What to communicate: selecting the information content to increase the demand for food products

RESEARCH ARTICLE

Elisa Garrido-Castro\textsuperscript{a}\textsuperscript{a} and Francisco José Torres-Ruiz\textsuperscript{b}

\textsuperscript{a}PhD student, Business, Marketing and Sociology Department, Faculty of Social and Law Sciences, University of Jaén, Campus Las Lagunillas, 23071 Jaén, Spain

\textsuperscript{b}PhD, associate professor, Business, Marketing and Sociology Department, Faculty of Social and Law Sciences, Centre for Advanced Studies in Olive Groves and Olive Oil, University of Jaén, Campus Las Lagunillas, 23071 Jaén, Spain

Abstract

Studies about how to convey a message through a communication campaign abound, but another important aspect is what to communicate, in other words, selecting the information content of the campaign. In many situations where the degree of knowledge is directly related to consumption of a food product, choosing what to communicate is crucial. The present study proposes a model to decide what the consumer needs to know in order to take the decision to consume. The model is based on a Qualitative Comparative Analysis method with some subsequent conversions. It was applied to a real situation: selecting the information content of a communication campaign to boost the consumption of virgin olive oils in the Spanish market. The main findings suggest that campaigns should inform about virgin olive oils are the highest quality and most healthy, that there is a basic difference between olive oil and virgin olive oil in that olive oil is a mixture of virgin and refined olive oils, and that they are both equally fattening.

Keywords: communication campaigns, information content, Qualitative Comparative Analysis, virgin olive oils

JEL code: D83, M31, M37

\textsuperscript{a}Corresponding author: egcastro@ujaen.es
1. Introduction

Nowadays people are living in a context of sophistication in which the purchase of food has soared in popularity. Consumers are more concerned about the relationship between diet and health and the hedonic role of food products. In addition to this growing concern and involvement, the range of products and information in the market has also raised owing to the globalisation process.

Indeed, the proliferation of new communication channels and television shows linked to diet, cooking, health and nutrition has implied that many people act as opinion leaders or influencers (bloggers, youtubers, etc.). However, this abundance of available information (not always correct) could involve consumer’s confusion and lack of real knowledge given that clichés and wrong ideas would be introduced in the mind of consumers. In this context, removing wrong beliefs and achieving a minimum level of knowledge in some food products could be crucial in order to manage effective communication campaigns. This minimum level of knowledge can be summarised as follows: what to communicate (information content of the message).

This aspect constitutes the core interest of the present paper. In short, the key question on which this paper is based is: what do consumers need to know in order to take the decision to consume a product (or to consume more of it)? This question may be particularly useful when conducting generic promotion campaigns to boost the consumption of a food product or the adoption of certain forms of behaviour, especially when the degree of knowledge the consumer has about the product is linked to consumption of that product (Landström et al., 2007; Tuorila and Cardello, 2002; Urala and Lähteenmäki, 2004; Urala, 2005; Wansink et al., 2005; Zhu et al., 2018) or when confusion in the market leads to wrong choices.

Nevertheless, it is not easy to choose which specific contents to introduce into the mind of the consumer when so many different aspects can be communicated. In this situation the literature on communication, from advertising campaigns to labelling, shows that it is impractical to try to teach consumers everything there is to know about the product. A lack of involvement and interest, an inability to assimilate technical information, the pressure and overload of information in society, people’s limitations as information processors and a shortage of resources in organisations are factors that make it inadvisable to saturate consumers with an excess of information that they are often neither able nor motivated to process. Moreover, not all the knowledge acquired by the consumer has the same impact on the purchase decision, as some information on the product and the results of consuming it are key factors, while other information has practically no effect. In short, identifying the minimum key information that consumers must have in order to change their behaviour can be very important in this type of campaign.

Nevertheless, it is not easy to choose which specific contents to introduce into the mind of the consumer when so many different aspects can be communicated. In this situation the literature on communication, from advertising campaigns to labelling, shows that it is impractical to try to teach consumers everything there is to know about the product. A lack of involvement and interest (Beharrell and Denison, 1995; Tanner and Raymond, 2016), an inability to assimilate technical information, the pressure and overload of information in society, people’s limitations as information processors (Dunbar, 2010; Hall and Osses, 2013; Jacoby et al., 1977; Loebnitz et al., 2015; Reutskaja et al., 2011; Scheibehenne et al., 2007; Sørensen et al., 2012; Wobker et al., 2015) and a shortage of resources in organisations are factors that make it inadvisable to saturate consumers with an excess of information that they are often neither able nor motivated to process. Moreover, not all the knowledge acquired by the consumer has the same impact on the purchase decision, as some information on the product and the results of consuming it are key factors, while other information has practically no effect (e.g. consumer who is concerned about his weight will pay greater attention to the information about product’s calories, while other person who do not have this worry will select other key information to take the purchase decision). In short, identifying the minimum key information that consumers must have in order to change their behaviour can be very important in this type of campaign.
However, the dimension of what to communicate has not been addressed by the literature. The research on managing communication campaigns has centred on how to communicate, both from the point of view of the most suitable media and channels for the message to reach the consumer and from that of achieving the highest impact (eye-catching, persuasive, etc.) in order to obtain the desired change in behaviour. All of this is defined by Tellis (2004) as ‘behaviour paradigm in advertising’. Consequently, there is an abundance of papers, theories, models and software regarding how to communicate the message in the most effective and efficient way possible within the time and budgetary restraints (Bartsch and Hartmann, 2017; Berger and Mitchell, 1989; Castronovo and Huang, 2012; Chang et al., 2015; Cline et al., 2003; Danaher and Rossiter, 2011; Darley and Smith, 1993; Eisend and Tarrahi, 2016; Feiereisen et al., 2009; Iyer et al., 2005; Jack et al., 2005; Li and Miniard, 2006; McKay-Nesbitt et al., 2013; Robson and Robinson, 2013; Stern, 1994; Steward and Pavlou, 2002; Vakratsas and Ambler, 1999; Wolin and Korgaonkar, 2003; Yang, 2012; Zhang and Buda, 1999). Despite this research has examined the impact of many variables in the consumer behaviour such as sidedness, comparisons, humour, popularity, attractiveness, framed messages, gender, social media, credibility, trust or source, etc.; it has not been proposed any method to select the information to be included in a message in order to achieve the most effective communication.

This paper proposes a new method for choosing the information content of communication campaigns by selecting a set of blocks of information, related to consumption of the food product, that are based on parsimony (minimum information content) and effectiveness, in the sense of achieving the desired behaviour (efficacy) with the least possible effort (efficiency). This method is based on analysing the connection between combinations of contents or blocks of information and consumption of the food product (or the behaviour to be achieved) in a large sample of consumers. The procedure bears a certain similarity to applications of Qualitative Comparative Analysis (QCA) models (Ragin, 1987, 2000), which makes it possible to employ the initial assumptions, procedures, indicators and algorithms of these models (with a few conversions).

Traditionally, QCA models have been used in disciplines such as political science or sociology (Hellström, 2001; Krook 2010; Miethe and Drass, 1999; Ragin, 1994; Redding and Viterna, 1999). They were conceived as a ‘macro-comparative’ approach (Berg-Schlosser and Quenter, 1996). However, with the passage of time, QCA models have been successfully applied in other disciplines (e.g. business management, Bernal-Jurado et al., 2017, 2018; Duarte et al., 2012; Fiss, 2011; Greckhamer et al., 2008; Huarng and Yu, 2015; Ordanini et al., 2014).

The first part of this paper sets out the problem of selecting combinations of contents or blocks of information (knowledge to be achieved) to influence changes in the consumers’ behaviour, taking Crisp-Set QCA (csQCA) as the frame of reference. Following a preliminary study of csQCA suitability for addressing the problem, certain conversions were proposed to fit it to the purpose for which this method was designed (identifying cause and effect relationships) and given that there has been no case of its being applied to choosing information contents. In a second stage, this new procedure or adaptation was tested on a real case: choosing the information contents to be included in a communication campaign to increase the consumption of virgin olive oils in Spain.

### 2. The problem of selecting the knowledge to be communicated, using a Qualitative Comparative Analysis approach

The proposed model for selecting the information content is based on a series of starting points or initial hypotheses: (1) consumption of a food product is directly related to the degree of objective knowledge about it; (2) knowledge can be specified and measured through a set of specific, relevant blocks of information related to the consumer’s decision to consume or not consume the food product; (3) consequently, the decision to consume depends on the consumer’s overall configuration of blocks of information (knowledge) and acting on this overall configuration can alter the consumer’s response; (4) the problem or goal is to find the combination or combinations of pieces of information that give rise to a particular response (consumption). Among the possible combinations, the aim is to achieve maximum impact (the greatest probability of a
change in behaviour by the greatest number of people) while communicating the fewest possible, in other words, on the basis of efficacy and efficiency.

All this could be expressed as follows: let N be members of the target population such that N+ consume a product and N- do not; and let there be a set of consumption-related significant pieces of information (SPIs) \((A, B, C, D, E, F, G,...)\) in which each SPI may adopt one of two values \((A \text{ or } a; B \text{ or } b, \text{ etc.})\), depending on whether the consumer knows them or whether he or she is mistaken about them or does not know them. In order for an SPI to be operative, it must meet three conditions: (1) be related to objective knowledge; (2) be processable by the consumer; and (3) be related to the consumption or behaviour being studied. Regarding the first of these conditions, there must be no ambiguity in its interpretation, in the sense that there must be enough scientific or logical evidence for it to be indisputable. The idea is that consumers who assimilate or learn this piece of information will increase their knowledge of the food product. In relation to the second condition, interpreting the piece of information communicated must not require any expert knowledge or special training. For this reason, excessively technical or complicated language must not be used. The third condition is related to usefulness to the consumer, on the assumption that the information will be relevant if it is directly related to the consumer’s problems, needs and wishes. Two examples of SPIs could be ‘salt is bad for your blood pressure’ or ‘oily fish are good for your heart’. It is easy to see that this information increases knowledge of the food product, is easy to learn and could change the behaviour of many consumers who are concerned about their health or have cardiovascular problems.

The overall objective is to find combinations of SPIs that are strongly associated with consumption of the food product (those who have this knowledge consume it) and are only minimally present among the non-consumer group. In this way, sending a message that attempts to introduce the set of knowledge with these characteristics into the consumers’ minds will be particularly effective, given its persuasiveness (strong association with the desired result) and the number of people who do not possess this body of knowledge but could possess it if exposed to the message.

It is important to note that in order to be effective, the number of SPIs communicated must be as small as possible, as: (1) it would be unrealistic to expect consumers to be experts on the product; (2) for financial and operative reasons, not everything can be communicated; and (3) an information overload can hinder the apprehension and reception of the message. Parsimony is therefore an aspect that needs to be borne in mind.

It is also important to highlight that not all SPIs or their combinations have the same impact and that there may be effects due to interactions between them. Consequently, the aim is not to build up cognitive content in the consumer’s mind but to select the key content to communicate, the information that is really important for achieving the desired impact. Bearing these considerations in mind, one possibility would be to examine all the possible SPI combinations, relate them to the outcome variable and analyse which are most effective. However, when the possible SPIs are numerous, the number of possible combinations is excessive \((2^n; n=\text{number of SPIs})\) and is often greater than the sample size. In this situation it is useful to be able to use a procedure to simplify the problem and, at the same time, identify the key or truly relevant SPIs for analysis. In such cases, csQCA models can provide a frame of reference, as by simplifying and reducing the number of possible configurations, they attempt to identify the different combinations of dichotomous variables that explain or cause an outcome. However, this depends on parallels or correspondence between the theoretical characteristics of the models and the praxis of the relationship between knowledge and consumption.

The essence of QCA\(^1\) is identifying causal relationships. Their underlying logic is based on the postulates of John Stuart Mill, particularly the method of agreement and the method of difference. These methods are the basis for identifying the causal factors of complex phenomenon through systematic comparison of cases where it is seeking the set of variables (potential causes) which produce the same outcome or the set of variables (potential causes) which produce different outcomes. Each of these cases is considered a ‘configuration’

\(^1\) For a general description of the different types of QCA and their basis, uses, etc. see Rihoux and Ragin (2009) and Schneider and Wagemann (2012).
and the aim is to choose the set of configurations that produce the desired effect and therefore are its causes. These configurations are subjected to a Boolean simplification process, which makes it possible to reduce both the number of them and their complexity (shorter set/combinations of variables). These variables will constitute ‘sufficient’ conditions for the outcome studied to obtain, in other words, whenever they occur, so does the result, although there may be others that also cause it. For example, if a combination ABCDE gives a result and another combination ABCDe gives the same result, the solution can be simplified by stating that in order to achieve the result, only ABCD need to occur simultaneously.

The basis for QCA models is the assumption of multiple conjunctural causality (Rihoux and Ragin, 2009), which, rather than identifying a single model that fits the data well, emphasises the need to identify all the possible models that have explanatory potential. Underlying multiple conjunctural causality is the consideration that different combinations of factors may achieve the same result and that a variable may or may not be a cause depending on the context, alluding directly to the interactive nature of causal relationships. For example, AB or aC may cause effect E. In this case there are two (multiple) causal configurations and in both, furthermore, presence (A) or absence (a) is a cause depending on the context (conjuncture), i.e. on the other variables with which they interact. For example, thinking that virgin olive oils are healthy (A) is one of the many alternative causes that may lead to their consumption, as it is also possible for them not to be thought healthy (a) but considered to give the hair a better appearance (C).

Accepting the principle of multiple conjunctural causality means not accepting a large part of the principles that underlie classic statistical modelling approaches (regression, ANOVA, etc.), such as permanent causality (no SPI will always have the same impact, this may depend on the degree of knowledge in the market), incrementality (individual SPIs do not have a separate effect independently of the rest), or symmetry (the presence or absence of consumption may require different explanations). In the same way, owing to its effect on SPI analysis, it should be emphasised that the principle of uniformity is not accepted either, which means that in some cases an SPI can act (in conjunction with other SPIs) to favour a certain result and in other cases to prevent it. Consequently, the SPI that ‘olive oil is a blend of virgin and refined oils’ (A) could have a positive effect on virgin olive oil consumption if at the same time it is also thought that ‘refined oils are of worse quality’ (B) and/or that ‘refined oils are less healthy’ (C) (ABC → consumption). However, if refined oils are thought to be of better quality (b) and/or healthier (c), this could affect the consumption of virgin olive oil negatively (Abc → non-consumption). If there is no knowledge of the latter two, the SPI (A) might have no effect on consumption, so both forms (A and a) of this variable would have to be eliminated from the potential explanatory configurations.

At all events, it is worth highlighting that the knowledge-consumption relationship could fit interactive models better than additive ones, where analysis of the effects of a separate or isolated SPI could be very limited. A superficial approach to the SPI selection context, attempting to choose those that individually have more effect on consumption (e.g. using bivariate analysis to relate each SPI to the fact of consumption or non-consumption) would ignore the possibility that certain SPIs may only affect consumption in the presence of others or in conjunction with others, or may even affect it differently depending on the values of other SPIs.

3. Proposed analytical procedure based on adapted Qualitative Comparative Analysis

A csQCA approach can be applied to identifying sets, combinations or configurations of consumption-related SPIs, considering that these can be expressed as dichotomous variables according to whether or not the consumer knows them. For operational purposes, the initial generation and selection of SPIs is crucial. The SPIs must meet the requirements already mentioned, namely being related to objective aspects of knowledge, being related to product consumption and being understood by consumers (e.g. not excessively technical). For this, the researcher’s prior knowledge and, particularly, the involvement of experts can be very important. The SPIs can be applied on a scale of true, false or don’t know/no answer and then recoded dichotomously according to whether the respondents knew the SPI or answered correctly. Each consumer
would have a particular knowledge pattern or configuration and the result would be a dichotomous variable (e.g. consumption/non-consumption).

Initially, applying csQCA models could identify sufficient combinations, in other words, sets of SPIs that are interpreted as constituting the minimum knowledge that achieves the desired result (consumption +). For this purpose, the consistency of the combinations is considered the key indicator for discovering which sets of SPIs are most effective. Consistency is the degree to which the empirical evidence is consistent with the theoretical question (Ragin, 2006), in other words, the number of cases that are positive for the combination (value=1 for the combination and value=1 for the outcome, called $X^+$) out of the total number of cases that possess the combination (in which the outcome value could be 0 or 1, expressed as the sum of $X^-$ and $X^+$). Thus, when a combination is $X$ and $X^+$ is the number of people who present this combination of SPIs and the desired result, and $X^-$ the number of people who present this combination of SPIs and not the desired result, the consistency of a combination (Equation 1) gives an indication of its efficacy. Consequently, combinations with high consistency values are more persuasive.

$$\text{Consistency} = \frac{X^+}{(X^+ + X^-)} \quad (1)$$

Another important dimension for choosing the most suitable configurations is efficiency, which in this context means that when a message (a chosen combination $X$) is sent, the number of non-consumers that do not possess this combination of beliefs, namely $\sim X \& (-)$ or $(\sim X^-)$, is as large as possible. In this way, sending out an effective message attempts to bring about a change in behaviour in the greatest possible number of people. In an ideal situation (all non-consumers receive the message) the number of changes in behaviour will be the result of multiplying the consistency by $\sim X \& (-)$. This criterion is the key factor in selecting the most suitable messages. Figure 1 shows how sending out a message with combination X of SPIs will have more impact than combination Y and will therefore be better, as: (1) more non-consumers do not have this SPI configuration ($\sim X \& (-) > \sim Y \& (-)$) or $\sim X^- > \sim Y^-)$; and (2) it is more efficient since the proportion who consume the product when they have this combination of SPIs is higher, it can be expressed as:

$$[X^+/(X^+ + X^-)] > [Y^+/(Y^+ + Y^-)] \quad (2)$$

The initial SPI combinations can be obtained as direct solutions from the csQCA model, as they relate combinations of dichotomous variables to a particular outcome and make it possible to summarise and select the most suitable combinations through a process of simplification, which allows a large number of configurations to be processed.

![Figure 1. Comparison of combinations X and Y of significant pieces of information.](chart1.png)
However, certain adjustments have to be made to the proposed solutions in order to adapt them to the specific problems of selecting SPI configurations. In the configurations obtained from the QCA models, the starting point is the hypothesis that both thee values of any dichotomous variable (for instance, the presence or absence of a factor), in conjunction with others, can be potential causes of an effect. In the specific case of consumption of a food product where the degree of knowledge is related to consumption, it is difficult to accept that both values of any SPI (presence and absence of the objective knowledge) will have the same explanatory power, unless it is nil. Also, obviously, the communication activity may generate knowledge or may have no effect, but it cannot reduce the target’s knowledge. In other words, it is not possible to increase a lack of knowledge, or to communicate something which is untrue, leading consumers into error, for obvious ethical and legal reasons.

Consequently, the second stage is to process the set of initial combinations so that only the conditions or SPIs that are present (right answers) for a particular output are considered. As a result, the combination ABcDe will become ABD. It is important to bear in mind that this conversion changes the initial consistency rating, as because it is less restrictive, a greater number of people possess this configuration. The consistency therefore needs to be recalculated (Equation 3), where $\chi$ is the combination initially obtained but without the conditions that have been removed.

\[
\text{Recalculated consistency} = \left[\frac{\chi^+}{\chi^+ + \chi^-} \right] \quad (3)
\]

Based on this new consistency (adapted consistency), the best combinations of SPIs are again selected. Their hierarchy may be different from the initial order. In this way, optimum combinations of SPIs can finally be reached and integrated as a set into a campaign. These combinations satisfy the twin requirements of potential (efficacy and efficiency) to change a high number of behaviours, owing to their efficacy and to the number of people who do not possess this combination, and parsimony, as they are composed of little information as a result of the simplification effected by the software.

Although the first stage of the process of selecting the set of SPIs is based on the process for selecting causal relationships used in csQCA models, there are some clear differences between the two procedures (Table 1) due to their different aims. The main differences can be summarised as follows:

1. The use of consistency as a basic indicator and its adaptation through removal of absent conditions.
2. Not using the coverage (the proportion of total cases covered by a combination) as a criterion for selecting combinations. Although this is fundamental in QCA, as it indicates the frequency with which this causal combination occurs in reality, in the case of SPI selection the opposite effect must be sought in order to achieve a campaign that can insert a new and effective combination of SPIs into the minds of the greatest possible number of people.
3. Choosing the complex solution in order not to introduce counterfactuals (configurations with no cases) into the simplification processes (Ragin, 2008; Ragin et al., 1996; Ragin and Sonnett, 2005; Rihoux, 2001) and to avoid researcher subjectivity bias$^2$.

$^2$ The counterfactuals can provide different solutions and they are very useful when the number of cases (evidence empirical) is small.
Table 1. Comparison between traditional crisp-set Qualitative Comparative Analysis (QCA) models and the adapted model to the problem of significant pieces of information configuration.

<table>
<thead>
<tr>
<th>Objective</th>
<th>Crisp-Set QCA (traditional)</th>
<th>Crisp-Set QCA (application)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1: Initial data</td>
<td>Set of dichotomised explanatory variables (0, 1). Dichotomised effect variable (0, 1).</td>
<td>To identify which combinations of variables can alter an outcome the most. Pragmatic approach, seeking efficacy, efficiency and parsimony. Choosing the most suitable combination(s).</td>
</tr>
<tr>
<td>Phase 2: Constructing the truth table</td>
<td>The truth table shows all the configurations and how frequently they appear. Analysis of contradictory configurations (the same combination of conditions leading to different outcomes) and theoretically plausible configurations with no observed cases (counterfactuals or logical remainders).</td>
<td>Using the complex solution (no counterfactuals) in order to minimise researcher subjectivity, given that SPI patterns have no theoretical underpinning (there is no theory to provide guidance on the logic of the counterfactual).</td>
</tr>
<tr>
<td>Phase 3: Boolean minimisation Process of eliminating/ simplifying variables with no effect on the dependent variable when found in combination with others.</td>
<td>Three possible solutions depending on the use of counterfactuals in the minimisation process: complex (no counterfactuals used), parsimionous (all counterfactuals used) and intermediate (the only counterfactuals entered are those that fit and make sense and have been reviewed by the researcher).</td>
<td>Stage 1. Selecting the most suitable set of combinations based on high consistency, as the objective is to insert the specific SPI combinations with the greatest effect into the target’s mind.</td>
</tr>
<tr>
<td>Phase 4: Analysis and interpretation</td>
<td>Selecting the set of combinations with the maximum explanatory power (largest possible number of cases presenting the outcome to be explained), saturation of all possible causes based on high consistency (relationship between combination and effect) and high coverage (frequency of occurrence of this combination). Preference in the literature for the intermediate solution, which simplifies but not excessively (does not lose important information or eliminate necessary conditions).</td>
<td>Stage 2. Calculating the modified consistency, eliminating the absent variables from the combinations. Selecting the most suitable combination(s), based on individual and joint consistency.</td>
</tr>
<tr>
<td>Phase 5. Proposed solution</td>
<td>Set of alternative causal combinations that explain the greatest possible quantity of a phenomenon. Holistic and parsimionous explanation.</td>
<td>Selecting the most effective combination(s) of SPIs to change the largest possible number of behaviours.</td>
</tr>
</tbody>
</table>

4. An empirical study: choosing significant pieces of information to increase the demand for virgin olive oil in Spain

4.1 The problem of knowledge and consumption of olive oils in Spain

Although Spain is the biggest olive oil producing country in the world and consumption of this product is deeply rooted in the culinary culture of the country and is one of the basic symbols of its food, the level of ignorance and confusion about olive oil and its different classes, uses and characteristics is surprisingly high. According to a number of studies over the years, this is a persistent situation (Langreo, 2000; MAGRAMA, 2005; Navarro et al., 2010) and gives a clear indication of the ineffectiveness, non-existence and/or lack of intensity of information campaigns targeting the end consumers. In some of these studies, this ignorance is directly related to important effects in the market, increasing the visibility of price as a purchase factor, reducing the price differential for different qualities, increasing the competition between different qualities
and discouraging the producers from taking on the cost of improving the quality when it will not be rewarded by the market.

As a result of this ignorance and confusion, olive oil (a mixture of virgin olive oil and refined oil) is the most consumed in Spain, to the detriment of virgin olive oils. This is paradoxical, considering that virgin olive oils are a natural juice, of greater quality and with greater health benefits, and that there is no great price difference. Consequently, the starting point was that if knowledge of the product were to increase, the demand for higher quality oils would rise. However, what specific knowledge would it be useful to try to introduce into the mind of consumers through a communication campaign? What is it that consumers should know in order for them to consume virgin olive oils?

4.2 Method

A computer-assisted telephone interview (CATI) using a structured questionnaire was conducted in ten Spanish provinces (Malaga, Valencia, Madrid, Barcelona, Santander, Valladolid, Zaragoza, La Coruña, Almería and Salamanca) in September 2014 by means of a restricted random sample with the following quotas: maximum 60% women, not more than 10% unemployed, and minimum 40% with a university education. In each city, 40 interviews were conducted with people aged between 20 and 45 years and another 40 with people aged between 46 and 70 years. The final number of interviews was 829, representing an approximate sampling error of ±3.5% for the overall percentages (k=2; p=q=0.5).

The questionnaire included 11 items related to knowledge of olive oils, to each of which the respondents had to answer true, false or don’t know, as well as other questions concerning the uses and frequency of use of different oils and vegetable fats.

4.3 Measurement scales

The SPIs inserted into the questionnaire (Table 2) were developed with the help of experts from Centro de Estudios Avanzados del Olivar y Aceite de Oliva (Centre for Advanced Studies in Olive Groves and Olive Oil) and other olive oil market experts. The recommendations set out in the previous section were taken into account for this purpose.

The respondents’ answers (true, false, don’t know) were recoded as dichotomous variables according to whether each consumer answered each one correctly (1) or not (0). Use of the different types of oil was recorded through open questions (spontaneous recall) asking which oils and fats the respondents used for different purposes (frying, cooking hot dishes, in salads and dressings, and on bread). This information was used to generate a dichotomous variable recording whether or not each respondent consumed virgin olive oils. Finally, the degree of product use was recorded on another scale (used daily, weekly, monthly or seldom).

3 Items are the result of many years of work, different surveys and qualitative research. Nowadays, there is a bank of questions with more than 30 items in order to measure the knowledge level about olive oils. This range of questions has been designed for more involved groups. We have successfully used it in a survey in which the respondents were chefs (some of them with stars in the Michelin guide). However, a shorter list of items was used with consumers given that they have lower involvement and knowledge. Moreover, surveys should be shorter in order to measure other variables. Naturally, the items have been filtered and tested in three aspects: (1) they are well understood (unambiguous); (2) experts agree with the correct answer; and (3) they are linked to expressions that consumers often use to explain their behavior about the product. (For further information about the complete list of items, please contact the authors).
4.4 Results

Considering that virgin olive oil (VOO) consumption can be increased by non-consumers’ consuming the product and/or by consumers’ increasing the quantities they consume, and that the arguments or SPIs may be different for the two groups, two separate analyses were performed.

However, it should be remembered that the entire procedure is based on the hypothesis of a causal relationship between knowledge and consumption. Consequently, prior to the analysis, two hypotheses were tested: (1) knowledge is positively related to consuming higher quality olive oils and (2) knowledge is positively related to greater use or frequency of use of these oils (among the group of consumers). For the analysis, the number of correct answers (C), the number of errors (E) and the number of items not answered (D) was counted for each respondent. A new variable was also calculated by subtracting the errors and don’t knows from the number of correct answers (C-D-E) to give the ‘degree of overall knowledge’, with a range of -11 to 11 (there are eleven knowledge items). Table 3 shows the relationship between this degree of knowledge and the use of virgin olive oils (hypothesis 1) and frequency of use (hypothesis 2).

---

**Table 2. Items inserted into the questionnaire.**

<table>
<thead>
<tr>
<th>Item</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPI1. Olive oil is pure olive juice without manipulation</td>
<td>Virgin olive oils are exclusively pure olive juice. It was suspected that a large number of consumers think that olive oil is also pure olive juice, affecting the demand for virgin olive oils.</td>
</tr>
<tr>
<td>SPI2. Refined olive oils are of higher quality</td>
<td>It was suspected that there was confusion in the market about the term ‘refined’ because of its positive connotations (purifying, improving). In the context of olive oils, it means using physical and chemical processes (e.g. a jet of air heated to 200 °C) to remove all the flavour, odour and colour from oils of poor quality. This refined oil is then mixed with virgin olive oils to give it flavour and the result is called olive oil.</td>
</tr>
<tr>
<td>SPI3. Refined olive oils are healthier</td>
<td>This item was introduced to assess the consumers’ knowledge of the meaning of the term ‘refined’ and its association or otherwise with health. The refining process removes all the minor components with proven beneficial effects for health.</td>
</tr>
<tr>
<td>SPI4. Oils with higher acidity have a stronger flavour</td>
<td>The connotations of the word ‘acidity’ (its association with flavour) and the use of acidity in communication campaigns has led to confusion in the market. In the context of olive oils, it is unrelated to flavour.</td>
</tr>
<tr>
<td>SPI5. Olive oils are a mixture of virgin and refined</td>
<td>The main characteristic of olive oil is that it is a mixture of refined olive oil with a small proportion of virgin olive oil. This is considered a basic or primary piece of information for an understanding of olive oils and their different classes and qualities.</td>
</tr>
<tr>
<td>SPI6. The main factor in differentiating between qualities is acidity</td>
<td>This is related to item 4, has been used in communication campaigns and is misleading. There are other criteria for measuring quality (tasting scores, peroxides index, etc.)</td>
</tr>
<tr>
<td>SPI7. Olive residue (US pomace) oil is made from the residue of olives</td>
<td>Potential confusion between olive oils and olive residue oils. The latter are extracted from the residue of olives and also need to be refined. They are of much lower quality.</td>
</tr>
<tr>
<td>SPI8. Olive oil is as healthy as virgin olive oil</td>
<td>Owing to undifferentiated communication campaigns, attempts have been made to convey that everything labelled as ‘olive’ is healthy and good. However, there are clear differences between virgin olive oils and the rest.</td>
</tr>
<tr>
<td>SPI9. Virgin olive oils are of higher quality than olive oil</td>
<td>Another of the statements that helps to form a minimum level of knowledge that will help consumers to avoid confusion when choosing.</td>
</tr>
<tr>
<td>SPI10. Extra virgin olive oils are organic</td>
<td>Another potential source of confusion owing to the connotations and evocative power of the term ‘organic’. The quality of the oil is independent of whether it is organic.</td>
</tr>
<tr>
<td>SPI11. Olive oil is less fattening than virgin olive oil</td>
<td>One of the false beliefs is that some oils are less fattening than others.</td>
</tr>
</tbody>
</table>

1 SPI: significant pieces of information.
To summarise, there would appear to be clear signs that knowledge about virgin olive oils is positively related to consumption and, to a lesser degree, to frequency of use. This conclusion warrants the assumption that it is appropriate to attempt to select combinations of SPIs to increase demand for the product.

**Analysis 1: respondents who do not currently consume virgin olive oils for any use (19% of the sample, 157 consumers)**

The initial configurations were formed using all the cases in the questionnaire, taking all the processed (dichotomised) SPIs as explanatory variables except for items 7 and 10, which were understood not to affect the consumption or otherwise of virgin olive oils, and the dichotomous variable of consumption or otherwise of virgin olive oils as the dependent variable. As all the variables, both dependent and independent, were dichotomous, the QCA program was used in Crisp-Set mode (Ragin, 1987, 2000; Rihoux and De Meur, 2009).

The initial configurations chosen presented at least four cases (0.5% of the sample) in order to avoid rare or random solutions and a consistency level of at least 95% in order to obtain only SPI combinations that were really effective, in other words, ones that led to use of the product at least 95% of the times that the consumer possessed that specific SPI combination. After choosing the complex solution and restricting the analysis to consistencies higher than 0.95, the absent conditions were eliminated, and the consistency of each combination was recalculated. The results are shown in Table 4.

According to these results, several of the solutions would be suitable. It should be noted at this point that these solutions fulfil our initial hypothesis: the SPI combinations are closely related to consumption of virgin olive oils. Although there are few differences among these solutions, the best would seem to be 6: in theory, or assuming the absence of other variables influencing consumption and assuming that all the non-consumers received the message and considered it credible, it could have an effect on 96.47% of this group. In this context, the SPIs or knowledge that should be included in a communication campaign should make it clear that refined olive oil is not of higher quality, that olive oil is a mixture of virgin and refined olive oils, that olive oil is not as healthy as virgin olive oil, that virgin olive oil is of higher quality than olive oil and that olive oil is not less fattening than virgin olive oil. In short, it should be stressed that virgin olive oils are the highest quality and most healthy, that there is a basic difference between olive oil and virgin olive oil in that olive oil is a mixture of virgin and refined olive oils, and that they are both equally fattening.

<table>
<thead>
<tr>
<th>Knowledge about olive oils (11 items)</th>
<th>VOO use</th>
<th>VOO frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Total correct</td>
<td>5.4</td>
<td>4.3</td>
</tr>
<tr>
<td>Total errors</td>
<td>3.3</td>
<td>3.4</td>
</tr>
<tr>
<td>Total don’t know</td>
<td>2.2</td>
<td>3.2</td>
</tr>
<tr>
<td>Overall knowledge (c-d-e)&lt;sup&gt;3&lt;/sup&gt;</td>
<td>-0.1</td>
<td>-2.3</td>
</tr>
</tbody>
</table>

<sup>1</sup> Numbers are median values.

<sup>2</sup> The hypothesis of equal numbers of correct answers, errors, etc. among those using and not using VOOs and among high or low frequencies of VOO use was tested by the Mann-Whitney U test; *** = 𝑃<0.01 and * = 𝑃<0.1. The symbols (–) and (+) indicate an inverse or direct relationship between the item and VOO use or high frequency of use.

<sup>3</sup> c: correct answers, d: don’t know answers, e: erroneous answers.

![source](https://www.wageningenacademic.com/doi/pdf/10.22434/IFAMR2019.0004 - Sunday, June 30, 2019 9:57:42 AM - University of Minnesota - Twin Cities IP Address:134.84.17.144)
The effect variable, ‘frequency of use of VOO’, was recoded as two values, where 1 represented frequent consumption of these oils (daily or weekly) and zero the absence of frequent consumption, in other words, monthly or very occasional use (1 = frequent consumption, 0 = infrequent consumption). In this case too, items 7 and 10 were excluded from the analysis as they were considered not to influence frequent or infrequent consumption of virgin olive oils. The subset of the sample analysed in this case comprised 672 respondents, of whom 413 (61.46%) were infrequent consumers of virgin olive oils.

Table 4. Analysis 1: non-consumers of virgin olive oils.\textsuperscript{1,2}

<table>
<thead>
<tr>
<th>SPI1 original combinations</th>
<th>Original consistency</th>
<th>SPI1 adapted combinations</th>
<th>Adapted consistency</th>
<th>Theoretical effectiveness (consumers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SPI1<em>SPI2</em>SPI3<em>SPI6</em></td>
<td>1.0000</td>
<td>SPI1<em>SPI2</em>SPI3*</td>
<td>0.9394</td>
<td>147</td>
</tr>
<tr>
<td>2 SPI1<em>SPI3</em>SPI8<em>SPI15</em></td>
<td>0.9565</td>
<td>SPI2<em>SPI11</em>SPI5*SPI9</td>
<td>0.9577</td>
<td>150</td>
</tr>
<tr>
<td>3 SPI1<em>SPI2</em>SPI3<em>SPI8</em></td>
<td>0.9647</td>
<td>SPI2<em>SPI8</em>SPI11<em>SPI5</em></td>
<td>0.9147</td>
<td>143</td>
</tr>
<tr>
<td>4 SPI1<em>SPI2</em>SPI3<em>SPI8</em></td>
<td>0.9045</td>
<td>SPI1<em>SPI2</em>SPI3*</td>
<td>0.9355</td>
<td>147</td>
</tr>
<tr>
<td>5 SPI1<em>SPI2</em>SPI3<em>SPI6</em></td>
<td>0.9297</td>
<td>SPI2<em>SPI11</em>SPI5*SPI9</td>
<td>0.9129</td>
<td>143</td>
</tr>
<tr>
<td>6 SPI1<em>SPI2</em>SPI3<em>SPI6</em></td>
<td>0.9535</td>
<td>SPI2<em>SPI8</em>SPI11<em>SPI5</em></td>
<td>0.9148</td>
<td>143</td>
</tr>
<tr>
<td>7 SPI1<em>SPI2</em>SPI3<em>SPI6</em></td>
<td>0.9524</td>
<td>SPI2<em>SPI11</em>SPI5*SPI9</td>
<td>0.9111</td>
<td>143</td>
</tr>
</tbody>
</table>

\textsuperscript{1} All significant pieces of information (SPIs) were analyzed. None of them were necessary conditions in which case the SPIs would have to be excluded from the analysis (Ragin, 2006).

\textsuperscript{2} * means Boolean operator ‘AND’.

\textbf{Analysis 2: respondents who currently consume virgin olive oils for some use (81\% of the sample, 672 consumers)}

The effect variable, ‘frequency of use of VOO’, was recoded as two values, where 1 represented frequent consumption of these oils (daily or weekly) and zero the absence of frequent consumption, in other words, monthly or very occasional use (1 = frequent consumption, 0 = infrequent consumption). In this case too, items 7 and 10 were excluded from the analysis as they were considered not to influence frequent or infrequent consumption of virgin olive oils. The subset of the sample analysed in this case comprised 672 respondents, of whom 413 (61.46%) were infrequent consumers of virgin olive oils.
With a consistency level of 0.95 and following the same steps as in analysis 1, the final solution shown in Table 5 was obtained, in which SPI combination number 5 appears to be the most appropriate.

It is worth noting that in this analysis, the consistency of the chosen SPI combinations fell considerably after the final step (the best of them barely reached 0.53), which is coherent with the results of Table 3 and indicative of the smaller effect of knowledge in this group, whether because they already knew more or because once the product has been consumed it is not so easy to modify the consumer’s behaviour in the direction of consuming greater quantities.

Considering the two groups of interest (non-consumers and infrequent consumers), an indicator that can be calculated is the global theoretical efficacy (GTE), which is the proportion of respondents that would change their behaviour under optimum conditions (they all receive the message, all consider it credible, and there are no limiting factors such as taste, product availability or lack of purchasing power). Since non-consumers made up 19% of the total and infrequent consumers comprised 61% of consumers, and that the theoretical effectiveness was 151 and 220 respectively, the GTE could be calculated as follows:

$$GTE = \frac{(151+220)}{(157+413)} = 0.65 \quad (4)$$

In other words, under the above optimum conditions it would be possible to change the behaviour of 65% of the target at most. This indicator complements traditional ones used in campaigns from the point of view of efficacy and also presents certain strengths that make it useful: It is easy to calculate, valid and realistic, as it directly links the information understood to consumption (in other words, there is a direct causal relationship) and the maximum efficacy of a campaign can be calculated (at a theoretical level) before launching it on the market. This is important, since the existing models and indicators to measure efficacy are used after launching these campaigns.

### Table 5. Analysis 2: consumers of virgin olive oil.\(^1,2\)

<table>
<thead>
<tr>
<th>SPI original combination</th>
<th>Original consistency</th>
<th>SPI adapted combination</th>
<th>Adapted consistency</th>
<th>Theoretical effectiveness (consumers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SPI1<em>SPI2</em>~SPI3<em>~SPI4</em> <em>SPI6</em>~SPI8<em>~SPI11</em> <em>SPI5</em>~SPI9</td>
<td>1.0000</td>
<td>SPI1*SPI2</td>
<td>0.3600</td>
<td>149</td>
</tr>
<tr>
<td>2 SPI2<em>SPI8</em>SPI11<em>~SPI1</em> <em>SPI3</em>~SPI4<em>~SPI6</em> <em>SPI5</em>~SPI9</td>
<td>1.0000</td>
<td>SPI2<em>SPI8</em>SPI11</td>
<td>0.3990</td>
<td>165</td>
</tr>
<tr>
<td>3 SPI1<em>SPI6</em>SPI11<em>SPI9</em> <em>SPI2</em>~SPI3<em>~SPI4</em> <em>SPI8</em>~SPI5</td>
<td>1.0000</td>
<td>SPI1<em>SPI6</em>SPI11*SPI9</td>
<td>0.4400</td>
<td>182</td>
</tr>
<tr>
<td>4 SPI2<em>SPI6</em>SPI11<em>SPI15</em> <em>SPI9</em>~SPI1<em>~SPI3</em> <em>SPI4</em>~SPI8</td>
<td>1.0000</td>
<td>SPI2<em>SPI6</em>SPI11*SPI9</td>
<td>0.4054</td>
<td>167</td>
</tr>
<tr>
<td>5 SPI2<em>SPI3</em>SPI4<em>SPI6</em> <em>SPI8</em>SPI9<em>~SPI1</em> <em>SPI11</em>~SPI5</td>
<td>1.0000</td>
<td>SPI2<em>SPI3</em>SPI4* <em>SPI6</em>SPI5<em>SPI8</em> *SPI9</td>
<td>0.5333</td>
<td>220</td>
</tr>
</tbody>
</table>

\(^1\) SPI: significant piece of information.
\(^2\) * means Boolean operator ‘AND’.
5. Conclusions

Nowadays people are increasingly interested in the purchase of food. In addition, the range of products and information in the market has raised owing to the globalisation process. In this context, it is of pivotal importance that consumers have the correct information in order to take the decision to consume a food product (or to consume more of it). However, in the area of communication management, the literature has focused on decisions concerning how to communicate. In this paper the attention has shifted to a prior question that is often overlooked: what to communicate. This question can be very important when devising generic communication campaigns. The main contribution proposed is a model for choosing the information content of campaigns, which will be particularly useful in situations where the consumers’ level of objective knowledge is positively related to food product consumption.

Since the problem of content selection is to some extent analogous with csQCA models, the solutions these provide are used as the initial reference framework. However, a series of conversions and indicators is proposed to adapt them to the specific purpose of this paper: selecting the most appropriate information content to provoke a greater market response.

The model was applied to the specific case of increasing the demand for virgin olive oil, enabling five SPIs to be selected as the main body of information content for a communication campaign to transform non-consumers into consumers and six SPIs to increase consumption among existing users of this product. Considering the observed relationship between possession of this knowledge and consumption, it was estimated that in optimum message coverage and repetition conditions it would be possible to influence the behaviour of a maximum of 96.4% of non-consumers and 53.3% of infrequent consumers. This information could be useful for the bodies responsible for promoting the sector, at both government and producer association level (International Olive Council, Olive Oil Interprofessional Organisation, Ministry of Agriculture, etc.), when selecting the information content to be included in promotion campaigns to increase the demand for the product. It should be note that the different combinations of SPIs are very similar in analysis 1, in this context of virgin olive oils. For this reason, it would be interesting to study this method in other type of products.

An element of this study that may be highlighted is the proposal of new indicators (theoretical efficacy and global theoretical efficacy) that complement existing ones for the management or evaluation of communication campaigns and have the advantage of simplicity, validity and the possibility of calculating them before conducting the campaign. In essence, these indicators measure the capacity of one or more combinations of SPIs to alter behaviour or, in other words, how many consumers in the target group would change their behaviour under optimum conditions. This should be considered a maximum level, assuming global dissemination of the message, total credibility of its content and the absence of limiting factors such as an insufficient income to purchase the product. These limitations mean that the indicators are optimistic approximations, although they fulfil the basic and essential function of selecting which information configurations are most appropriate. In this respect, it would be possible to deepen the analysis by relating each SPI configuration to media management data or consumer data (socio-demographic, psycho-graphic, etc.), as each configuration corresponds to a group of consumers with particular specific knowledge who may present additional shared characteristics. Another interesting line of enquiry could be to develop more complex indicators or integrate them with others used for media planning or effectiveness measurement in order to improve the instruments used in designing communication campaigns.

References


Rihoux, B. and C.C. Ragin. 2009. *Configurational Comparative Methods, Qualitative Comparative Analysis (QCA) and related techniques*. Sage, California, CA, USA.


