresponded to a farmed shrimp, for the former, the information provided corresponded to a yellowfin tuna. The order is randomized, to control for the effect of learning.

After completing the two choice tasks, participants answer a post-experimental questionnaire. On average, the survey takes around 20 min. Participants receive a standard participation fee and one euro if their answer matches the modal response. Only one randomly selected

[3] Notice that participants will win the prize if they coordinate on any available claim. This may create the phenomenon of sunspot equilibria. Think for example at a situation where all the claims are red and only one is blue, if they only choose the blue because it is different from the other, they will win the prize. The instructions prime participants to look for saliency in shared beliefs of trustworthiness and should avoid generating other sources of focality.

[4] The use of coordination game to measure shared perception or classifications is well established and the literature agrees that it is more robust than other survey methods. The discussion is beyond the scope of this paper, but a detailed discussion is in Goerges and Nosenzo (2020), Erkut (2020), d'Adda et al. (2016), Veselý (2015), Houser and Xiao (2011).

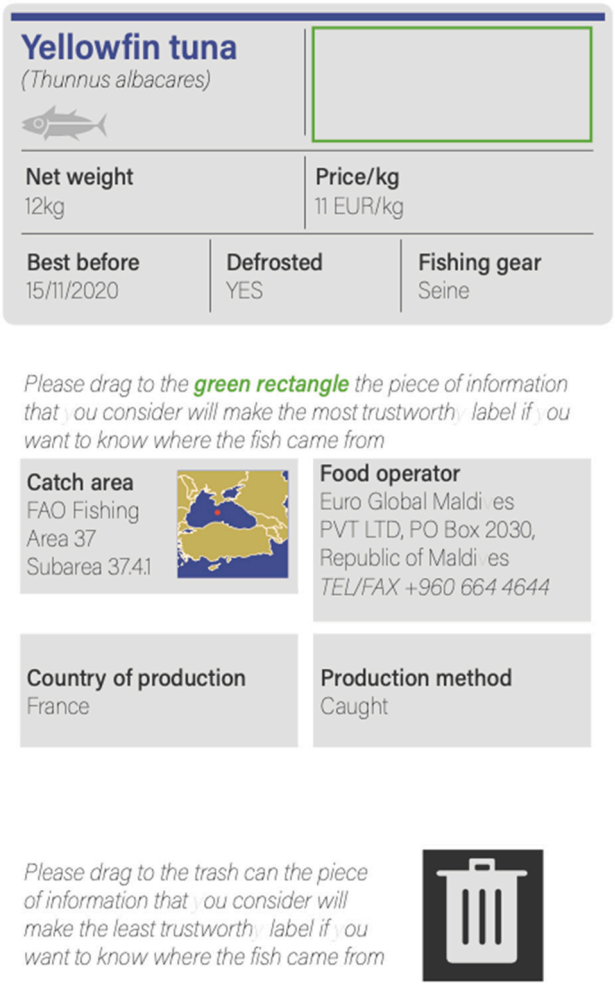

Fig. 2. The main Task in Study II.

task is incentivized (with both choices), to avoid income effects. The incentive system was thoroughly explained (including two comprehension questions) prior to any decisions. In the explanation of the incentives, we made clear that: a) Payment was postponed after all participants completed the survey to determine the modal response; b) Modal responses were collected at the country level.

Participants are split in two groups, time pressure (10 s) and control (60 s).

The experiment was run in four countries, Spain, Hungary, Germany, and Sweden. Participants are randomly sampled from an online panel. The access to the panel is provided by the same market company as in Study I. The criteria of representativeness are the same: participants match population on a restricted set of observable characteristics. In drawing our sample, we included an additional filter of previous experience in purchasing FAP. We surveyed 180 participants per country. The full experimental protocol, in English, is available in the SOM: Section 2.

4.2. Hypothesis and analysis

To analyze the data, we assess the distribution of claims across tasks using a Kolmogorov Smirnov test. The null hypothesis is the lack of a treatment effect, against the alternative that the distributions are different.

Then, we compute the likelihood of selecting a claim. We build a score variable for each claim in each task which is equal to + 1 if the claim is selected as trustworthy and −1 if selected as not trustworthy. In this way, the total sum over the claims and participants, separately for each task, sum to zero. If a claim has a positive mean, it has been consistently chosen as more trustworthy, and vice versa.

4.3. Results

There is no effect of time pressure on choices. In other words, when comparing the distribution across claims between the two time-conditions, we cannot reject the null hypothesis that the distribution is the same. For the caught fish, we cannot reject the null hypothesis either for the selection of the most trustworthy claim (K-S, $D = 0.04$, $p = 1.00$) or for the least trustworthy claim (K-S, $D = 0.09$, $p = 0.76$). A similar conclusion holds for the farmed fish (K-S, $D = 0.06$, $p = 0.99$ and $D = 0.08$, $p = 0.92$ respectively).

In Fig. 3, we plot the relative frequencies per claim, separately for caught and farmed FAP. The claims are reported in order of relevance: first the mandatory claims, then the voluntary claims, then the fixed attributed (Food operator). For caught fish, the mandatory claims are caught area and "processed in" and the voluntary claims are Vessel's Flag, Port of Landing, and Quality Claim. For farmed fish, the only difference is that the first mandatory claim is "Farmed in". It is possible to test the direction of the selection using a *t*-test, with the null hypothesis that the relative frequency is not different from zero. In the caught FAP task, the flag ($t = 3.46$, $p < 0.01$) and the catch area are systematically selected as more trustworthy ($t = 3.56$, $p < 0.01$). "Processed in" ($t = -4.34$, $p < 0.01$) and the quality claim ($t = -5.13$, $p < 0.01$) are systematically discarded. Port of landing ($t = 0.46$, $p = 0.64$) and food operator ($t = -0.29$, $p = 0.76$) are neither selected not discarded. This evidence suggests that caught area is indeed the most relevant information but indicates that visual information is relevant (because of the flag and supporting REFCAM model).

In the farmed FAP task, the flag ($t = 3.85$, $p < 0.01$) and the quality claim are systematically selected as more trustworthy ($t = 2.37$, $p = 0.01$). The "Farmed in" claim is systematically discarded ($t = -3.76$, $p < 0.01$). Port of landing ($t = -0.43$, $p = 0.66$), "Processed in" ($t = -0.64$, $p = 0.51$), and the food operator ($t = -0.66$, $p = 0.50$) claims are neither selected not discarded. The evidence confirms that visual elements are important (again supporting the model). Notice that the vessel's flag in this case is not informative (for farmed fish), but respondents are giving it a high valuation in terms of trustworthiness because they may be interpreting it to indicate the country where it was farmed instead of the country of the vessel. Additionally, in this case, quality certification is considered more trustworthy than disaggregated information over country of processing. This is important because it suggests that respondents construe processed food in a different way and hold a shared perception that "composite" indicators summarizing the underlying process or disclosing traceability (Hinkes and Schulze-Ehlers 2018) are more trustworthy. In other words, the fact that "processed in" is perceived as unreliable while the quality claim as reliable suggests that consumers perceive independent evaluators along the value chain as more trustworthy.

5. Study III

In Study III we assess the label behaviorally. We set out to understand to what extent a label meets consumers' need at the point of sale. Since observational data only informs us on what consumers actually do and not what they aimed to do, we control the environment by giving respondent a list of instructions. Respondents are asked, following the set of instructions, to complete a purchasing task by selecting a list of items from a set of options. One instruction involves FAP and the other ones involve other products. We did not ask only about FAP because we wanted to avoid experimental demand: had we only inquired about FAP, our main interest would have been obvious, and respondents may have either paid extra attention to labels or changed their behavior to react to our research question. Our online experiment on a programmed

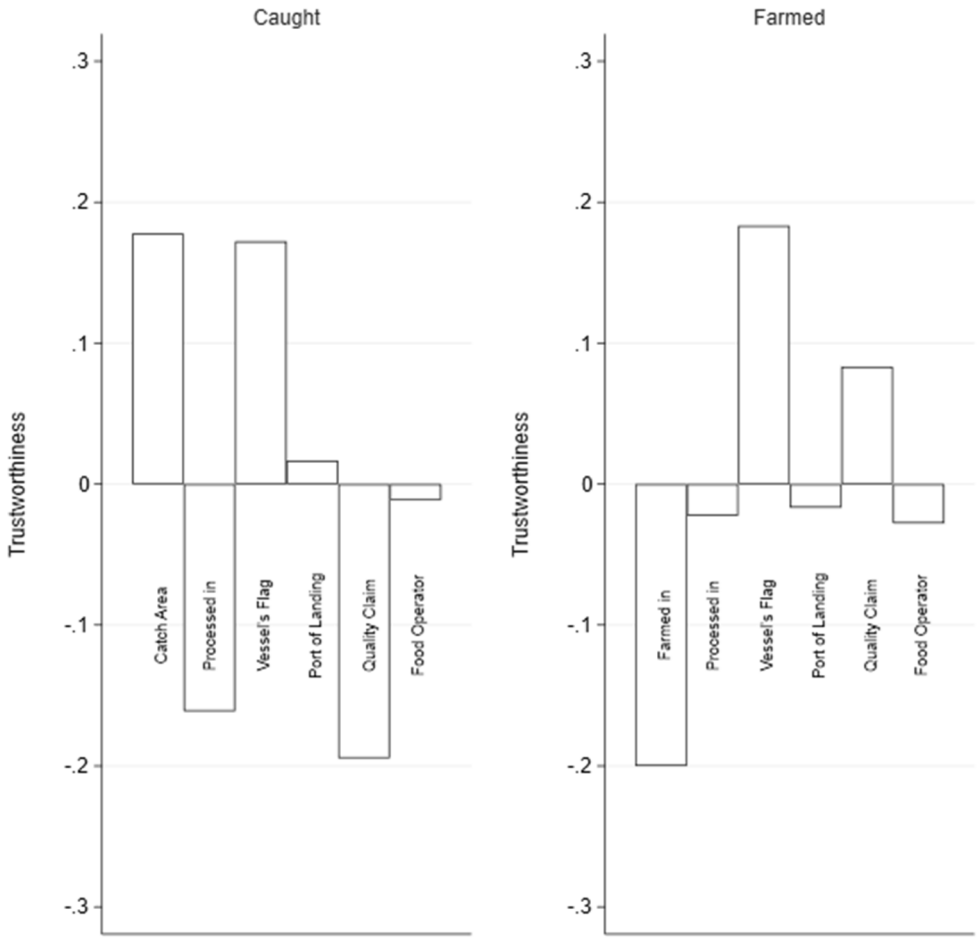

Fig. 3. Relative frequencies of choice. Labelling Design Experiment.

platform manipulates the label attached to FAP that respondents should read to complete the instruction. One instruction in the list is randomly drawn for payment: participants receive their variable incentive if the instruction has been correctly executed.

5.1. *Experimental design and procedure*

In this experiment, a platform is programmed to be like those used for a food delivery service. The platform is a 7 by 5 matrix, embedded in a food delivery platform (Fig. 4). By row, respondents see classes of products and by column, the product varieties within each class. The classes of products are: milk, meat, bread, fruits, cereals, vegetables, and fish. Each cell is associated with a label with the full set of characteristics, that can be opened with a click. This label in the marketplace displays a set of general claims (i.e., Type of Product, Commercial Name, Brand, Price, Net Weight, Price per Kilogram, Best-Before Date, Nutritional Content, Storage Conditions) and some mandatory origin-related claims (i.e., Country of Production, Processed in, and Food Operator).

The list of instructions comprises five items. An example is shown in Fig. 5. The target instruction regards FAP, the other four are distractors. Respondents are instructed to find the variety of a product matching a required feature. Having selected an item, the respondent places the product in the shopping cart, as in a realistic setup (Fig. 6). Given the complexity of the task, we make participants perform a practice round with feedback.

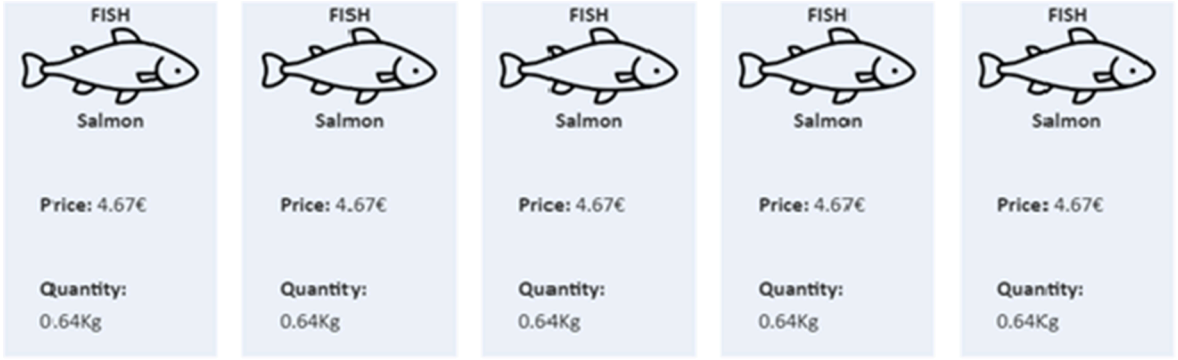

Fig. 4. A section of the online platform: the fish section.

Here is the list of products you need to choose in the market page.

- **A fish produced in Scotland**
- **The lowest price per kilogram cereal**
- **The cheapest milk**
- **The apricots with the farthest expiry date**
- **The meat with the farthest expiry date**

The platform that you will use has 4 buttons at the bottom of the screen:

1. Go to instruction page
2. Go to the shopping list (this page you are reading now)
3. Go to the marketplace
4. Go to the shopping cart

Remember that once you added the five required products to the cart and are satisfied with your choice you have to go to the shopping cart and click on the "finish" button!

Fig. 5. The shopping list.

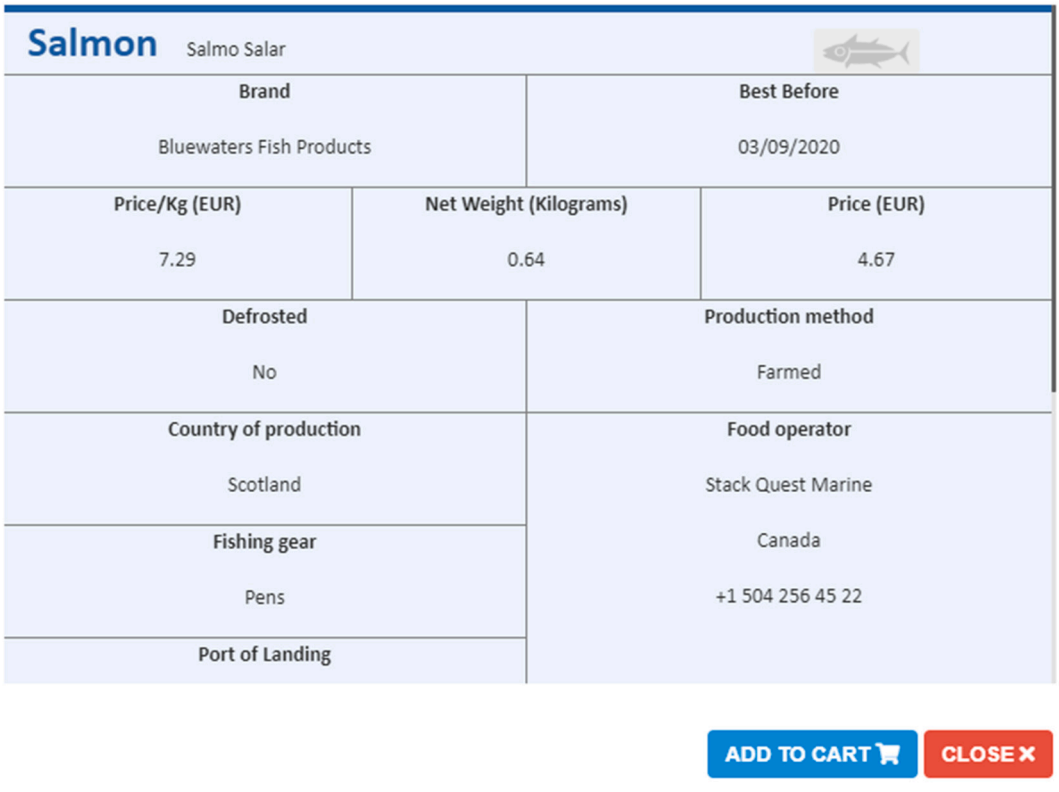

Fig. 6. Screenshot from the online platform in Study III.

This experiment follows a 5 by 3 factorial design, where factor one is the FAP label and factor two, the instruction for the FAP in the shopping list. The label treatment consisted of five different conditions related to the type of origin-related information shown in the label of FAP. The five conditions we used were:

- **Control group:** Mandatory only information.
- **Control + Vessels flag state:** In addition to the mandatory information, the labels contained an image of the vessel's flag state along with its written name.
- **Control + Port of landing:** In addition to the mandatory, the label contained the written name of the port of landing of the ship.

- **Control + Vesseĺs flag state & Port of landing:** In addition to the mandatory information, the label contained both the vessel's flag state and the port of landing as described in the previous conditions.
- **Information Overload:** This is the same as the previous treatment but with the addition of a non-origin related piece of information related to the environmental impact of the product́s fishing practices.

The second factor was a "question treatment", which consisted of three levels depending on the country-of-origin attribute that was required to check. The three possible instructions for fish were:

- Country of production.
- Food Operator.
- Processed in.

This experiment was administered via an online panel to 8 EU member states: Croatia, Czech Republic, Denmark, Germany, Italy, Romania, Sweden, and Spain. In each country, 800 respondents aged 18 to 74 were interviewed (The final sample was 6643). The sample was balanced by gender and three age groups in each country, we included an additional filter of having bought FAP in the previous 30 days. The access to the panel is provided by the same market company as in Study I and II. The criteria of representativeness are the same: participants match population on a restricted set of observable characteristics. The full experimental protocol, in English, is available in the SOM: Section 3.

5.2. Hypothesis and analysis

The outcome variable is an indicator variable for whether the instruction for FAP is correctly executed. The main Hypothesis (H1) is that including additional information within the label will reduce the mean accuracy of execution with respect to mandatory information only.

To test H1, we estimate by OLS the following model:

$$Y_i = \beta_0 + \beta_1 Vessel + \beta_2 Port + \beta_3 Vessel\&Port + \beta_4 InfoOverload + T_i\gamma + X_i\delta + \varepsilon_i \quad (2)$$

Where Y_i is the outcome (indexed by the respondent *i*), Vessel, Port, Vessel&Port, InfoOverload are the dummies for each label condition, T_i is a set of dummies for the question treatment, X_i is a set of socio-demographic (gender, education, civil status, household size, income[5]). As for Study I, this equation is a reduced form, and not a structural model. The coefficients should simply be interpreted as average causal parameters (differences in outcome) and not as structural parameters in the sense of a specific behavioral model.

To provide a statistical assessment of H1 we look at the estimated coefficients of the treatment dummies. They could be interpreted as mean outcome difference with respect to the control treatment, which is the mandatory information only.

5.3. Results

Accuracy in the five conditions is reported in Fig. 7. The instruction is correctly executed 47 percent of the time in the control condition, with mandatory information. It is notable how accuracy decreases in all other conditions regardless of the additional amount of information provided. In fact, accuracy is 40 percent in the Vessel condition, 45 percent in the Port, 43 percent in the Vessel-Port, and 42 percent in the Information Overload treatment. Additionally, information accuracy decreases monotonically when more information is added, from Port, to Vessel & Port, to Information Overload. For the case of Vessel alone the accuracy is even lower than information overload, although the difference between the two is not statistically significant (t = 0.83, p = 0.40).

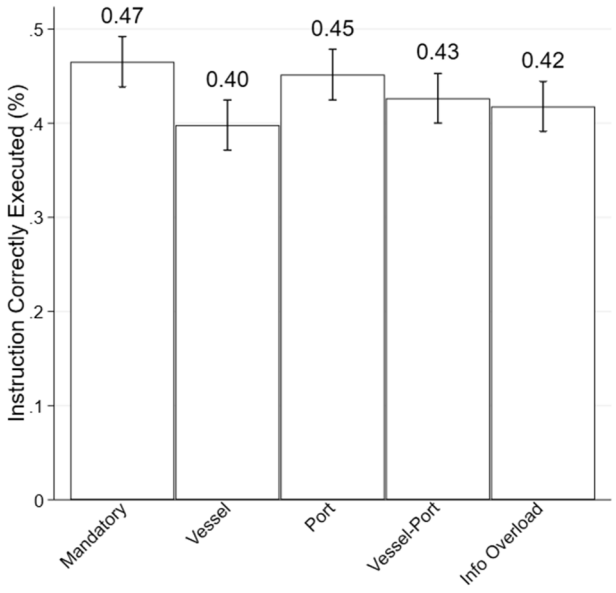

Fig. 7. Average accuracy by treatment. The list of treatment is displayed according to order of presentation in Experiment III Design.

A regression analysis provides a statistical test of the difference in accuracy. We estimate Equation [2] by OLS. The results are reported in Table 5. The introduction of Vessel's Flag reduces accuracy by 6.1 percentage points (pp) (t = -3.30, p < 0.01); the introduction of the Port of Landing claim reduces accuracy by 1.2 pp (t = -0.68, p = 0.49); the introduction of both claims reduces accuracy by 3.8 pp (t = -2.13, p = 0.03), and the Information Overload treatment reduces by 4.5 pp (t = -2.48, p = 0.01).[6]

Table 5
OLS Regression Results, Experiment III.

Variables	(1) Instruction Correctly Executed
Vessel	−0.061***
	(0.019)
Port	−0.013
	(0.019)
Vessel&Port	−0.039**
	(0.018)
Info Overload	−0.046**
	(0.018)
Constant	0.598***
	(0.029)
Observations	6,643
R-squared	0.088
Question Dummy	Yes
Controls	Yes

Note: Robust standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1.

[5] Differently from Study I, we do not include age because due to a coding error some respondents, where allowed to report text or leave it blank, and we do not want to exclude observations. However, its inclusion is immaterial for the results and identification is independent from age because of random assignment.

[6] Notice that differences across treatments are not in general statistically significant but testing for these differences was not part of the original set of hypotheses and we do not report these additional tests. It is also possible that the statistical power needed to identify effects of each specific claim is larger than our current sample. Regarding effect sizes, it is important to mention that standard deviation is slightly<50%, thus the largest effect is 12% of a standard deviation.

6. Discussion

Typically, experiments should be assessed for their internal and external validity (Campbell and Stanley 1966; Cronbach 1983; Guala 2005). Internal validity is the likelihood with which a fact or mechanism has been correctly identified. External validity is the plausibility of generalizing an identified finding, usually referring to populations, treatments, observations, and settings. We will analyze both aspects of validity and point out strengths and limitations.

As we mentioned above, a study can be judged to be externally valid if the sample is representative, if the treatments are generalizable, if the setting of data collections does not bias the results, and if the outcome collected are parallel to the real-life behavior that we aim to capture. In all three studies, the samples have been randomly selected from an online panel and among those participants with recent experience in the purchase of FAP. All the studies cover different countries and geographical areas in Europe. The use of multiple countries is a strength of the study, although for budgetary reason the number of countries and the sample size varies across the three experiments. The panel is an opt in panel with a limited number of observables matched with the underlying population. It is possible that the population with limited access to internet is underrepresented, although all the experiments were accessible via cellphone. In the discussion of Study I, we provided a systematic assessment of our sample with respect to the one from a Eurobarometer study which claimed to be representative. We did find qualitatively the same characteristics and some imbalance (e.g. the share of respondents with children). A more precise quantitative assessment was not possible because of the differences in the questions used to self report socio-demographic characteristics.

In terms of experimental treatments, we were attentive in the design of labels, that were variations of existing labels and fully compliant with the existing legislation and the proposed changes embedded in new policy proposals. This is a strength of the study. However, in Study I, the use of externally valid labels poses some restrictions over the placement of different claims within the label. Since the mandatory claims are always on top, the placement cannot possibly generate our results on information overload. If anything, it should have reduced the impact of the additional information. But the placement reduces generalization with respect to the effect of the claims in other parts of the label. However, by construction, labels in the mandatory plus voluntary are larger, thus a full randomization of claims within the labels would not have been possible.

Other elements of the treatments may have been less generalizable, for example the timing conditions which may be more extreme in comparison to real life purchase but were motivated by the focus on internal validity. The online setting improved the response rate in the three experiments and was made compulsory by the pandemic. However, for study III, although online shopping is becoming more and more common, a real setting such as a supermarket (Becchetti, Salustri, and Scaramozzino 2020) would be an ideally externally valid test for the purchase of FAP. A field experiment in markets for fresh food would also fit the research question but was not viable. Another limitation was a product of the need to limit the varieties of fish in the final experiment, the inclusion of different FAPs would have provide a more complete picture.

Finally, one strength of the current study is that we elicit an array of outcomes and some of these outcomes (such as comprehension) in different settings. As a result, we are not overly concerned about measurement error and lack of generalizability, although the final word can only come from further studies. Nonetheless, one of the limitations is that we lack a measurement of a real purchase (e.g. in a supermarket), where more variables influence consumers' decisions.

In terms of internal validity, induced value (Smith 1982; 1976) and random assignment to treatments, in order to control for selection bias (Angrist and Pischke 2009) are benchmarks for causal inference. All three experiments complied with these basic procedures. While the online setting naturally induces a loss of control, in-person laboratory experimentation was not possible due to the pandemic. The use of data from a controlled laboratory experiment and from a field study would be a good complement to the current study. For experiment I, one could argue that the reduced time of exposure to stimuli may have favored the observability of the information overload. Nevertheless, this was a controlled environment where subjects knew the specific labels they had to memorize whereas in natural settings the additional time is counterbalanced by additional external stimuli that should be factored in. In study II, the recollection of both first order and second order beliefs of trustworthiness would have been preferable. In study III, there is no role of supply in the market: respondent knew that they were not interacting with real suppliers who may possibly have the intention to distort information (Bogliacino et al. 2022). As a result, our estimation of the information effect is probably a lower bound. Additional, controlling time of purchase (e.g. time delay versus time pressure) as a treatment may have provided additional input to policy.

If perfection is not taken to be the enemy of the good, the three experiments reported here offer significant issues for the competent regulatory authorities to consider for the projected European FAP labelling Directive.

7. Concluding remarks

This article presents the results of three experiments in a multi-country study on the design of labels for FAP, focusing on the provision of country-of-origin information.

This research was designed to inform European regulation. The existing legislation in Europe requires mandatory claims in the label, which differ across subtypes of FAP, such as farmed and caught products, processed and unprocessed, but allows sellers to include voluntary information. As outlined in the introduction, there is a tendency towards information overload as a result of increased consumer demand for FAP and the need for additional information regarding quality, environmental impact, and potential risks (Wilson and Lusk 2020). This demand-led shift puts sellers and the competent authorities under pressure. It makes label design a challenge as the requirement for more information conflicts with real world purchasing decisions where only limited information is processed.

To identify the optimal provision of country-of-origin information, the research comprised three studies. In the first study, respondents were incentivised to correctly recall information from a label. The results showed that very simple claims (e.g. 'catch area', 'farmed in') that are usually included as mandatory information are better understood, in comparison to other voluntary claims. This may be due to simplicity or the fact that consumers are more accustomed to them. However, increasing the information on the label decreased the accuracy in the recall, *ceteris paribus*. This is the so called *'less is more'* effect, showing that less information sometimes leads to more accurate consumer response. Evidence from study I points towards the possibility that, even though mandatory claims were relatively better recalled than voluntary claims, they are less accurately recalled than other (e.g. nutritional) claims that were included in the label.

While avoiding information overload emerged as a clear finding from study I, it was not sufficiently robust evidence to make a policy related recommendation. The reason being that differences in recall may overlook the *normative* evaluation of the label – the extent to which the information on the label meets the consumers' wishes, and the *accurate processing* of the information. Secondly, it can be argued that a controlled test of comprehension of the label (although an explicit objective of the legislation) is not enough to capture comprehension in an ecologically valid setting. To answer these research questions, two additional studies were conducted.

In study II, a label design task was conducted in four countries using a novel procedure aimed at eliciting shared beliefs about the trustworthiness of the attributes in a label. Respondents showed that farmed and

caught fish are perceived as different types of products: whereas for the latter, simple claims such as "catch area" is valued positively, in the former, composite claims (such as quality certifications) are found to be more trustworthy. Visual information such as flags are salient and emerge as trustworthy, but can also create confusion, as when the vessel flag is selected as trustworthy for the farmed fish. The lack of differences across two time conditions (10 and 60 s) suggests a stable level of discriminability among options (Krajbich et al. 2015).

In study III, an online platform for food delivery was programmed, in which participants were incentivized to correctly perform a purchasing task following set of instructions related to product categories and product types. To minimize experimenter demand effects, the FAP item was included in a list of other distractors. A between subject experimental design with increasing amounts of information overload was administered. The negative effect of too much information emerges again as the significant finding, with accuracy decreasing monotonically when additional claims are added to the label. The effect is small but robust. In line with previous experiment, the flag appears as misleading.

In conclusion, these results lead to the following recommendations. The literature points out that FAP origin is an important attribute. According to data from the Eurobarometer, origin is the third most important aspect for respondents when buying FAPs (41 % of respondents mention it as important aspects when buying FAPs), after only the product's appearance (59 %) and the cost of the product (52 %). Carlucci et al. (2015) show similar findings. However, the competent regulatory authorities should acknowledge that origin may be understood differently across consumers, and it is often used as a signal for quality and sustainability (Dekhili & d'Hauteville, 2009; Verbeke & Ward, 2006). It is well possible that consumers that process origin claims as informational cues coexist with consumers who use quality claims as a mental shortcut for evaluating quality or sustainability. For example, a country may evoke a tradition of cooking a certain type of fish, or may be associated with poor environmental record, or simply the distance from the consumer's home may evoke a judgment of freshness or a greater number of transportation emissions. Based on our finding, our first recommendation would be to increase saliency of country of origin information but also clarify that country has no nutritional or environmental implications.

The literature suggests there may exist an attitude-behavioral gap: consumer interested in origin may actually avoid checking labels (Wang et al., 2012). This effect is magnified because of the presence of information overload in labels, as we documented in the studies I and III, and the intrinsic difficulties of processing different claims, as shown in study I. Our second recommendation is to explicitly reduce the possibility of claims that may distract or confuse the information coming from mandatory claims, as this will likely reduces the risk of label avoidance.

Third, while the existence of information overload may apply across a variety of food categories, labels for FAP may, in particular, call for simplification. The evidence from experiment III is relevant here. Even in the control condition, the accuracy of following the choice selection criteria is below 50 %. In everyday life and with no instructions, the accuracy of delivering on consumers' preferences can be assumed to be far lower.

The fourth recommendation is that there are subtleties across type of FAPs that should be taken into account. The literature has established that visual cues attract attention and is more impactful than other features of labels and packaging. Experiments with eye-tracking technologies confirm this finding (Rramani et al. 2020) and argue in favor of the traffic light display of calories. In the context of FAP, the flag has a similar role as evidenced in study II, which found that flags invoke a shared perception of trustworthiness. It is notable that whereas in caught and unprocessed fish the flag can be useful, for the other categories the presence of a long value chain may be hidden behind the saliency of the flag.

Finally, the shared perception of trustworthiness of quality claims suggests that there is a space for independent auditing and verification, and of sanctioning against manipulation by companies with greenwashing and other illegitimate commercial practices. This risk can be significant especially given the strong evidence for the effect of information overload provided by study I and III. Future research should address this point.

Funding

European Commission, Behavioural Study on origin claims on fishery and aquaculture products (European Commission - Framework Contract with reopening of competition - Behavioural studies).

Comments by the Editor and four anonymous referees were very helpful during the reviewing process. We are grateful to Sabela Perez Mais, Gerd Heinen, Karolina Majewska, Antonios Proestakis for their comments and suggestions during the development of the project. All the remaining errors are ours. We thank Setecem for the programming of Study III and BdI-Schlesinger Group for the data collection. The content of this article represents the views of the authors and is its sole responsibility; it can in no way be taken to reflect the views of the European Commission and/or DG MARE or other body of the European Union. The European Commission and/or DG MARE do not guarantee the accuracy of the data included in this manuscript, nor do they accept responsibility for any use made by third parties thereof.

CRediT authorship contribution statement

Francesco Bogliacino: Overall Design of the study, Experimental Design, Analysis, Writing – original draft, Writing – Revision. **Rafael Charris:** Overall Design of the study, Experimental Design, Analysis, Writing – Revision. **Cristiano Codagnone:** Overall Design of the study, Writing – Revision. **Frans Folkvord:** Overall Design of the study, Writing – original draft, Writing – Revision. **George Gaskell:** Overall Design of the study, Writing – original draft, Writing – Revision. **Camilo Gómez:** Overall Design of the study, Experimental Design, Analysis, Writing – Revision. **Giovanni Liva:** Overall Design of the study, Experimental Design, Writing – Revision. **Felipe Montealegre:** Overall Design of the study, Experimental Design, Analysis, Writing – Revision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ajzen, I., 1991. The Theory of Planned Behavior. Organ. Behav. Hum. Decis. Process. 50 (2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Ajzen, I., Fishbein, M., 2005. The Influence of attitudes on behavior. In: D. Albarracín, B. Johnson, M. Zanna, The Handbook of Attitudes (pp. 173-222). Mahwah, NJ: Erlbaum.

Ampuero, O., Vila, N., 2006. Consumer perceptions of product packaging. Journal Of Consumer Marketing 23 (2), 100–112.

Angrist, J., Pischke, J.S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princenton University Press, Princeton.

Ankamah-Yeboah, I., Jacobsen, J.B., Olsen, S.B., Nielsen, M., Nielsen, R., 2019. The Impact of Animal Welfare and Environmental Information on the Choice of Organic Fish: An Empirical Investigation of German Trout Consumers. Mar. Resour. Econ. 34 (3), 248–266. https://doi.org/10.1086/705235.

Becchetti, L., Salustri, F., Scaramozzino, P., 2020. Nudging and Corporate Environmental Responsibility: A Natural Field Experiment. Food Policy 97 (July), 101951. https://doi.org/10.1016/j.foodpol.2020.101951.

Bertoletti, P., Etro, F., 2016. Preferences, Entry, and Market Structure. RAND J. Econ. 47 (4), 792–821. https://doi.org/10.1057/9780230226203.2478.

Bicchieri, C., 2016. Norms in the Wild: How to Diagnose, Measure, and Change Social Norms. Oxford University Press.

Birch, D., Lawley, M., Hamblin, D., 2012. Drivers and Barriers to Seafood Consumption in Australia. J. Consum. Mark. 29 (1), 64–73. https://doi.org/10.1108/07363761211193055.

Bogliacino, F., Charris, R., Codagnone, C., Folkvord, F., Montealegre, F., Lupiáñez-Villanueva, F., 2022. Unfair commercial practices in a pit market: Evidence from an artefactual field experiment. Behavioural Public Policy 1–18. https://doi.org/10.1017/bpp.2022.33.

Bogliacino, F., Codagnone, C., Veltri, G.A., 2015a. The Behavioural Turn in Consumer Policy: Perspectives and Clarifications. Intereconomics 50 (2), 108–114. https://doi.org/10.1007/s10272-015-0532-4.

Bogliacino, F., Codagnone, C., Veltri, G.A., Chakravarti, A., Ortoleva, P., Gaskell, G., Ivchenko, A., Lupianez-Villanueva, F., Mureddu, F., Rudisill, C., 2015b. Pathos & Ethos: Emotions and Willingness to Pay for Tobacco Products. PLoS One 10 (10), 1–25. https://doi.org/10.1371/journal.pone.0139542.

F. Bogliacino C. Cristiano G.A. Veltri An Introduction to the Special Issue on 'the Behavioural Turn in Public Policy: New Evidence from Experiments Econ. Polit. 33 3 2016 323 332 10.1007/s40888-016-0044-0.

Bordalo, P., Gennaioli, N., Shleifer, A., 2013. Salience and Consumer Choice Nicola Gennaioli Andrei Shleifer. J. Polit. Econ. 121 (5), 803–843.

Brécard, D., Hlaimi, B., Lucas, S., Perraudeau, Y., Salladarré, F., 2009. Determinants of Demand for Green Products: An Application to Eco-Label Demand for Fish in Europe. Ecol. Econ. 69 (1), 115–125. https://doi.org/10.1016/J.ECOLECON.2009.07.017.

Buijzen, M., Van Reijmersdal, E.A., Owen, L.H., 2010. Introducing the PCMC model: An investigative framework for young people's processing of commercialized media content. Commun. Theory 20 (4), 427–450.

Camerer, C., 1995. Individual Decision Making. In: Kagel, J.H., Roth, A.E. (Eds.), The Handbook of Experimental Economics. Princeton University Press, pp. 587–704. https://doi.org/10.2307/j.ctvzsmff5.12.

Campbell, D., Stanley, J.C., 1966. Experimental and Quasi-Experimental Designs for Research. Rand McNally and Co, Chicago.

Cardoso, C., Lourenço, H., Costa, S., Gonçalves, S., Nunes, M.L., 2013. Survey into the Seafood Consumption Preferences and Patterns in the Portuguese Population. Gender and Regional Variability. Appetite 64 (May), 20–31. https://doi.org/10.1016/J.APPET.2012.12.022.

Carlucci, D., Nocella, G., De Devitiis, B., Viscecchia, R., Bimbo, F., Nardone, G., 2015. Consumer Purchasing Behaviour towards Fish and Seafood Products. Patterns and Insights from a Sample of International Studies. Appetite 84, 212–227. https://doi.org/10.1016/j.appet.2014.10.008.

Caswell, J.A., Majduzka, E.M., 1996. Using Informational Labeling to Influence the Market for Quality in Food Products. Am. J. Agric. Econ. 78 (5), 1248–1253.

Chaiken, S., Trope, Y., (Eds.)., 1999. Dual-process theories in social psychology. Guilford PressClaret, Anna, Luis Guerrero, Enaitz Aguirre, Laura Rincón, M. Dolores Hernández, Inmaculada Martínez, José Benito Peleteiro, Amàlia Grau, and Carmen Rodríguez-Rodríguez. 2012. "Consumer Preferences for Sea Fish Using Conjoint Analysis: Exploratory Study of the Importance of Country of Origin, Obtaining Method, Storage Conditions and Purchasing Price." Food Quality and Preference 26 (2): 259–66. 10.1016/J.FOODQUAL.2012.05.006.

Claret, A., Guerrero, L., Ginés, R., Amàlia Grau, M., Hernández, D., Aguirre, E., Peleteiro, J.B., Fernández-Pato, C., Rodríguez-Rodríguez, C., 2014. Consumer Beliefs Regarding Farmed versus Wild Fish. Appetite 79 (August), 25–31. https://doi.org/10.1016/J.APPET.2014.03.031.

Codagnone, C., Alessandro Veltri, G., Lupiáñez-Villanueva, F., Bogliacino, F., 2014. The Challenges and Opportunities of 'Nudging. J. Epidemiol. Commun. Health 68 (10), 909–911. https://doi.org/10.1136/jech-2014-203948.

Codagnone, C., Veltri, G.A., Bogliacino, F., Lupiáñez-Villanueva, F., Gaskell, G., Ivchenko, A., Ortoleva, P., Mureddu, F., 2016. Labels as Nudges? An Experimental Study of Car Eco-Labels. Economia Politica 33 (3), 403–432. https://doi.org/10.1007/s40888-016-0042-2.

Codagnone, C., Liva, G., Folkvord, F., Puppo, F., Gaskell, F., Bogliacino, G., Olaya, A.F. M., 2021. Behavioural Study on Origin Claims on Fishery and Aquaculture Products. Brussels. https://doi.org/10.2771/846769.

Conner, M., 2001. Health Behaviors. Int. Encyclop. Soc. Behav. Sci. 6506–12 https://doi.org/10.1016/B0-08-043076-7/03871-7.

Cowburn, G., Stockley, L., 2005. Consumer Understanding and Use of Nutrition Labelling: A Systematic Review. Public Health Nutr. 8 (1), 21–28.

Cronbach, L.J., 1983. In: "Course Improvement Through Evaluation". In Evaluation Models. Springer, Dordrecht, pp. 101–115.

Crosetto, Paolo, Gaudeul, Alexia, 2017. Choosing Not to Compete: Can Firms Maintain High Prices by Confusing Consumers? J. Econ. Manage. Strat. 26 (4), 897–922. https://doi.org/10.1111/jems.12212.

d'Adda, G., Drouvelis, M., Nosenzo, D., 2016. Norm Elicitation in Within-subject Designs: Testing for Order E!ects. J. Behav. Exp. Econ. 62, 1–7.

Dekhili, S., d'Hauteville, F., 2009. Effect of the region of origin on the perceived quality of olive oil: An experimental approach using a control group. Food Qual. Prefer. 20 (7), 525–532. https://doi.org/10.1016/j.foodqual.2009.05.008. ISSN 0950-3293.

Eagly, A.H., Chaiken, S., 1993. The psychology of attitudes. Harcourt Brace Jovanovich College Publishers.

Enax, L., Krapp, V., Piehl, A., Weber, B., 2015. Effects of social sustainability signaling on neural valuation signals and taste-experience of food products. Front Behav Neurosci 9, 247. https://doi.org/10.3389/fnbeh.2015.00247.

Enax, L., Krajbich, I., Weber, B., 2016. Salient nutrition labels increase the integration of health attributes in food decision-making. Judgm Decis Mak 11, 460–471.

Eppler, M.J., Mengis, J., 2004. The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. Inf. Soc. 20 (5), 325–344. https://doi.org/10.1080/01972240490507974.

Erkut, H., 2020. Incentivized Measurement of Social Norms Using Coordination Games. Analyse Kritik 42 (1), 97–106.

Eurobarometer, 2018. 89.3: Europeans' perceptions of the Schengen Area, EU consumer habits regarding fishery and aquaculture products, and EU citizens and development cooperation.

European Commission, 2006. europa.eu. Retrieved 28 March 2022, from https://ec.europa.eu/food/safety/labelling_nutrition/claims/nutrition_claims_en.

Fallucchi, Francesco, Daniele Nosenzo, 2020. The Coordinating Power of Social Norms. Aarhus Economics Working Papers. Department of Economics and Business Economics, Aarhus University. https://ideas.repec.org/p/aah/aarhec/2020-06.html.

FAO/WHO, 2011. Report of the Joint FAO/WHO Expert Consultation on the Risks and Benefits of Fish Consumption. Rome, Geneva.

Fay, M., 2003. A 50-year longitudinal study of changes in the content and form of food advertising in New Zealand magazines. Int. J. Advert. 22 (1), 67–91.

Fehr, E., 2009. On the Economics and Biology of Trust. J. Eur. Econ. Assoc. 7, 235–266.

Folkvord, F., Anschütz, D., Boyland, E., Kelly, B., Buijzen, M., 2016. Food advertising and eating behaviour in children. Appetite 107, 681.

Folkvord, F., Anastasiadou, D.T., Anschütz, D., 2017. Memorizing fruit: The effect of a fruit memory- game on children's fruit intake. Prev. Med. Rep. 5, 106–111.

Folkvord, F., Codagnone, C., Bogliacino, F., Veltri, G., Lupiañez-Villanueva, F., Ivchenko, A., Gaskell, G., 2019. Experimental Evidence on Measures to Protect Consumers of Online Gambling Services. J. Behav. Econ. Policy 3 (1), 20–29.

Garretson, J.A., Burton, S., 2000. Effects of Nutrition Facts Panel Values, Nutrition Claims, and Health Claims on Consumer Attitudes, Perceptions of Disease-Related Risks and Trust. J. Public Policy Mark. 19 (2), 213–227. https://doi.org/10.1509/jppm.19.2.213.17133.

Goerges, L., Nosenzo, D., 2020. Measuring Social Norms in Economics: Why It Is Important and How It Is Done. Analyse Kritik 42 (2), 285–311.

Gomila, R., 2021. Logistic or Linear? Estimating Causal Effects of Experimental Treatments on Binary Outcomes Using Regression Analysis. J. Exp. Psychol. Gen. 150 (4), 700–709.

Gravel, K., Doucet, E., Herman, C.P., Pomerleau, S., Bourlaud, A.S., Provencher, V., 2012. "Healthy","diet", or "hedonic". How nutrition claims affect food-related perceptions and intake? Appetite 59 (3), 877–884.

Grunert, K.G., Hieke, S., Wills, J., 2014. Sustainability Labels on Food Products: Consumer Motivation, Understanding and Use. Food Policy 44 (February), 177–189. https://doi.org/10.1016/J.FOODPOL.2013.12.001.

Gu, Y., Wenzel, T., 2015. Putting on a Tight Leash and Levelling Playing Field: An Experiment in Strategic Obfuscation and Consumer Protection, 120–28 Int. J. Ind Organiz 42. https://doi.org/10.1016/j.ijindorg.2015.07.008.

Gu, Y., Wenzel, T., 2020. Curbing Obfuscation: Empower Consumers or Regulate Firms? Int. J. Ind. Organiz. 70, 102582 https://doi.org/10.1016/j.ijindorg.2020.102582.

Guala, F., 2005. The Methodology of Experimental Economics. History. Cambridge University Press, Cambridge and London.

Hansen, H., Sallis, J., 2011. Extrinsic Cues and Consumer Judgments of Food Product Introductions: The Case of Pangasius in Norway. J. Food Prod. Mark. 17 (5), 536–551. https://doi.org/10.1080/10454446.2011.618791.

Hare, T.A., Camerer, C.F., Rangel, A., 2009. Self-control in decision-making involves modulation of the vmPFC valuation system. Science 324, 646–648. https://doi.org/10.1126/science.1168450.

Hare, T.A., Malmaud, J., Rangel, A., 2011. Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice. J. Neurosci. 31, 11077–11087. https://doi.org/10.1523/JNEUROSCI.6383-10.2011.

Hinkes, C., Schulze-Ehlers, B., 2018. Consumer Attitudes and Preferences towards Pangasius and Tilapia: The Role of Sustainability Certification and the Country of Origin. Appetite 127 (August), 171–181. https://doi.org/10.1016/J.APPET.2018.05.001.

Houser, D., Xiao, E., 2011. Classification of Natural Language Messages Using a Coordination Game. Exp. Econ. 14 (1), 1–14.

Jeffery, R.W., 2004. How can Health Behavior Theory be made more useful for intervention research? Int. J. Behav. Nutr. Phys. Act. 1 (1), 10.

Juhl, H.J., Poulsen, C.S., 2000. Antecedents and Effects of Consumer Involvement in Fish as a Product Group, 261–67 Appetite 34 (3). https://doi.org/10.1006/APPE.2000.0318.

Kahneman, D., 2011. Thinking Fast and Slow. Farrar, Straus and Giroux, New York City.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk, 293–98 Econometrica 47 (2). https://doi.org/10.2307/j.ctv1kr4n03.21.

Kalaycı, K., 2016. Confusopoly: Competition and Obfuscation in Markets. Exp. Econ. 19 (2), 299–316. https://doi.org/10.1007/s10683-015-9438-z.

Kalayci, K., Potters, J., 2011. Buyer Confusion and Market Prices. Int. J. Ind Organiz 29 (1), 14–22. https://doi.org/10.1016/j.ijindorg.2010.06.004.

Karnal, N., Machiels, C., Orth, U., Mai, R., 2016. Healthy by design, but only when in focus: Communicating non-verbal health cues through symbolic meaning in packaging. Food Qual. Pref. 52, 106–119.

Kim, K., Cheong, Y., Zheng, L., 2009. The current practices in food advertising. Int. J. Advert. 28 (3), 527–553.

Klassen, M., Wauer, S., Cassel, S., 1991. Increases In Health And Weight Loss Claims In Food Advertising In The Eighties. J. Advert. Res. 30 (6), 32–37.

König-Kersting, Christian., 2021. On the Robustness of Social Norm Elicitation. https://ideas.repec.org/p/inn/wpaper/2021-02.html.

Krajbich, I., Bartling, B., Hare, T., Fehr, E., 2015. Rethinking Fast and Slow Based on a Critique of Reaction-Time Reverse Inference. Nat. Commun. 6 (July), 1–9. https://doi.org/10.1038/ncomms8455.

Krupka, E.L., Weber, R.A., 2013. Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary? J. Eur. Econ. Assoc. 11 (3), 495–524. https://doi.org/10.1111/jeea.12006.

Lawley, M., Birch, D., Hamblin, D., 2012. An Exploratory Study into the Role and Interplay of Intrinsic and Extrinsic Cues in Australian Consumers' Evaluations of Fish, 260–67 Australas. Mark. J. 20 (4). https://doi.org/10.1016/j.ausmj.2012.05.014.

Mani, A., Mullanaithan, S., Shafir, E., Zhao, J., 2014. Poverty impedes cognitive function. Science 341 (6149).

Masterson, T.D., Bermudez, M.A., Austen, M., Lundquist, E., Bruce, A.S., Pearce, A.L., Keller, K.L., 2018. Food commercials do not affect energy intake in a laboratory meal but do alter brain responses to visual food cues in children. Appetite.
Mehta, J., Starmer, C., Sugden, R., 1994. The Nature of Salience: An Experimental Investigation of Pure Coordination Games. Am. Econ. Rev. 84 (3), 658–673.
Menozzi, D., Nguyen, T.T., Sogari, G., Taskov, D., Lucas, S., Castro-Rial, J.L.S., Mora, C., 2020. Consumers' Preferences and Willingness to Pay for Fish Products with Health and Environmental Labels: Evidence from Five European Countries. Nutrients 12 (9). https://doi.org/10.3390/nu12092650.
Nguyen, T.T., Haider, W., Solgaard, H.S., Ravn-Jonsen, L., Roth, E., 2015. Consumer Willingness to Pay for Quality Attributes of Fresh Seafood: A Labeled Latent Class Model. Food Qual. Prefer. 41 (April), 225–236. https://doi.org/10.1016/J.FOODQUAL.2014.12.007.
O'Reilly, I.I.I., Charles, A., 1980. Individuals and Information Overload in Organizations: Is More Necessarily Better? Acad. Manag. J. 23 (4), 684–696. https://doi.org/10.1016/0024-6301(81)90222-3.
Petty, R., Cacioppo, J., 1996. Attitudes and persuasion: Classic and contemporary approaches, new ed. Westview Press, Boulder, CO.
Petty, R., Cacioppo, J., Strathmann, A., Priester, J., 2005. To think or not to think: Exploring two routes to persuasion. In: Brock, T., Green, M. (Eds.), Persuasion: Psychological Insights and Perspectives. Sage, Thousand Oaks, CA, pp. 81–116.
Rangel, A., 2013. Regulation of dietary choice by the decision-making circuitry. Nat. Neurosci. 16 (12), 1717–1724.
Reynolds, C.J., Buckley, J.D., Weinstein, P., Boland, J., 2014. Are the Dietary Guidelines for Meat, Fat, Fruit and Vegetable Consumption Appropriate for Environmental Sustainability? A Review of the Literature. Nutrients 6 (6), 2251–2265. https://doi.org/10.3390/nu6062251.
Roe, B., Levy, A.S., Derby, B.M., 1999. The Impact of Health Claims on Consumer Search and Product Evaluation Outcomes: Results from FDA Experimental Data. J. Public Policy Mark. 18 (1), 89–105. https://doi.org/10.1177/074391569901800110.
Rramani, Q., Krajbich, I., Enax, L., Brustkern, L., Weber, B., 2020. Salient Nutrition Labels Shift Peoples' Attention to Healthy Foods and Exert More Influence on Their Choices. Nutr. Res. 80, 106–116. https://doi.org/10.1016/j.nutres.2020.06.013.
Sanjari, S.S., Jahn, S., Boztug, Y., 2017. Dual-Process Theory and Consumer Response to Front-of-Package Nutrition Label Formats. Nutr. Rev. 75 (11), 871–882. https://doi.org/10.1093/nutrit/nux043.
Shafir, E, ed. 2013. The Behavioral Foundations of Public Policy.
Silayoi, P., Speece, M., 2004. Packaging and purchase decisions. Br. Food J. 106 (8), 607–628.
Smith, V.L., 1976. Experimental Economics: Induced Value Theory. Am. Econ. Rev. 274 https://doi.org/10.1017/cbo9780511528354.008.
Smith, V.L., 1982. Microeconomic Systems as an Experimental Science. Am. Econ. Rev. 72 (5), 923–955. https://doi.org/10.1017/cbo9780511528354.018.
Sørensen, K., Van der Broukle, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., Brand, H., 2012. Health Literacy and Public Health: A Systematic Review and Integration of Definitions and Models. BMC Public Health 12, 80. https://doi.org/10.1017/dmp.2020.100.
Spielvogel, I., Matthes, J., Naderer, B., Karsay, K., 2018. A treat for the eyes. An eyetracking study on children's attention to unhealthy and healthy food cues in media content. Appetite 125, 63–71.
Teisl, M.F., 2003. What We May Have Is a Failure to Communicate*: Labeling Environmentally Certified Forest Products. For. Sci. 49 (5), 668–680. https://doi.org/10.4324/9781351154529-28.
Thorsdottir, F., Sveinsdottir, K., Jonsson, F.H., Einarsdottir, G., Thorsdottir, I., Martinsdottir, E., 2012. A Model of Fish Consumption among Young Consumers. J. Consum. Mark. 29 (1), 4–12. https://doi.org/10.1108/07363761211193000.
Tomić, M., Matulić, D., Jelić, M., 2016. What Determines Fresh Fish Consumption in Croatia? Appetite 106 (November), 13–22. https://doi.org/10.1016/J.APPET.2015.12.019.
Trondsen, T., Braaten, T., Lund, E., Eggen, A.E., 2004a. Health and Seafood Consumption Patterns among Women Aged 45–69 Years. A Norwegian Seafood Consumption Study. Food Qual. Prefer. 15 (2), 117–128. https://doi.org/10.1016/S0950-3293(03)00038-7.
Trondsen, T., Braaten, T., Lund, E., Eggen, A.E., 2004b. Consumption of Seafood—the Influence of Overweight and Health Beliefs. Food Qual. Prefer. 15 (4), 361–374. https://doi.org/10.1016/S0950-3293(03)00083-1.
Tversky, A., Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases. Science 185 (4157), 11124–11131. https://doi.org/10.4324/9781912282562.
Uchida, H., Onozaka, Y., Morita, T., Managi, S., 2014. Demand for Ecolabeled Seafood in the Japanese Market: A Conjoint Analysis of the Impact of Information and Interaction with Other Labels. Food Policy 44 (February), 68–76. https://doi.org/10.1016/J.FOODPOL.2013.10.002.
Underwood, R.L., Klein, N.M., Burke, R.R., 2001. Packaging communication: attentional effects of product imagery. J. Prod. Brand Manag. 10 (7), 403–422.
van Rompay, T., Deterink, F., Fenko, A., 2016. Healthy package, healthy product? Effects of packaging design as a function of purchase setting. Food Qual. Prefer. 53, 84–89.
Vanhonacker, F., Altintzoglou, T., Luten, J., Verbeke, W., 2011. Does Fish Origin Matter to European Consumers? Br. Food J. 113 (4), 535–549. https://doi.org/10.1108/00070701111124005.
Verbeke, W., Ward, R.W., 2006. Consumer Interest in Information Cues Denoting Quality, Traceability and Origin: An Application of Ordered Probit Models to Beef Labels. Food Qual. Prefer. 17 (6), 453–467. https://doi.org/10.1016/J.FOODQUAL.2005.05.010.
Verbeke, W., 2008. Impact of Communication on Consumers' Food Choices. In: Proceedings of the Nutrition Society 67 (3): 281–88. 10.1017/S0029665108007179.
Veselý, Š., 2015. Elicitation of Normative and Fairness Judgments: Do Incentives Matter? In: Judgment and Decision Making 10 (2), 191–197.
Wang, W., Jin, Y., Yao, S., 2012. Exploring attitude-behavior gap in sustainable consumption: based on regulatory focus theory. Ecol. Econ. 9, 55–60.
Williams, P., 2005. Consumer Understanding and Use of Health Claims for Foods. Nutr. Rev. 63 (7), 256–264. https://doi.org/10.1111/j.1753-4887.2005.tb00382.x. PMID: 16121480.
Wilson, L., Lusk, J.L., 2020. Consumer Willingness to Pay for Redundant Food Labels. Food Policy 97 (May), 101938. https://doi.org/10.1016/j.foodpol.2020.101938.
Wolgast, A., Tandler, N., Harrison, L., Umlauft, S., 2020. Adults' Dispositional and Situational Perspective-Taking: a Systematic Review. Educ. Psychol. Rev. 32, 353–389. https://doi.org/10.1007/s10648-019-09507-y.
Zizzo, D.J., 2010. Experimenter Demand Effects in Economic Experiments. Exp. Econ. 13 (1), 75–98.