



## **Journal of Internet Banking and Commerce**

*An open access Internet journal (<http://www.icommercecentral.com>)*

*Journal of Internet Banking and Commerce, August 2021, Vol. 26, No. 8*

# **A Multicriteria Model to Evaluate E- Commerce Websites Under the Perspective of the Customer**

---

**Luís Valadares Tavares\***  
Lusiadauniversity, Portugal  
Tel: +351929127497  
Email: [lv@lis.ulusiada.pt](mailto:lv@lis.ulusiada.pt)

**Pedro Arruda**  
LisbonUniversity,Portugal

**José Antunes Ferreira,**  
LisbonUniversity, Portugal

**Ana Cristina Saraiva**  
LusiadaUniversity, Portugal

**AfonsoVinagreAmaral**  
LisbonUniversity, Portugal

---

### **Abstract**

E-Commerce is reaching a new stage of conspicuous and widely disseminated development also due to the recent challenges of COVID 19. Thus, the process of shopping includes a first decision problem quite critical to the success of any business: which site of e-commerce should be selected by the consumer?

This means that the selection of such site should be studied in terms of the satisfaction of the consumer when using it for shopping compared with other competitive sites rather on its descriptive features. The list of studies comparing and evaluating the E-Commerce sites is quite long but almost no attempts have been made to model the satisfaction function of the consumer and so they are not particularly relevant to study the competitive choice of sites by the consumer.

In this paper a new model of consumer satisfaction is proposed using an approach–TRIDENT-based on the Multi Attribute Utility Theory (MAUT) and the critical stage of estimating the weights of multiple criteria is solved using an original method OptionCards which avoids the shortcomings of more traditional surveys.

The utility function describing the satisfaction function concerning websites of E-Commerce is estimated using the OptionCards method for a group of young professionals and university students confirming a similar importance assigned to the three major criteria.

Such utility function was used to estimate the rating of the 14 major Portuguese websites of E-Commerce using the answers of a group of young professionals using E-Commerce and the overall score confirms their relative level of popularity.

**Keywords: E-commerce; Customer satisfaction; Multicriteria model; Optioncards method.**

## **Introduction**

### **Why a Multicriteria Model to Evaluate Websites for E-commerce?**

E-commerce is becoming a prevailing type of commerce which is based on the ubiquitous development of digital society and on the internet penetration as there are more than 2 billion digital buyers [1] and more than 250 M internet domains [2] including about 125 M domains belonging to the class of “dot com” [3]. Several trends are quite promising [4] and recent challenges as the pandemic spread of COVID19 are accelerating the adoption of e-commerce even by older consumers replacing traditional retail shopping commerce by online business and so the last estimates of e-commerce are sub-estimating its growth as they are related to 2019. Even so, the average penetration (percentage of population with 16 to 75 years old using e-commerce) for EU is about 71% and more than 80% is already achieved by States such as UK, Germany, Netherlands and the Nordic countries [5].

Therefore, the evaluation of websites promoting e-commerce has been studied by many authors as it is clear from the reviews published by Chiou et al. but unfortunately their main output is the presentation of long lists of features which are supposed to be relevant for Goisuch evaluation without any systemic or taxonomic structure as it is clear from the summary table [6-8].

Furthermore, some of these features are overlapping such as “usability” and “quality” and doubts remain about the procedures to evaluate each feature.

An evaluation model of a website is not a neutral instrument because it may be constructed to assess the achieved level of Information and Communications Technology (ICT) professional skills as it seems to be the case of world best enterprises, Sorum et al. emphasizing functionality, design, originality, etc. or it may be considered as a key instrument to facilitate the use of e-commerce by consumers as it should be the case for the objectives of this paper. Therefore, in this case, the evaluation should reflect the customer satisfaction for the purpose of e-commerce [9].

The technical quality of the site may not even be highly correlated with the customer satisfaction as it was shown by Mangiaracina et al. studying the customer satisfaction for awarded websites in terms of their quality and concluding that there is “lack of correlation between website quality and user satisfaction”.

The evaluation instrument should be independent of the commercial contents because the purpose is to assess the ability to implement e-commerce rather than assessing the commercial strategy of the vendor but this distinction was not clear for some authors such as they suggest the inclusion in the website evaluation process, features such as price and

promotion which belong to the evaluation of the commercial strategy rather than to the website.

The evaluation model of a website for e-commerce should represent the evaluation by the customer and so no valid model can be proposed without aiming to estimate the customer satisfaction in terms of the relevant attributes, under the perspective of the consumer and without assessing the relative importance assigned by the consumer to such attributes [10].

The previous published research has not formulated the evaluation of websites for e-commerce through a utility function modelling the customer satisfaction although the need to consider the level of satisfaction was pointed out by several authors. Actually, Mangiaracina et al. and Barnes et al. emphasized the need to study the customer journey experience but such study is just based on unweighted sums of scores [11]. Between 150 and 170 parameters are required (!) making their estimation rather unfeasible. Recently, only the perspectives of information display, pages organization and graphical design were considered [12-16] and two other authors Rouyendegh et al. and Kaya et al. presented complex and fuzzy models using the AHP approach to compare alternatives sites but only based on pairwise comparisons and so the number of experts and alternatives was very small and subject to rank reversal [17-19].

Other authors such as surveyed the level of satisfaction for information websites concluding that “navigation and content” are important but no evaluation metrics was proposed and suggested that the customers weights should be studied but no procedure was developed and no further results were published along this line of research [16].

Thus, the final conclusions of “This absence of coordinated theory development causes the research in Internet marketing to appear haphazard and unfocused” and of “This study found that the existing literatures do to have any commonly agreed-upon standards or techniques for website evaluation” are well justified.

### **The Research Question and Methodologic Options**

The main research question addressed by this paper is: How can the satisfaction of the customer are modeled for the purpose of evaluating websites of e-commerce?

The adopted methodology is the Multi-Attribute Utility Theory (MAUT) [18] and the structure of the model is based on an additive linear function of multiple criteria and sub-criteria which have to be defined and constructed in order that all the relevant perspectives contributing for the level of satisfaction of a customer when is using an website for e-commerce will be captured and conveniently described and represented [19].

The proposed model will be constructed in terms of a system of attributes following a tree structure in order that each node the number of attributes will be reasonably small (e.g.,  $n=3$ ), to facilitate the estimation of weights by the customer as many authors of decision theory have shown that such estimation is much harder and less rigorous if  $n$  is higher [20].

The evaluation model should adopt a scoring function so that the consumer can assess each attribute and the OptionCards method [21] is used to elicit the weights assigned by the customer to each attribute.

Summing up, the objective of this paper is to develop a multicriteria model representing the satisfaction of the customer due to using a specific website for E-Commerce in order that the estimated evaluation will represent such satisfaction.

The application of this model implies the estimation of the appropriate value function

representing the satisfaction of the population of customers which are supposed to use the e-commerce website and such estimation depends on the type of population to which the website is dedicated [22].

## **The Proposed Model**

The proposed model aims to describe the level of satisfaction of the customer for the target group using a specific website for e-commerce in terms of a tree-structure of most relevant attributes and so the following components of the model have to be proposed:

- a) A tree-structure of criteria in order so that all relevant perspectives for the consumer will be considered.
- b) Scoring functions for each attribute.
- c) A satisfaction model using a linear additive function MAUT.
- d) A procedure recommended to elicit attributes weights.

The proposed model follows the lines of the TRIDENT model meaning that at each node the number of attributes will not exceed 3 and that the weights functions will be studied in terms of the 2-dimensional representation of indifference lines between pairs of evaluated alternatives.

The scoring function for each criterion or sub-criterion adopts a Likert scale ranging from 1 ("very bad") to 5 ("excellent") and the aggregation of scores at each level is obtained by a linear additive function in terms of weights which have to be elicited from the consumers.

These four components of the model are presented in the following sections.

## **Results and Discussion**

### **1. The Major Attributes And Scoring Systems To Assess Consumer Satisfaction**

**1.1 Major attributes and scoring functions:** The proposed three major attributes to describe the customer satisfaction are:

- a) The quality of the website navigation experience;
- b) The level of trust of the consumer to the whole commerce transaction;
- c) The quality of information concerning the delivery logistic system.

These attributes as well as the sub-criteria within each one have been chosen after a review of existing literature European Union [23] which has identified these dimensions as the most critical ones to a successful experience of e-business.

**1.2 On the navigation experience:** Online commerce is based on the interaction of the potential consumer and a website which means that the traditional strolling along the main street and the consequent window shopping is replaced by the quite different experience of site opening, browsing and searching for the most appropriate product or services. This experience is particularly critical for the success of any online purchase because an unpleasant, tiring or cumbersome experience is responsible for the loss of the potential consumer and eventual "immunity" against the website in question.

In this domain the disparity of results is quite stunning, ranging from very high attractive levels to very unsuccessful results [24].

An analytical decomposition of these attribute can be suggested considering three sub attributes:

Registration easiness: A best practice [25] can be identified as requiring just the email and password leaving for a [26] further stage the collection of additional data, such as address, Tax Identification Number, Payment options, etc.

Therefore, the proposed descriptor, I1, describing the registration easiness is defined by number of items, N, and so the following scoring is proposed (Table. 1).

Descriptor	Score
$N \leq 2$	5
$N = 3$	4
$N = 4$	3
$N = 5$	2
$N \geq 6$	1

**Table 1 Descriptor I1 (Registration easiness)**

Search engine and advisory support: All websites include a search option, but the level of the semantic inference engine and its learning ability can range from very low intelligent algorithms to high quality expert systems.

Alternative indicators to evaluate the capabilities of each of expert system have been proposed but in this case a descriptor of their quality can be based on the ability of finding the keyword representing the desired object using incorrect variations of that keyword and measuring the distance between the two words (original word and incorrect variation) in terms of the number of incorrect letters. For example, assuming the consumer looks for a smartphone (keyword, x) and introduces the word “smarkfone” (search word, y) does the engine still provide the correct answer? The semantic distance between two words can be formulated in alternative ways but in this model the following definitions are adopted:

$$D(x, y) = \frac{1}{2} (A(y \rightarrow x) + s(y \rightarrow x))$$

Where:

$D(x,y)$  is the absolute semantic distance between the x word and the word y

$A(y \rightarrow x)$  is the number of letters to be added to y so that x will be obtained

$S(y \rightarrow x)$  is the number of letters to be subtracted from y so that x will be obtained

Therefore, the relative semantic distance  $d(x,y)$  can be determined by:

$$d(x, y) = \frac{D(x, y)}{L}$$

where L is the average dimension of the two word given by:

$$L = \frac{1}{2} (I(x) + I(y))$$

where:  $I(i)$  is the number of letters of the i word.

Consequently, the quality of the search engine will be studied in terms of the probability P of finding the correct keyword in terms of  $d(x,y)$  and usual statistical parameters can be used to describe such relationship such as its average, its variance, or its quantile for any statistical level.

This means that the descriptor of the quality of the search engine (I2(s)) will be expressed by a statistical measure of P for the specified level of  $d(x,y)$  measured by the following scale (Table. 2).

Descriptor	Score
P ≥ 80%	5
80% > P ≥ 60%	4
60% > P ≥ 40%	3
40% > P ≥ 20%	2
20% > P ≥ 0%	1

**Table 2 Descriptor I2(s) (Quality of the Search Engine)**

The application of this measuring scale implies testing the website using random keywords which describe the search relevancy for the e-consumer including corrupted random variations of two letters in each of these keywords.

Another important feature concerns the ability to receive advisory support helping the consumer to make the best choice. This support can be given by different types and levels of instruments concerning the specific good or service being considered by the customer (Table. 3).

Descriptor	Score
A direct video call line, phone call line and an email contact allowing direct interaction between the customer and the staff+a menu of frequently asked questions (FAQ)	5
A direct phone call line and an email contact allowing direct interaction between the customer and the staff+FAQ	4
An email contact allowing interaction between the customer and the staff+FAQ	3
FAQ	2
No advisory support	1

**Table 3 Descriptor I2(a) (Quality of the advisory support)**

I2 can be given by:

$$I2 = \frac{1}{2} I2(s) + \frac{1}{2} I2(a)$$

Full specification of the product or service: The online customer experience is very sensitive to these attributes because direct interaction between customer and service/ product does not exist any longer and therefore the website should replace the three-dimensional process.

The following descriptor and measuring scale are proposed (Table. 4).

Measurement	Score
Full description of the product with at least 3 HD pictures and/or video	5
Full description of the product with less than HD pictures and no video	4
Incomplete description and one or more pictures and no video	3
Incomplete description without pictures or video	2
No description and no picture or videos	1

**Table 4 Descriptor I3 (Full specification of the product or service)**

**1.3 On the level of trust:** Any commercial transaction implies an underlying contract which

will be not accepted if there is no sufficient trust by the consumer on the merchant and on that specific transaction. Actually, the issue of trust is quite critical to the development of online commerce as there is no real physical experience of direct contact between buyer and seller.

Therefore, this major concept of trust should be considered under a threefold perspective:

- 1) Level of security of the website implying that the whole site is secure and not just the payment system avoiding leaks of information about the object or service to be purchased, passwords and registration data, etc.
- 2) Level of confidence on the transaction meaning that if the consumer is not satisfied, the transaction can be reversed easily and without additional costs for the consumer.

This explains why the European Union adopted a Directive on online commerce imposing a minimal period of 14 days to give the consumer the possibility of reversing the transaction, returning the product, and being fully reimbursed. It should be noted that some marketplaces like eBay or Amazon have adopted even more favorable rules managed by the so called “dispute centers”. However, other marketplaces based on RPC adopt different rules, namely on the return cost of the product.

- 3) Level of security of the payment system assuring that no frauds or leaks of data will occur.

Thus, the proposed descriptors and measuring scales for these three perspectives are presented in (Tables. 5-7).

Measurement	Score
If the whole website is a secured one (encrypted using Transport Layer Security (TLS))	5
If just some pages of the web site are secured	3
If no pages are secured	1

**Table 5 Descriptor I4 (Trust in the website (excluding the payment system))**

Measurement	Score
If reversing the transaction can be easily requested by the consumer through the site approved on less than 2 days and the consumer reimbursed also in less than 2 days without having to pay any returning costs.	5
If 5 if reversing the transaction can be easily requested by the consumer through the site approved on more than 2 days and less than 14 days and being the consumer reimbursed in the same period without having to pay any returning costs.	3
If reversing the transaction cannot be easily requested through the website and if the customer has to pay the return costs	1

**Table 6 Descriptor I5 (Trust in the transaction)**

Measurement	Score
If the payment system is fully secure and multiple options are offered including systems like ATM payment or PayPal avoiding having to introduce bank card data	5
If the payment system is fully secure but no options avoiding the introduction of bank card data are included	3
If the payment system is not fully secure	1

**Table 7 Descriptor I6 (Trust in the payment system)**

**1.3 On the website information concerning logistics:** This dimension is also quite critical because the purpose of the transaction is receiving the acquired product or service as soon as possible and without additional uncertainty. However, this model is devoted to the evaluation of the website rather than of the commerce process and so the object of this evaluation should concern the display of information concerning this dimension.

Therefore, this perspective can be evaluated according to the following three descriptors (Tables. 8-10):

Measurement	Score
Full information about the time and transportation mode is presented before completing the purchase.	5
Partial information about the time and transportation mode is presented before completing the purchase.	4
Full information about the time and transportation mode is just presented when the purchase is completed.	3
Partial information about the time and transportation mode is just presented when the purchase is completed.	2
No information is presented	1

**Table 8 Descriptor I7 (Delivery time)**

Measurement	Score
If online real time tracking is available and notices are sent before the delivery	5
If tracking is available just near the delivery time and notices are sent before the delivery	4
If no tracking is available but notices are sent before the delivery	3
If no tracking is available and no notices are sent	1

**Table 9 Descriptor I8 (Tracking information)**

Measurement	Score
Full information about flexibility of delivery time is presented before completing the purchase.	5
Partial information about flexibility of delivery time is presented before completing the purchase.	4
Full information about flexibility of delivery time is just	3



presented when the purchase is completed.	
Partial information about flexibility of delivery time is just presented when the purchase is completed.	2
No information is presented	1

**Table 10 Descriptor I9 (Flexibility of Delivery Time)**

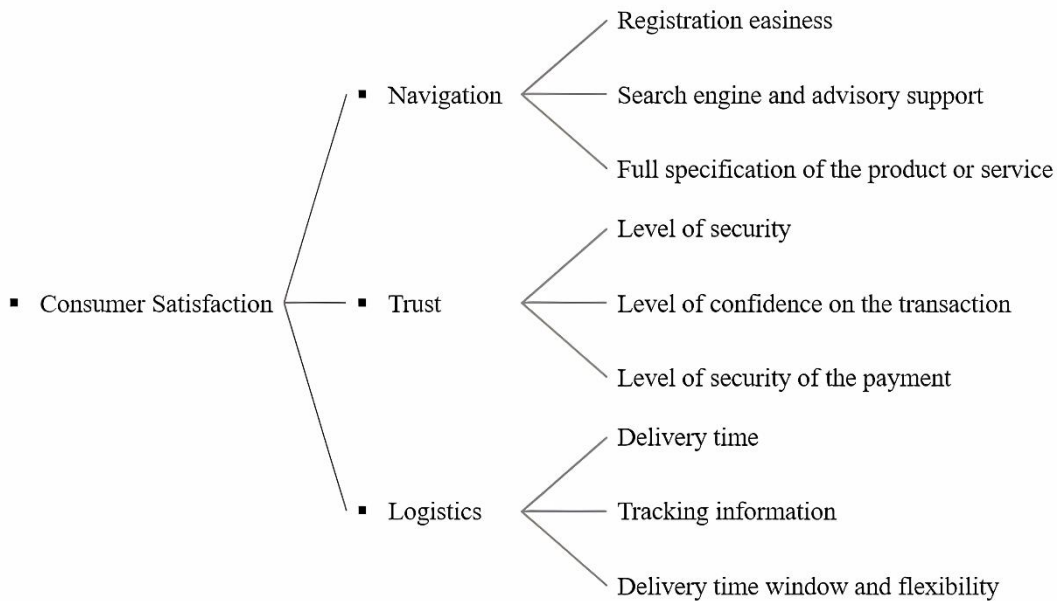
## 2. Multicriteria Evaluation

### 2.1 The Trident Model

This model was proposed by and is based on an application of Multi-Attribute Utility Theory (MAUT) to a multicriteria tree based on three branches stemming from each node and subdividing the weight space into sub-areas with the same ranking of alternatives. In this case the main three criteria are:

- Navigation
- Trust
- Logistics

and each of these nodes will be subdivided according to the presented sub-criteria. The evaluation tree is presented in (Figure. 1).



**Figure 1 Multicriteria evaluation tree of the customer satisfaction**

Therefore, for each node associated to three criteria, the MAUT evaluation function can be defined by:

$$U(i) = \sum_{j=1}^3 \lambda_j \cdot U_{ij}$$

Where  $U_{ij}$  is the utility of alternative  $i$  according to criteria  $j$ ,  $\lambda_j$  is the weight of criteria  $j$  with  $\sum \lambda_j = 1$  and  $U(i)$  is the global utility of alternative  $i$ .

Now, the application of this MAUT model at each node of the presented tree implies the estimation of the weights representing the values of each decision maker which is always a difficult problem to be solved. The most often adopted approach is based on enquires of the type "Do you consider criterion  $j$  equally important, more or less important than  $j'$ ?" so that estimates of the weights will represent such answers but the main criticism is that decision makers do not know exactly what does it mean to be equally, more or less important than. The method Option Cards proposed by adopts a complete different approach based on successive binary questions which are unequivocally understood by the decision maker avoiding the previous shortcomings of the traditional approach and thus it is presented in the next section.

## 2.2 The OptionCards Method

The weights selected by any decision maker can be represented by a point in a 2 dimension triangular space  $S_{1,2}$  bounded by the axes corresponding to  $\lambda_1, \lambda_2$  and  $\lambda_1 + \lambda_2 = 1$ .

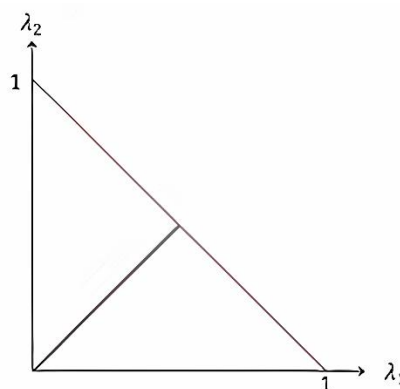
The subset of points for which there is indifference  $I_{ik}$  between 2 alternatives is a straight line defined by:

$$U(i) = U(k)$$

or

$$\lambda_1.(U_{i1}-U_{i3})+\lambda_2.(U_{i2}-U_{i3})+U_{i3}=\lambda_1.(U_{k1}-U_{k3})+\lambda_2.(U_{k2}-U_{k3})+U_{k3}$$

and this straight line (Figure. 2) subdivides the space into two subspaces:  $D_{ik}$  where  $U(i) > U(k)$  meaning that  $i$  is preferred to  $k$  and  $D_{ki}$  meaning that  $k$  is preferred to  $i$ .



**Figure 2 The representation of an indifference line in the weights space**

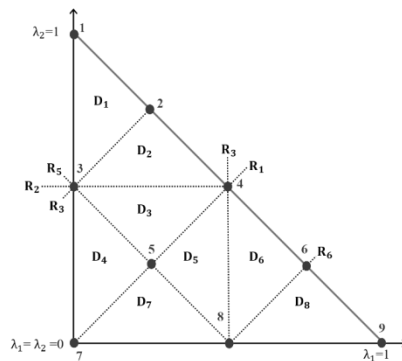
Each value of any arbitrated alternatives follows the adopted Likert scale.

Thus, if the decision maker selects  $I_{ik}, D_{ik}$  or  $D_{ki}$  then information can be elicited about the sub-domain containing his preference and so the OptionCards method is based on a sequence of questions comparing arbitrated alternatives so that the sub-domain containing the point

belonging to  $S_{1,2}$  and describing the weights selected by the decision maker will be progressively narrowed.

Each binary question is presented by an Option Card, OC(1) to the decision maker including two alternatives ( $a_m, b_m$ ) who can select the option ( $(a>b), (a=b), (a<b)$ ) and the next Option Card, OC(m+1) will be presented in terms of such answer.

The questions to be asked should concern the comparisons related to the subspaces indicated in Figure 3 where the indifference lines ( $R_1, \dots, R_6$ ) and the subspaces ( $D_1, \dots, D_8$ ) are represented.



**Figure 3 The subdomains of the weights space**

Thus, the sequence of questions to be asked is as follows:

- 1) OC(1): the indifference line to be considered,  $I_{ab}$  corresponds to  $R_1$  and so  $D_{ab}$  should include  $(D_1 \cup D_2 \cup D_3 \cup D_4)$  and  $D_{ba}$  should include  $(D_5 \cup D_6 \cup D_7 \cup D_8)$ . Furthermore, if the answer corresponds to  $I_{ab}$  then the weights of the decision maker should satisfy the equation of  $R_1$ .

The alternatives  $a, b$  can be freely arbitrated providing that they should respect the equation of  $I_{ab}$ . The adopted alternatives are:

- a:  $U_{a1}=4; U_{a2}=1; U_{a3}=2$
- b:  $U_{b1}=1; U_{b2}=4; U_{b3}=2$

- 2.1) If  $D_{ab}$  from OC (1), then OC(2.1) should be based on the indifference line  $I_{ab}$  represented by  $R_2$  and so  $D_{ab}$  corresponds to  $(D_1 \cup D_2)$  and  $D_{ba}$  corresponds to  $(D_3 \cup D_4)$ . If the answer corresponds to the indifference, then the equation corresponding to  $R_2$  should be respected.

The arbitrated alternatives are:

- a:  $U_{a1}=4; U_{a2}=1; U_{a3}=4$
- b:  $U_{b1}=1; U_{b2}=4; U_{b3}=1$

- 2.2) If  $D_{ba}$  from OC(1), then OC(2.2) should be based on the indifference line  $I_{ab}$  represented by  $R_3$  and so  $D_{ab}$  corresponds to  $(D_5 \cup D_7)$  or if  $D_{ba}$  then one has  $(D_6 \cup D_8)$ . If the answer corresponds to the indifference then the equation of  $R_3$  should be respected.

The arbitrated alternatives are:

- a:  $U_{a1}=4; U_{a2}=1; U_{a3}=1$
- b:  $U_{b1}=1; U_{b2}=4; U_{b3}=4$

- 3.1.1) If  $D_{ab}$  from OC (2.1), then OC(3.1.1) should be based on the indifference line  $I_{ab}$  represented by  $R_4$  and so  $D_{ab}$  corresponds to  $D_1$  and if  $D_{ba}$  then one has  $D_2$ . If the decision

maker chooses the indifference, then R4 should be respected.

The arbitrated alternatives are:

a:  $U_{a1}=3; U_{a2}=1; U_{a3}=3$

b:  $U_{b1}=1; U_{b2}=3; U_{b3}=1$

3.1.2) If  $D_{ba}$  from OC (2.1), then OC(3.1.2) should be based on the indifference line  $I_{ab}$  represented by R5 and so  $D_{ab}$  corresponds to (D3) and if  $D_{ba}$  then one has (D4). The equation related to R5 should be respected if the answer corresponds to indifference.

The arbitrated alternatives are:

a:  $U_{a1}=4; U_{a2}=4; U_{a3}=1$

b:  $U_{b1}=1; U_{b2}=1; U_{b3}=4$

3.2.1) If  $D_{ba}$  from OC (2.2), then OC(3.2.1) should correspond to the indifference line  $I_{ab}$  represented by R5 and so  $D_{ab}$  corresponds to D5 and  $D_{ba}$  to D7. The equation relate to R5 should be respected if the answer corresponds to indifference.

The alternatives corresponding to R5 were already arbitrated.

3.2.2) If  $D_{ba}$  from OC (2.2), then OC(3.2.2) corresponds to the indifference line  $I_{ab}$  represented by R6 and so if  $D_{ab}$  then one has D6 and if  $D_{ba}$  then one has D8. The equation related to R6 should be respected if the answer corresponds to indifference.

The arbitrated alternatives are:

a:  $U_{a1}=3; U_{a2}=1; U_{a3}=2$

b:  $U_{b1}=2; U_{b2}=4; U_{b3}=3$

Obviously, further subdivisions of the space  $S_{1,2}$  can be progressively done using further Option cards.

The application of this method is quite easy as it just implies the presentation of the option cards and receiving the answers following the enquires tree presented in Figure 5.

Finally, the estimated of the weights of the decision maker is obtained by averaging the coordinates of the extreme points of the sub-domain identified and such coordinates are presented in Table. 11.

Subdomain	$\lambda_1$	$\lambda_2$
P(D1)	1-Dec	3-Apr
P(D3)	1-Apr	7-Dec
P(D4)	1-Apr	5-Dec
P(D5)	1-Dec	1-Apr
P(D6)	5-Dec	1-Apr
P(D7)	7-Dec	1-Apr
P(D8)	1-Apr	1-Dec
P(D8)	3-Apr	1-Dec

**Table 11 Average coordinates of each subdomain**

### 3. APPLICATION

The proposed model is applied to the evaluation of the 14 most visited e-commerce websites of the Portuguese market, excluding booking.com because traveling was severely affected by

## COVID19.

A focus group of 43 young and qualified respondents was used to estimate the evaluation of each website according to the presented 9 sub-criteria and the estimated average scores obtained in 2020 are presented in (Table. 12).

Website	I1	I2	I3	I4	I5	I6	I7	I8	I9
Worten	3.75	3.5	4.25	3.75	3	4.5	3.75	4	3.25
Fnac	3	5	5	5	1	5	5	5	5
Continente	2	3.25	5	4	5	5	5	5	5
Aliexpress	5	4.25	5	5	5	5	2.5	4	3
Amazon	4	4	5	3	5	5	5	5	3
Leroy Merlin	5	4.5	5	3	3	5	5	5	4
Bershka	5	4.5	3	3	3	3	1	5	3
Ikea	2	3.75	4.5	4.5	3	4.5	3.5	3.5	3.5
Wook	5	4	4	5	3	5	5	5	4
Ebay	3	4	5	5	5	5	1	5	1
Zara	1	4	5	5	3	5	4	5	2
Stradivarius	2.5	4	4.5	4	3	5	2	3.5	1.5
Mango	2.5	4.25	4.5	4.5	3.5	4.5	3	3	3.5
Auchan	5	4.5	3	5	1	5	2	5	4
Average	3.64	4.17	4.5	4.27	3.31	4.82	3.45	4.67	3.4

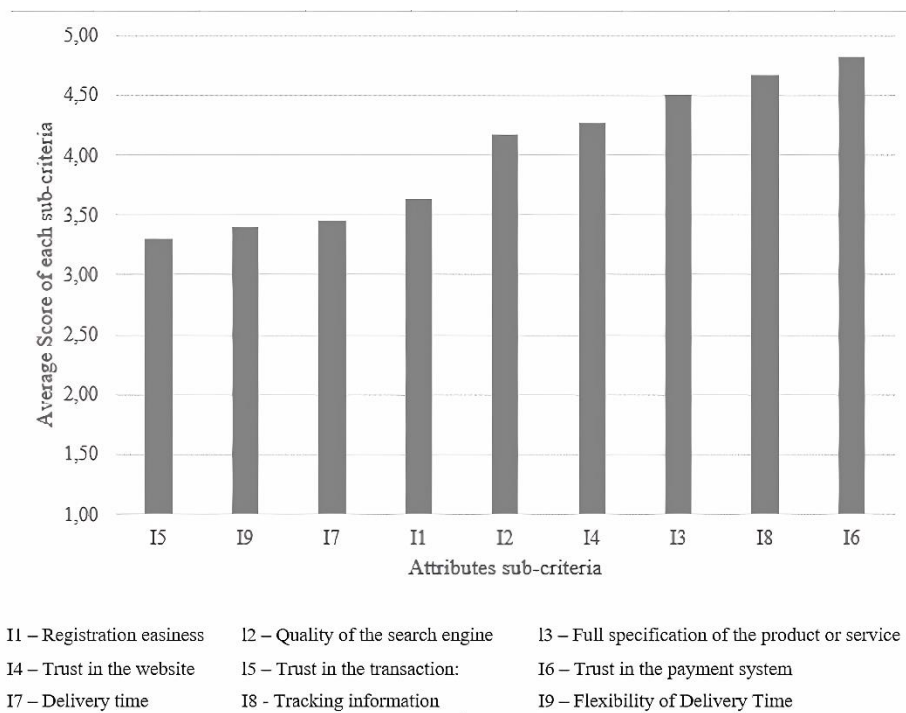
**Table 12 Website Evaluation sub-criteria**

The average score of each sub-criteria is also presented in the last row showing that the full specification of products, the level of security of the payments as well as tracking information are highly rated by the respondents while the level of confidence of the transaction, the delivery time and its flexibility are subpar next to the other attributes (Figure. 4).

The coefficient of variation of the scores assigned by the respondents to each site according to each criterion, CV, is low and under 10% meaning that the limits, L, for the confidence interval (95% level) concerning the estimated mean, M, for each site and subcriterion are given by :

$$L = M. \left( 1 \pm \frac{1.95 \cdot CV}{\sqrt{43}} \right) = M. (1 \pm 0.03)$$

Confirming that a larger sample is not required.



**Figure 4 Average Score of each sub-criteria**

A broader focus group of 131 respondents including not just young professionals but also university students (bachelor or master level), was used to estimate the criteria weights because the estimation of their preferences does not imply to be frequent consumers and so the students group was considered relevant and included in this sample. This study is carried out using the OptionCard Method implemented through an interactive software.

The following results were obtained:

- a) The average weights for navigation, trust and delivery are 0,33; 0,34 and 0,33 respectively;
- b) The estimated coefficients of variation are under 10% and so the confidence limits, L (95% confidence level) for the estimated means, M, are quite close to each other meaning:  $L=M.(1\pm 0.017)$ .
- c) The variation of weights in terms of scientific areas ranges from 0.31 to 0.37 (Table. 13).

Gender	$\lambda_1$	$\lambda_2$	$\lambda_3$
F	0,32	0,32	0,36
M	0,33	0,36	0,31

**Table 13 Criteria weights according to gender**

- d) The variation of weights in terms of scientific areas ranges from 0.28 to 0.38 (Table. 14).

Major Scientific Domain	$\lambda_1$	$\lambda_2$	$\lambda_3$
Arts (A)	0.37	0.31	0.32
Management and Economics (B)	0.28	0.36	0.36

Social Sciences and Law (C)	0.3	0.32	0.38
Technology (D)	0.35	0.36	0.29
BioSciences (E)	0.34	0.32	0.34

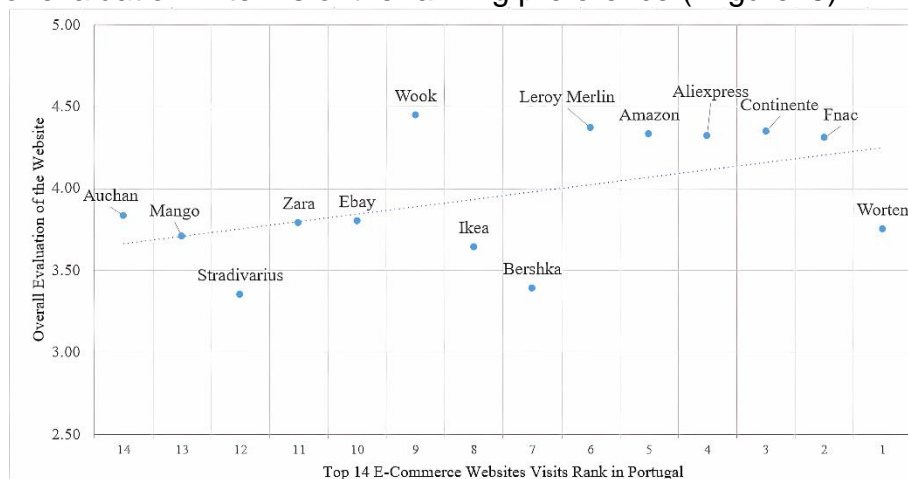
**Table 14 Criteria weights according to scientific area**

Meaning that the weights variations are not statistically significant. The overall evaluation of each website according to the preferences of each evaluator can be now estimated using the estimated averages of  $\lambda_1=0.33$ ,  $\lambda_2=0.34$  and  $\lambda_3=0.33$ , with the results presented in Table 15:

Website	$v_i$
Worten	3.75
Fnac	4.33
Continente	4.37
Aliexpress	4.31
Amazon	4.33
Leroy Merlin	3.38
Bershka	3.64
Ikea	4.44
Wook	3.8
Ebay	3.79
Zara	3.79
Stradivarius	3.34
Mango (Outlet)	3.7
Auchan	3.83

**Table 15 Overall website evaluation**

Finally, the relationship between such overall evaluation and the ranking of the number of visits of each site is presented in Figure 5 showing that the most visited websites deserve quite positive evaluations according to our model, ranging from 3.34 to 4.44 and that there is a positive slope of evaluation in terms of the ranking preference ( Figure. 5).



**Figure 5 Relationship between overall evaluation and visits ranking**

## Conclusion

- The evaluation of websites of e-commerce has been extensively studied but not giving

the required attention to the consumer satisfaction function because the relative preferences of the consumers have not been conveniently studied.

- b) The proposed model–Trident–is based on the MAUT and the proposed criteria cover the key dimensions of navigation, trust and delivery.
- c) This model includes an original approach–The OptionCard Method–to facilitate the estimation of the weights assigned by the consumers to the proposed criteria and therefore the evaluation results of the websites are not independent of the studied consumers.
- d) The proposed model is successfully applied to the evaluation of the 14 most popular websites of e-commerce in Portugal using an interactive software to implement the OptionCard method and the results are also presented herein.
- e) The application of the OptionCard Method was very well accepted by the 131 respondents because it takes no more than 6 minutes as it just requires the successive comparisons concerning the three levels of OptionCards.
- f) The obtained results confirm that the most popular websites have quite positive global evaluations and that there is a general positive trend between such evaluations and the ranking position of each website.
- g) Summing up, the proposed instrument is an appropriate metric tool to evaluate the websites for e-commerce in terms of the specific target group of consumers to be considered.

## Acknowledgements

This research has been developed by COMEGI: Research Centre on Organizations, Markets and Industrial Management of the Universidades Lusíadas (Foundation Minerva) (<http://comegi.ulusiada.pt/>), CiTUA: Center for Innovation in Territory, Urbanism, and Architecture of IST-University of Lisbon (<http://citua.tecnico.ulisboa.pt/>), OPET: Portuguese Observatory of Technology Foresight (<https://www.opet.pt/>), APMEP: Portuguese Association of Public Markets (<http://www.apmep.pt/>) and VORTAL, Connecting Business <https://www.vortal.biz/>.

This project received continuous support of a Steering Group including the following colleagues: Ruben Assis (Vortal Portugal), José Luís Arístegui (Vortal Spain), Pedro Coimbra (Fund for Municipalities Development), Marco Coelho (Municipality of Sintra), Ana Sá (IST, University of Lisbon), Gonçalo Mendes (OPET and APMEP), Vasco Moreira (OPET and APMEP) and João Branco (Electricity of Portugal, EDPC) plus other invited colleagues from other corporations and municipalities. This Steering Group has been chaired by Luís Valadares Tavares and José Antunes Ferreira.

## References

1. eMarketer. “2017 Global number of digital buyers 2014-2021”., 2020.
2. Denic. “2020 Domain numbers of the largest TLD 2020”., 2020.
3. W3Techs. “2020 Usage of top level domains for websites 2020”., 2020.
4. Balaram P., Kosalram K. “E-Commerce Evaluation and E Business Trends. International”, IJIEEB., vol. 4, no. 4, pp. 9–16, 2012.
5. Eurostat. “2019 E-commerce statistics for individuals”., 2020.
6. Chiou W C., Lin C C., and Perng C. “A strategic framework for website evaluation based on a review of the literature from 1995–2006”, Information & Management., vol. 47, no. 5, pp. 282–290, 2010.
7. Goi C L. “A review of web evaluation criteria for e-commerce web sites”, The Journal of Internet Banking and Commerce. J.I.B.C., vol. 17, pp.1–10, 2012.



8. Corley JK., Jourdan Z., and Ingram WR. "Internet marketing: A content analysis of the research", *Electronic Markets.*, vol. 23, 2013.
9. World Best Enterprises. "2004 quality criteria for website excellence".
10. Sørsum H., Andersen K N., and Vatrapu R. "Public websites and human-computer interaction: an empirical study of measurement of website quality and user satisfaction", *Behaviour & Information Technology.*, vol. 31, no. 7, pp.697–706, 2011.
11. 11. Mangiaracina R., Brugnoli G., and Perego A. "The eCommerce customer journey: A model to assess and compare the user experience of the ecommerce websites"., *Journal of Internet Banking and Commerce. J.I.B.C.*, vol.14, pp.1-11, 2009.
12. 12. Barnes S.,and Vidgen R. "An integrative approach to the assessment of e-commerce", *Journal of Electronic Commerce Research.*, vol. 3, pp.114–127, 2002.
13. 13. Kaya T. "Multi-attribute evaluation of website quality in e-business using an integrated fuzzy AHPTOPSIS methodology", *International Journal of Computational Intelligence Systems.*, vol.3, no. 3, pp.301–314, 2010.
14. 14. Kincl T., and Štrach P. "2012 Measuring website quality: asymmetric effect of user satisfaction", *Behaviour & Information Technology.*, vol. 31, No. 7, pp.647–657.
15. 15. Fishburn P C. "Utility theory for decision making", Wiley, 1970.
16. 16. Dyer J S. "Multiattribute Utility Theory MAUT", *Multiple Criteria Decision Analysis.*, pp.285–314, 2016.
17. 17. Roy BI. "Paradigms and challenges in Greco S, Ehrgott M, Figueira JR (eds.)" *Multiple Criteria Decision Analysis. State of the Art Surveys.*, pp.19–39, 2016.
18. 18. Tavares L V. "The TRIDENT approach to rank alternative tenders for large engineering projects", *Foundations of Control Engineering.*, vol. 9, no. 4, pp.181–193, 1984.
19. 19. Keeney R L. "Value-focused thinking: A path to creative decision making", Harvard University Press, 2009.
20. 20. Pedro Maia G., Tavares L V. "Optioncards: An interactive approach to elicit qualification and award criteria weight in Piga G, Treumer S (eds.)", *The Applied Law and Economics of Public Procurement* Routledge., 2013.
21. 21. Likert R. "A technique for the measurement of attitudes", *Archives of Psychology.*, vol. 140, pp. 1–55, 1932.
22. 22. Ingaldi M., Ulewicz R. "How to make e-commerce more successful by use of Kano's model to assess customer satisfaction in terms of sustainable development," *Sustainability*, vol. 11, no. 18, pp.1–22., 2019.
23. 23. European Union: European Commission. "Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce')", *Official Journal of the European Union.* vol.178, pp.1–16, 2000.
24. 24. Mousseau V. "Are judgments about relative importance of criteria dependent or independent of the set of alternatives ? An experimental approach", *Cahier du Lamsadevol.*111, 1992.
25. 25. Keeney R L., and Raiffa H. "Decisions with multiple objectives: Preferences and value trade-offs",

1976.

26. 26. Netrica. "Behavioral panel book. Netquest", 2020.