Journal of Supply Chain Management Science

https://journals.open.tudelft.nl/jscms

http://dx.doi.org/10.18757/jscms.2020.5454

Technology acceptance and return management in apparel e-commerce

Ruchika Kalpoe [a]

[a] *Faculty of Technology, Policy and Management, Delft University of Technology, 2628 BX Delft, The Netherlands*

Article history: received 13-11-2020, accepted 30-12-2020

Abstract – Returns management, especially in apparel e-commerce, has gained increased attention due to the ecological and economic implications it imposes. However, research which explores the relationship between (i) reasons which drive customers' apparel returns and (ii) customer-based instruments designed to reduce online apparel returns, has not yet been empirically examined in literature, especially from the point of view of customers. This research aims to examine the customers' technology acceptance of four technological alternatives designed to prevent unnecessary apparel returns. To determine the customers' technology acceptance, the Technology Acceptance Model (TAM) is used. To operationalize TAM, a Multi-Criteria Decision-Analysis (MCDA) approach is applied, wherein the Bayesian group Best-Worst Method (BWM) is used to infer the weights of the indicators (i.e., criteria) that contribute to the customers'(users') technology acceptance. This is done within the context of apparel e-commerce and with the application of an online BWM survey and expert interviews. The results show that reliable fit & size information is the most important sub-indicator contributing to the customers' technology acceptance. Furthermore, it seems that whilst per subsequent alternative, the reliability of information provision regarding apparel attributes increases, the perceived user-friendliness (ease of use) of the technologies decreases, privacy and security concerns increase, and the managerial implications increase as well.

Keywords: Returns Management; Apparel e-commerce; Technology Acceptance Model; Multi-Criteria Decision-Analysis; Bayesian Best-Worst Method

1. Introduction

Nowadays, more people are purchasing apparel online instead of in physical shops. As a result of this growing apparel e-commerce business, the number of apparel returns is also increasing (Minnema et al. 2016). For instance, approximately 30 percent of online purchased products in the Netherlands are returned (Minnema et al. 2016), of which, 40 percent are apparel items (Edwards et al. 2010).

Aside from unsuccessful purchases and a reduced amount of revenues for apparel e-commerce retailers, as indicated by studies conducted by for example Griffiset al. (2012), other negative societal implications are inflicted as well. For example, more retours result in more use of transportation vehicles, hence more CO_2 emission, less traffic safety, reduced air quality and overall living environment of cities and more traffic congestion. Due to these ecological and economic implications, returns management has gained increased attention, especially in the online apparel retail domain (Difrancesco et al. 2018).

In order to reduce the number of returns and the negative externalities, Walsh et al. (2014) proposed three categories of preventive instruments to reduce return rates, namely (i) monetary instruments, (ii) procedural instruments and (iii) customer-based preventive instruments. According to Walsh et al. (2014), the distinction between the three instrument categories is necessary to study the performance of each preventive instrument more effectively. According to Walsh et al. (2014), monetary instruments "are aimed at financially disincentivizing (or financially incentivizing) customers from returning (retaining) products". Furthermore, "procedural instruments are designed to either reduce transparency (in relation to the return process) for customers, to identify 'return sinners' and to increase the efficiency of the order and delivery process" (Walsh et al. 2014). According to the authors, "customer-based instruments attempt to increase the ease of the order process from the consumer perspective by reducing consumers' perceived pre-purchase uncertainty" (Walsh et al. 2014).

However, literature studies conducted by Walsh and Möhring (2017) and Walsh et al. (2014) indicate that prior research has mainly focused on monetary instruments and that existing research about procedural instruments and mostly customer-based preventive product return instruments is sparse. Based on a literature study regarding apparel returns and preventive instruments, the observation could also be made that so far, many studies have mostly focused on addressing the logistic problems post purchasing. The results have shown that not much empirical research has been conducted so far on how to prevent apparel returns pre-purchasing or during the online screening/evaluation process of apparel items. Consequently, research regarding technologies and instruments which can be used to influence the customers' online pre-purchase decision in order to prevent unnecessary apparel returns is lacking. The literature study has also shown that no empirical studies have yet been carried out within the online apparel retail domain which examines and compares the perceived effectiveness of various customer-based technological concepts in addressing online purchased apparel returns and examine the perceived users' acceptance, especially from the customers perspective.

Therefore, this paper aims to address this gap by analyzing various technological alternatives designed to prevent unnecessary returns of online purchased apparel items. Since the technologies are designed to be used by customers, its success relies greatly on the customers usage. In other words, the customers' acceptance towards these technologies will determine the impact on unnecessary apparel returns. Therefore, the research is mainly approached from the users (customers) perspective. As a result, this paper also aims to explore what the customers' acceptance is regarding the technological alternatives. This is done by applying a more qualitative approach and operationalization of the Technology Acceptance Model (TAM), developed by Davis (1986). The Bayesian group Best-Worst Method (BWM) developed by Mohammadi and Rezaei (2020), is used to operationalize TAM, through which the users' technology preference is determined. Since this research is conducted in the empirical setting of apparel e-commerce, the applicability and reliability of the applied approach is presented in the apparel e-commerce sector.

Consequently, on a practical level, this research contributes to the apparel e-commerce sector by designing and analyzing technological alternatives which assists customers during the online screening process of apparel items, such that they can more accurately evaluate apparel items so that their online apparel purchase success increases and unnecessary apparel returns are prevented. Furthermore, on a scientific and methodological level, this research contributes to the knowledge of operationalizing the TAM in the empirical setting of apparel e-commerce, by identifying 11 evaluation criteria (i.e., indicators) which play a role in the users' (online shoppers') technology acceptance. To measure the customers' preference regarding the technological alternatives, the Bayesian BWM is applied. Lastly, as a result of the Multi-Criteria Assessment, decision-makers of online apparel retail shops can decide if and how they can adapt current arrangements in order to reduce online apparel returns.

The remainder of this paper is structured as follows. In Section 2, the theoretical model (TAM) is explained in virtue of which the customers' acceptance regarding the various technologies is determined. The applied qualitative approach used to operationalize TAM is then described in Section 3. In Section 4, the application of the research approach and the results are presented, followed by a discussion in Section 5. At last, in Section 6, the conclusion and recommendations are presented.

2. Technology Acceptance Model

In order to understand the users' acceptance or rejection of technologies, various theories exist in literature which can be used, such as the theory of reasoned action (TRA) developed by Fishbein and Ajzen (1975), the theory of planned behavior (TPB) developed by Ajzen (1985) and the Technology Acceptance Model (TAM) developed by (Davis 1986). To predict users' behavioral intention, which measures the likelihood of a behavior occurring, TRA uses the determinants relative importance of attitudes (i.e., the users' feeling towards a particular behavior) and subjective norms (i.e., the way in which the perceived social pressure from others affects the users' performance and behavior) (Fishbein and Ajzen 1975). To improve the predictive power of TRA, the TPB was developed by Ajzen (1985), wherein the additional determinant perceived behavioral control (i.e., the users' perceived control over expressing their own behaviors and attitudes) was included as well to predict the users' behavioral intention.

However, Davis (1986) observed that this additional determinant did not have a high correlation with the use of technologies. As a result of this, TAM was developed, to predict the technology acceptance, wherein subjective norm and perceived behavioral control were excluded and the actual use (i.e., behavior) was simply predicted by solely using two determinants which are: (i) the Perceived Usefulness (PU) and (ii) the Perceived Ease of Use (PEU) (Davis 1986). Since this model solely predicts the acceptance of technologies via two determinants PU and PEU, this model suits the practical research goal better compared to the aforementioned theories. As a result, the TAM developed by Davis (1986), which is a very prominent model in explaining the technology acceptance (especially for information technologies) is used. Figure 1 provides an overview of the TAM. The X1, X2, and X3 are the external predictors used to measure the two determinants (PU and PEU).

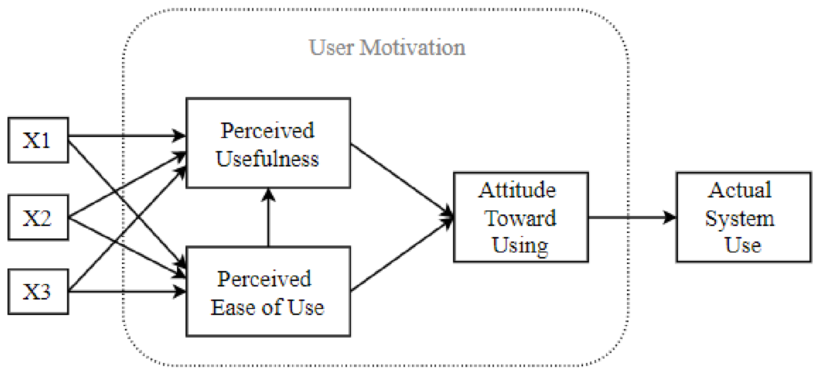

Figure 1. Technology Acceptance Model (retrieved from Davis (1986))

Perceived usefulness (PU) is defined as "the degree to which an individual believes that using a particular system would enhance his or her job performance" (Davis 1986). Consequently, PU expresses the effectiveness of a technology in addressing a specific function. Perceived ease of use (PEU) is defined as "the degree to which an individual believes that using a particular system would be free of physical and mental effort" (Davis 1986). According to Davis (1986), "PEU plays a crucial role in understanding an individual's response to information technology".

TAM finds its popularity in its simplicity, as it solely uses two determinants to predict technology acceptance, PU and PEU, which makes the model highly versatile and easy to apply (Vogelsang et al. 2013). In literature, other frameworks exist which can be used to evaluate the customers' acceptance regarding technologies, e.g., the Feitelson and Salomon (2004) framework and the multi-level perspective on technology transitions framework developed by Geels (2004). However, since these frameworks also include determinants such as political, institutional, and financial to predict the technology success, they are deemed too broad for this research goal, which is to solely determine the customers preference regarding various technologies.

In the literature, TAM has been mainly criticized on its predictive validity, as it is perceived as incomplete since it predicts the acceptance based on solely two determinants (Legris et al. 2003, Chuttur 2009). Based on an extensive literature study conducted by Marangunić and Granić (2015) regarding the application of TAM, many changes and extensions of the TAM were identified to increase the predictive validity of the model. Therefore, to increase the predictive validity of TAM in this research, the decision was also made to examine additional determinants to predict the customers' acceptance regarding the technological alternatives. This was done through a literature study regarding TAM and the inclusion of experts' opinion with a background in academia and the apparel e-commerce industry. In Section 4, this is further elaborated.

3. Methodology

Based on the extensive literature study conducted by Marangunić and Granić (2015) and a review of literature regarding TAM, it becomes clear that almost all publications used Structural Equation Modeling (SEM) to operationalize TAM, implying that the fraction of qualitative approaches is still very small. SEM is a statistical approach, mostly used in the field of psychology (Nachtigall et al. 2003). It can be considered as a combination of factor analysis, multiple regression analysis and path modelling, and it is applied to evaluate the structural relationship/correlation between indicators (measured variables) and latent variables (non-observable variables) without measurement error (Hox and Bechger 1999, Nachtigall et al. 2003). Since non-observable variables such as attitude toward using a technology and the users' acceptance of technologies cannot be measured directly, indicators (observable variables) through which they can be measured are required.

However, since the aim of this study is to explore the customers' acceptance regarding various technological alternatives based on evaluating a set of indicators (criteria), and not to determine the correlation between the indicators used to predict the users' technology acceptance, a more qualitative approach was used to operationalize TAM. Consequently, within this research an MCDA approach is applied whereby TAM is used as theoretical foundation to identify and theoretically underpin the indicators (criteria) which are necessary to evaluate the customers' technology acceptance and rank the technological alternatives in the context of apparel e-commerce returns management. An MCDA approach was applied, in order to quantify the importance of indicators (criteria) and determine which indicators are perceived as the most important for achieving users' technology acceptance. In the following section, this approach is described.

3.1. MCDA approach

Matrix (1) indicates the general form of an MCDA approach for the evaluation of a set of alternatives $\{a_1, a_2, \dots, a_m\}$ based on a set of decision-criteria $\{c_1, c_2, \dots, c_n\}$ and p_{ij} is the score of each alternative i with respect to each criterion j. The goal is to rank the alternatives and select the best one.

$$A = \begin{array}{c} \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{array} \begin{array}{c} \begin{array}{cccc} c_1 & c_2 & \cdots & c_n \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{pmatrix} \end{array} \quad (1)$$

By using for example, the additive value function, as presented in equation 2, the overall value of alternative i presented as V_i can be calculated. When for example the weight w_j is assigned to criterion j, then V_i is simply determined by multiplying the score p_{ij} with the respective weight w_j of criterion j $\left(w_j \geq 0, \sum w_j = 1\right)$ (Keeney and Raiffa 1976).

$$V_i = \sum_{j=1}^{v} w_j p_{ij} \quad (2)$$

Following this MCDA approach, first a set of alternatives is needed followed by a set of decision-criteria by which the alternatives can be evaluated. Then, by using a preference elicitation method, the criteria weights should be established. In literature, a variety of methods exist which can be applied to infer the criteria weights. In the next section the applied method is described.

3.2. Bayesian BWM

In this research, the BWM is applied as preference elicitation method, since it (compared to other MCDA methods such as the analytic hierarchy process (AHP)) (i) requires less comparisons data and (ii) leads to more consistent comparisons, implying that it produces more reliable weights (Rezaei 2015, Rezaei 2020). When using the BWM, $2n - 3$ comparisons are required (Rezaei 2015), while using AHP, the number of comparisons needed is $n(n-1)/2$ (Saaty 2004). Since the decision-maker chooses a best and worst criterion before conducting the pairwise comparisons when using BWM, a clear understanding regarding the range of evaluation is gained upfront which could lead to more consistent pairwise comparisons, hence more reliable weights (Rezaei 2020). Over the years, BWM has gained increased attention and has been used in various fields of study (bestworstmehod.com), such as ecosystem data governance (de Prieëlle et al. 2020), supplier selection in online fashion retail (Kaushik et al. 2020), circular economy (Moktadir et al. 2020), supply chain sustainability innovation (Gupta et al. 2020), and crowdsourcing delivery personnel (Li et al. 2020) to name a few.

Since this research examines what the technology acceptance is from the perspective of users (customers), a group decision-analysis version of BWM, namely Bayesian BWM is used to operationalize TAM.

The Bayesian BWM uses the same input data as the original BWM, see Step 1 till Step 4, as provided by Rezaei (2015). However, the application of the last Step (Step 5) which consists of computing the criteria weights differs when using the Bayesian BWM. The BWM Steps as provided by Rezaei (2015) are described below:

Step 1. Establishing a set of decision- criteria.

The first step of the BWM is to identify a set of n decision criteria $(c_1, c_2, c_3 \dots c_n)$ which the decision-maker can use to evaluate the designed alternatives.

Step 2. Defining the Best criterion and the Worst criterion.

In the second step, the decision-maker chooses the best (most important or most preferable) criterion and the worst (least important or least preferable) criterion.

Step 3. Obtaining the Best-to-Others (BO) comparison vector.

In the third step, the decision-maker determines the preference of the best (most important) criterion against all other criteria by using a scale from 1-9. A value of 1 implies that the two criteria are of equal importance, whereas a 9 suggests that the best criterion is absolutely more important than the other one. As a result, a BO vector is obtained:

$A_B = (a_{B1}, a_{B2}, a_{B3}, \dots, a_{Bn})$, where a_{Bj} is the preference of the best criterion B over the other criterion j.

Step 4. Obtaining the Others-to-Worst (OW) comparison vector.

In the fourth step, the decision-maker determines the preference of all other criteria over the worst (least important) criterion by using the same scale from 1-9. As a result, an OW vector is obtained:

$A_W = (a_{1W}, a_{2W}, a_{3W}, \dots, a_{nW})^T$, where a_{jW} is the preference of the other criterion j over the worst criterion W.

The Bayesian BWM has the following additional sub-steps which are undertaken in Step 5 to compute the optimal group weights of criteria (Mohammadi and Rezaei 2020).

Step 5. Establishing optimal group weights of criteria.

Step 5.1. Constructing the probability distribution.

Assume that there are k decision-makers *(k =1, 2, ..., K)*, there are n evaluation criteria ($c_j = c_1, c_2, \dots, c_n$), then A_B^k represents the Best-to-Others (BO) vector of decision-maker k and A_W^k the Others-to-Worst (OW) vector of decision-maker k. If the optimal weights of decision-maker k is w^k, the optimal group weight after aggregation is w^{agg}. The vector, $A_B^{1:K}$ represents the BO vectors of all decision-makers and $A_W^{1:K}$ indicates the OW vectors of all decision-makers. Based on this, the equation for the joint probability distribution of the group decision for the Bayesian BWM is formulated as:

$$P\left(w^{agg}, w^{1:K} \middle| A_B^{1:K}, A_W^{1:K}\right) \tag{3}$$

If the probability in (3) is calculated, the following probability rule can be used to compute the probability of each individual variable:

$$PP(x) = \sum_y P(x,y) \tag{4}$$

where, x and y represent arbitrary random variables.

Step 5.2. Calculating the optimal group weight.

The aggregated weight w^{agg} is dependent on the optimal weight of every individual decision-maker w^k, which is calculated by the input BO and OW vectors (A_B^k and A_W^k). The equation for the joint probability of the Bayesian BWM can be presented as:

$$P\left(w^{agg}, w^{1:K} \middle| A_B^{1:K}, A_W^{1:K}\right) \propto P\left(A_B^{1:K}, A_W^{1:K} \middle| w^{agg}, w^{1:K}\right) P\left(w^{agg}, w^{1:K}\right) \tag{5}$$

Equation 5, can further be presented as:

$$P\left(A_B^{1:K}, A_W^{1:K} \middle| w^{agg}, w^{1:K}\right) P\left(w^{agg}, w^{1:K}\right) = P\left(w^{agg}\right) \prod_{k=1}^{K} P\left(A_W^k \middle| w^k\right) P\left(A_B^k \middle| w^k\right) P\left(w^k \middle| w^{agg}\right) \tag{6}$$

Based on Equation 6, the corresponding probability can be found by specifying the distribution of each element. As a result, $A_B^k | w^k$ and $A_W^k | w^k$ can be defined as follows.

$$A_B^k \left| w^k \sim multinomial\left(\frac{1}{w^k}\right), \forall_k = 1, 2, \ldots, K; A_W^k \right| w^k \sim multinomial\left(w^k\right), \qquad \forall_k = 1, 2, \ldots, K. \tag{7}$$

Furthermore, w^k under w^{agg} conditioned can be composed as an underlying Dirichlet distribution:

$$w^k | w^{agg} \sim \mathrm{Dir}(\gamma \times w^{agg}), \ \forall_k = 1,2,\ldots,K \tag{8}$$

where w^{agg} is the averaged value of the distribution and γ is a non-negative parameter.

Since γ is a non-negative parameter, it needs to obey the underling gamma distribution where a and b represents the shape and the scale parameters of the gamma distribution.

$$\gamma \sim gamma(a, b) \tag{9}$$

Ultimately, the aggregated or group optimal weight w^{agg} abides to the Dirichlet distribution, with the parameter α being set to 1.

$$w^{agg} \sim Dir(\propto) \tag{10}$$

Once the probability distribution of all parameters is finalized, the posterior distribution is calculated by using the Markov-chain Monte Carlo (MCMC) technique (Mohammadi and Rezaei 2020).

Step 5.3. Credal ranking and Confidence level.

The Bayesian BWM provides a credal ordering of each and every pair of criteria (c_i, c_j)) for all $(c_i, c_j \epsilon\ C)$, with C being the set of criteria. The confidence level (CL) is computed for each pair of criteria to show how significant the difference between their weights is. The CL thus indicates the probability or confidence (P) that c_I is more preferred than c_j (Mohammadi and Rezaei 2020). To visualize this significance, a weighted directed graph can be utilized. The probability (P) that c_i is more preferred than c_j is computed as follows.

$$\mathrm{P}(c_i > c_j) = \int I(w_i^{agg} > w_j^{agg}) \mathrm{P}(w^{agg}) \tag{11}$$

In Equation 11, I represents a conditional parameter which can only be computed if $\left(w_i^{agg} > w_j^{agg}\right)$ is detained, or else it is zero. Evidently, the CL is obtained by the number of samples Q acquired by the Markov-chain Monte Carlo technique (MCMC).

$$\mathrm{P}(c_i > c_j) = \frac{1}{Q} \sum_{q=1}^{Q} I\left(w_i^{agg_q} > w_j^{agg_q}\right); \mathrm{P}(c_j > c_i) = \frac{1}{Q} \sum_{q=1}^{Q} I\left(w_j^{agg_q} > w_i^{agg_q}\right) \tag{12}$$

In Equation 12, w^{agg_q} represents q w^{agg} from MCMC samples. If $\mathrm{P}(c_i > c_j) > 0.5$, then criterion i is more important than criterion j (Mohammadi and Rezaei 2020). The total probability is equal to one, $\mathrm{P}(c_i > c_j) + \mathrm{P}(c_j > c_i) = 1$.

Through the provided credal ranking and the assigned confidence levels (CL) in the weighted directed graph, the group's perceived importance of one criterion over one another is visualized, which can provide decision-makers (in this case apparel e-commerce decision-makers) with more information on how to adapt current arrangements (Mohammadi and Rezaei 2020).

Compared to SEM, which determines the technology acceptance based upon the relationship between the indicators, this research attempts to determine the customers' technology acceptance through the assigned importance/preference to each indicator (criterion). Consequently, the contribution to technology acceptance is quantified through the computed weights of each indicator (criteria). Criteria with high aggregated weights are considered to have a significant impact on technology acceptance, suggesting that a high degree of users' (customers') technology acceptance could be realized once scoring well on each and every criterion. In the following section, the theory and methodology are applied.

Following the MCDA approach, within this research the following steps were initialized to determine the customers technology preference using the Bayesian BWM as a method to operationalize TAM.

In Section 4, each paragraph is devoted to addressing the four steps of the data collection process (Figure 2).

4. Results

4.1. Set of Alternatives

In order to obtain a set of alternatives, a literature study was conducted as described in Step 1 of Figure 2. Through the literature study, as presented in the previous section, various reasons for customers' apparel returns were identified along with various product return prevention instruments. However, the results showed that research so far has mainly focused on monetary and procedural instruments, and not so much on customer-based instruments, which according to Walsh et al. (2014) "attempt to increase the ease of the order process from the consumer perspective by reducing consumers' perceived pre-purchase uncertainty". As a result, these instruments are treated in this research.

Walsh, et al. (2014) indicate that "the purpose of using these instruments is to communicate suitable information about the product to customers, such that they can evaluate the personal fit more precisely and refrain from returning it because of a possible misfit". As a result, in this research, return reasons were included which can be addressed by these instruments.

The identified drivers of customers online purchase apparel returns were: (i) disconfirmation driven (Saarijärvi et al. 2017), (ii) size-chart driven (Saarijärvi et al. 2017), (iii) feeling driven (Saarijärvi et al. 2017) and (iv) benefit maximization driven (Saarijärvi et al. 2017, Brooks and Brooks 2014, de Leeuw et al. 2016). In Appendix A, an overview of the identified reasons linked to these drivers is presented. Based on the identified reasons for apparel returns, the following four apparel attributes were extracted, which are necessary for customers to evaluate apparel accurately online: (i) material information, (ii) color information, (iii) fit & size information and (iv) style information.

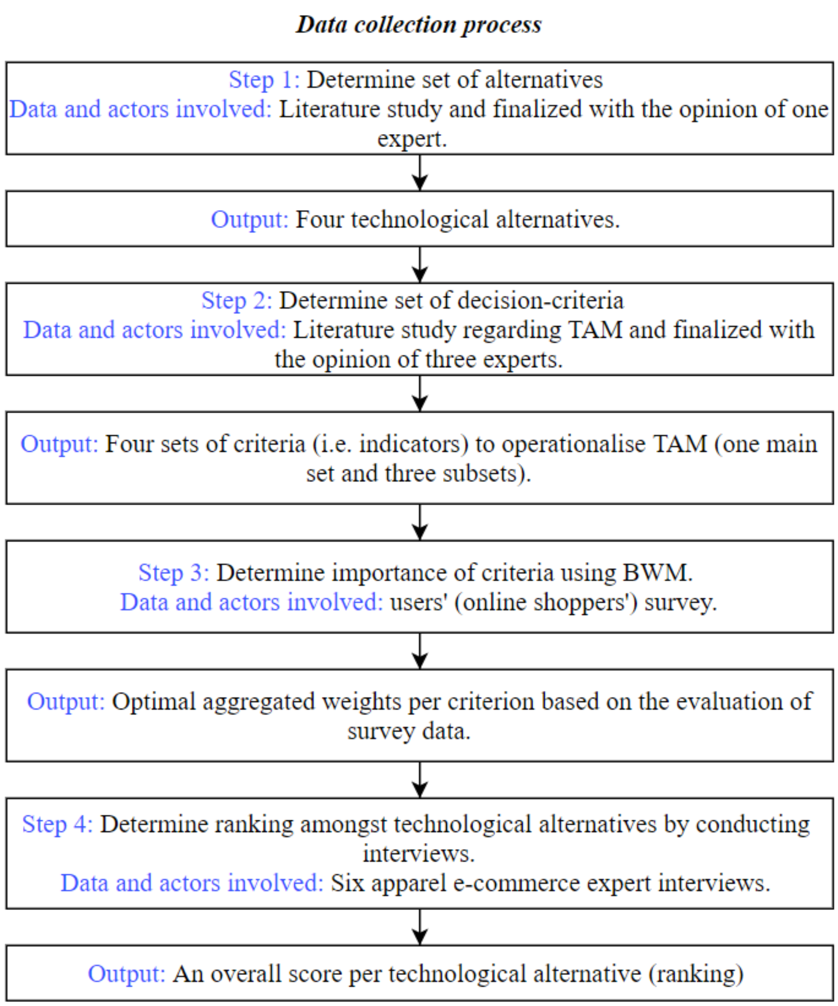

Figure 2. Stepwise data collection process

Furthermore, creating the right expectations regarding the apparel attributes, providing accurate product information, and creating a 'feel' for and perception of apparel items displayed online were also essential requirements that were identified, which the technologies should fulfil in order to prevent unnecessary apparel items. These features were based on the observation that apparel items were returned because the attributes were different from what was expected, the provided information was misleading, style related features were not visible, and customers were unable to get a 'feel' and perception of the apparel items.

This literature study has also led to the identification of various customer-based instruments and technologies such as: height/size chart, fit & size recommendation application, alternative product photo's, mix-and-match function, zoom technologies, avatars and virtual dressing rooms. Since returns management is getting increased attention, a thorough literature study regarding customer-based technologies was required to not only identify traditional instruments, but also examine various novel technologies such as avatars and virtual dressing rooms which are developed to increase online information accuracy and successful purchases as well.

However, since some of these identified instruments on their own cannot provide all the aforementioned requirements, it was necessary to combine some instruments such that they can fulfil the requirements and function as comparable alternatives against the technologies which on their own can fulfil all the requirements. The combined alternatives were based upon the current practices of apparel e-commerce retailers. This was done to provide practical solutions to apparel retailers. The set of alternatives was finalized with the opinion of an apparel quality assurance inspector (Expert 1 indicated in Table 1). The four alternatives, referred to as A1, A2, A3 and A4, are presented as follows.

A1: The bare minimum

A2: The bare minimum with a fit & size recommendation instrument

A3: Avatar (digital computer-based twin)

A4: Virtual Dressing Room (VDR)

In Appendix B, a description of each alternative is presented.

4.2. Set of Criteria

Following the second step of the MCDA approach (as indicated in Figure 2), a set of decision criteria (indicators) to operationalize TAM needs to be established. For the MCDA, criteria can be established through literature research when sufficient literature is available. Otherwise, the criteria can be established through interviews. Within this research, a literature study regarding TAM functioned as input since sufficient literature was available in the e-commerce field. The set particularity used for the apparel e-commerce case was finalized with three experts' opinions, of which two have a background in academia and one is an apparel quality assurance inspector at the fourth biggest e-commerce retailer in the Netherlands and the second biggest online fashion retailer in the Netherlands. Figure 3 gives an overview of the criteria sets.

As indicated in Figure 3, two hierarchy levels exist, namely main criteria (main indicators) and sub-criteria (sub-indicators). The main criteria were established based on the synergy between the identified significant sub-criteria through the literature study regarding TAM.

Based on the description of TAM, provided in Section 2, the original TAM only has two determinants, PU and PEU. However, based on literature study results regarding TAM, trust was also an important determinant and is therefore also included as determinant of technology acceptance. As a result, this research provides an extension of the original TAM. Hence, in this research, the following three determinants of technology acceptance are used: PU, Trust, and PEU.

Based on the literature study regarding TAM, the main criterion 'quality of provided information' is mostly perceived as significant external predictor of the determinant PU, the main criterion 'information gathering and handling' is mostly perceived as significant external predictor of Trust and the main criterion 'user-friendliness' is mostly perceived as significant external predictor of the determinant PEU. Based on the applicability for the online e-commerce case, the decision was made to also include the three main criteria and their sub-criteria as such in the research.

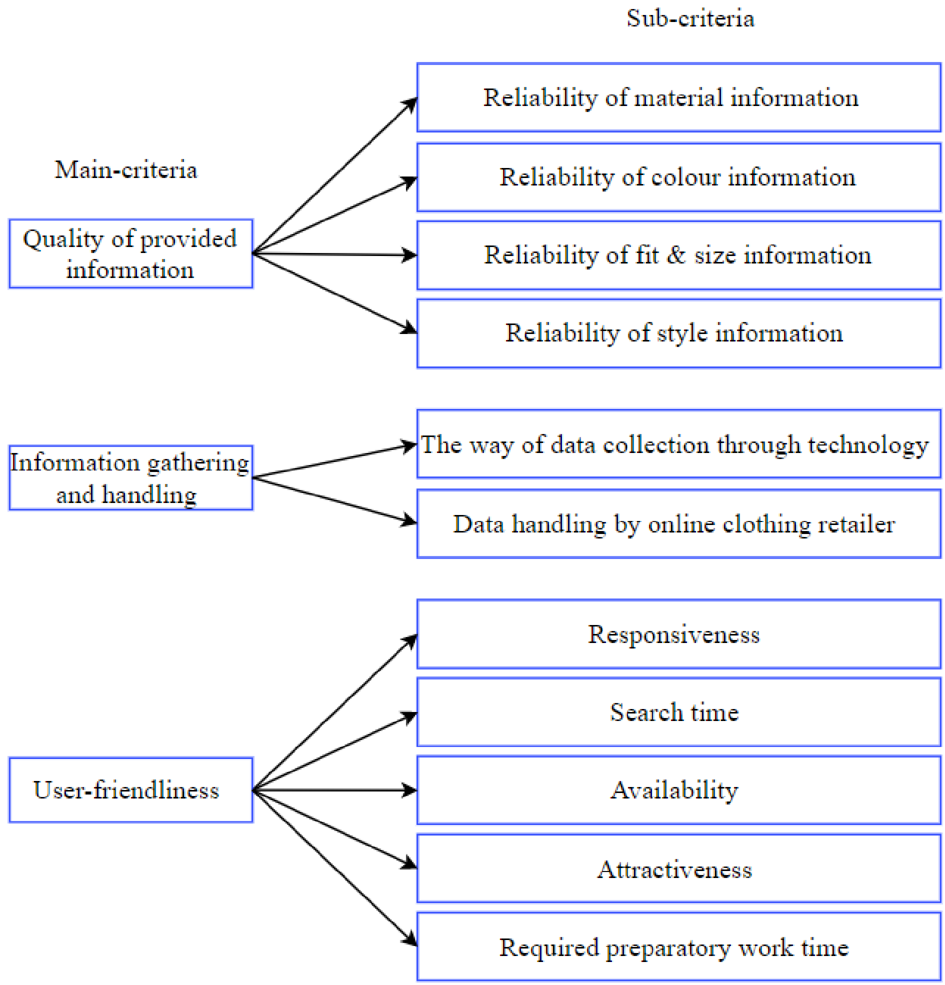

Figure 3. Hierarchy of criteria to evaluate the technologies

4.3. Users (Customers) Preferences

Criteria Weights

Since the technologies are designed to be used by customers, its success relies greatly on the customers usage. In other words, the customers' acceptance towards these technologies will determine the impact on unnecessary apparel returns. Therefore, the research is mainly approached from the users (customers) perspective. The optimal aggregated weight per criterion was established by applying the Bayesian BWM. The input data for the BWM was obtained through an online survey targeted at online apparel shoppers. In total, 216 respondents who have purchased apparel items online were reached.

Before calculating the optimal group weights, the consistency of the respondents was also checked (Liang et al. 2020) and the ones which were acceptable were considered. After excluding the pairwise comparisons with an unacceptable consistency ratio, different sample sizes for different levels of the model were acquired and used. As a result, a sample size of 113 was used to obtain the weights for the main criteria. As indicated in Figure 3, three sets of sub-criteria were analyzed. To obtain the weights of the sub-criteria belonging to the first set, a sample size of 77 was used. A sample size of 113 was used for the second set of sub-criteria and a sample size of 73 was used for the third set of sub-criteria.

Performance Scores

In this research, experts were approached to obtain the performance scores, since experts have the knowledge about the technologies and instruments and how effective each composed alternative is in addressing each criterion. The performance scores were collected through individual structured interviews with six industry experts, stemming from four online apparel retail companies in the Netherlands. To obtain the performance scores, the Bayesian BWM was again applied.

Using the BWM as scoring method resulted in more reliable results, compared to e.g., using a scale from 1 to 10 to obtain the performance scores per alternative with respect to each criterion. However, it was more time consuming to obtain and analyze the data. Since the Bayesian BWM was applied, the obtained scores are weights as well. In Table 3, the obtained performance scores per alternative with respect to each criterion is indicated. Table 1 provides an overview of the interviewed experts.

Table1. Characteristics of interviewees

Expert	Company (anonymized)	Function	Expertise	Years of experience
1	A	Quality Assurance Inspector	Technical translation from styling/design to the technical application and visualization of clothing on the web shop, lead of the returns management project.	5.5 years
2	B	Quality Assurance Inspector	Responsible for the fit & size of apparel and material quality for woman's department.	7 years
3	A	Local marketing manager	Online marketing and retour analysis.	10 years
4	C	Local marketing manager	Omnichannel marketing (physical and digital marketing).	3 year
5	C	Online product specialist	Retour analysis of apparel items, product information optimization.	3 years
6	D	Country online marketing manager	Product recommendations for online apparel items, online marketing campaigns, making the technical translation from styling/design of brands to the technical application and visualization of apparel on the web shop.	2 years

Interpreting Criteria Weights

To quantify the importance of the indicators (criteria) and determine which indicators have the largest contribution to technology acceptance, the Bayesian BWM was applied. Table 2 provides an overview of the obtained weights based on survey respondents.

Table 2. Customers' group weights of main criteria and sub-criteria

Main criteria	Weight	Sub-criteria	Local Weight	Global Weight
C1. Quality of provided information	0.441	c1.1. Reliability of material information	0.242	0.107
		c1.2. Reliability of color information	0.248	0.110
		c1.3. Reliability of fit & size information	0.318	0.140
		c1.4. Reliability of style information	0.192	0.084
C2. Information gathering and handling	0.235	c2.1. The way of data collection through technology	0.432	0.101
		c2.2. Data handling by online clothing retailer	0.568	0.133
C3. User-friendliness	0.324	c3.1. Responsiveness	0.207	0.067
		c3.2. Search time	0.223	0.072
		c3.3. Availability	0.190	0.062
		c3.4. Attractiveness	0.179	0.058
		c3.5. Required preparatory work time	0.201	0.065

Interpreting Main Criteria Weights

Looking at the main indicators, 'quality of provided information' is the most important main indicator for technology acceptance ($w^{agg} = 0.441$). This implies that individuals certainly prefer to obtain reliable apparel attribute information, compared to the perceived ease of use of the technology and the main indicator information gathering and handling. When looking at Figure 4, there can be observed that the criterion 'quality of provided information' has a high confidence level of 1 compared to the other two criteria 'data gathering and handling' and 'user-friendliness', implying that the degree of certainty about the criterion is also evident. In other words, we can be very sure about the superiority of C1 over C3 and C2, that 'quality of provided information' is certainly more important than 'user-friendliness' of the technology and 'information gathering and handling'.

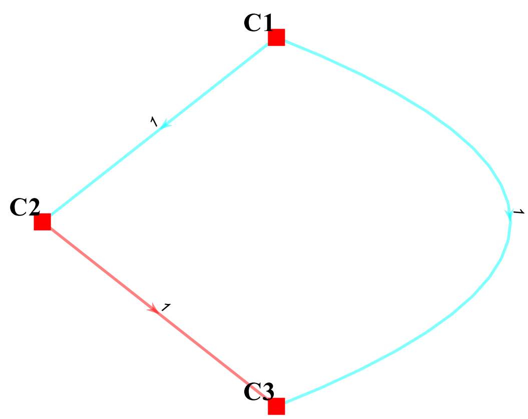

Figure 4. Credal ranking of main criteria

Interpreting Global Weights of Sub-Criteria

The results show that from all 11 sub-indicators, fit & size information is perceived as the most important for technology acceptance ($w^{agg} = 0.140$). This implies that individuals assign high value to obtaining reliable fit & size information. Slightly behind it is 'data handling by online clothing retailer' ($w^{agg} = 0.133$). This implies that the way the online apparel retailer uses and stores the collected information for its services significantly impacts the customers preference and technology acceptance. The third most important sub-indicator is 'reliability of color information' ($w^{agg} = 0.110$), implying that the accuracy, completeness and truthfulness of the provided information regarding the color of apparel items is the third most important sub-indicator affecting technology acceptance. The results also show that 'reliability of material information' and 'the way of data collection through technology' are the fourth and fifth most important sub-indicators for technology acceptance ($w^{agg} = 0.107$ *and* $w^{agg} = 0.101$). This implies that the provision of accurate, complete and truthful information regarding the material of apparel items which refers to material thickness, stretch-ability, texture and stitching (sewing) also significantly contributes to the customers' technology acceptance, followed by the way in which the technology acquires customers information (for example, through scanning, facial recognition or manually inserting body-measurement information).

Looking at the main indicator 'quality of provided information', the sub-indicator 'reliability of fit & size information' is perceived as the most important. Based on the assigned confidence level in Figure 5, the relationship is also evident, suggesting that 'reliability of fit & size information' is certainly more important (CL = 1) than 'reliability of material information', 'reliability of color information' and 'reliability of style information'. On the other hand, Table 2 shows that 'reliability of style information' is perceived as the least important. This implies that individuals who purchase apparel items online are the least interested in obtaining style information to evaluate apparel items online, compared to the other three apparel attributes: material, color, and fit & size information.

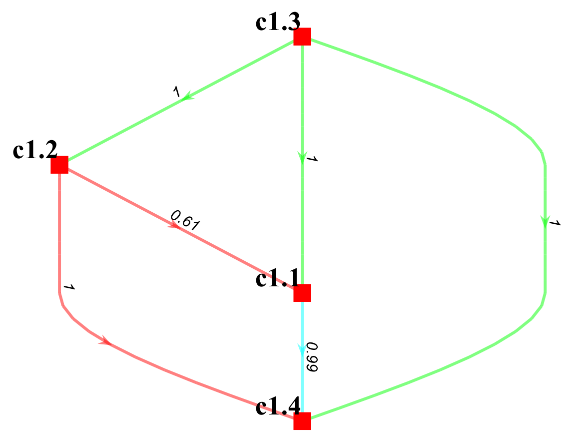

Figure 5. Credal ranking of sub-criteria related to 'quality of provided information'

Furthermore, the observation can be made that all two sub-indicators related to the main indicator 'information gathering and handling' are also perceived as highly important indicators for technology acceptance (aggregated weight higher than 0.1). Furthermore, of all the 11 sub-indicators, 'the way of data collection through technology' is perceived as the second most important indicator for technology acceptance. This implies that customers assign high value to privacy and security concerns, which has a high significant influence on the determinant Trust and through that on the technology acceptance. Looking at the assigned confidence levels in underling Figure 6, 'data handling by online clothing retailer' is certainly more important than 'the way of data gathering through technology' in determining the customers' technology acceptance.

Figure 6. Credal ranking of sub-criteria related to 'information gathering and handling'

Looking at the sub-indicators belonging to the main indicator 'user-friendliness', the results show that 'search time' is perceived as the most important for technology acceptance. This implies that the perceived ease of use mostly relies on the number of clicks/efforts an individual needs to perform when using the technology to evaluate apparel items online. Looking at the assigned confidence levels in Figure 7, 'search time' is more important than all sub-indicators related to user-friendliness, with a confidence level that is higher or equal to 0.81. This implies that 'search time' is certainly perceived as more important in determining the perceived ease of use of the technologies, compared to the other sub-indicators belonging to the main indicator 'user-friendliness'.

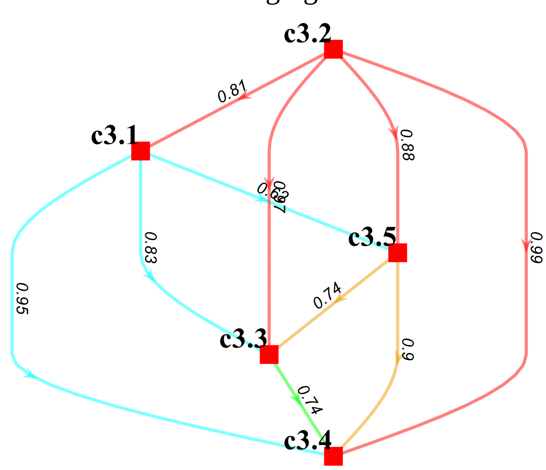

Figure 7. Credal ranking of sub-criteria related to 'user-friendliness'

The second most important sub-indicator related to the main indicator 'user-friendliness', is 'responsiveness', implying that the loading time of the technology is the second most important aspect contributing to the perceived ease of use of the technology. The results show that 'required preparatory work time' is the third most important sub-indicator, implying that the amount of work customers need to do upfront before they can start using the technology to evaluate apparel items with is the third most important indicator influencing the perceived user-friendliness of the technology. The fourth and fifth most important sub-indicators related to the main indicator 'user-friendliness' are 'availability' and 'attractiveness'. This implies that the ability to use the technology on any device is perceived as the fourth important aspect defining user friendliness and that customers technology acceptance is the least influenced by the aesthetic features of the technologies.

Comparing the Alternatives

In Table 3, an overview of the obtained scores (i.e., weights) of each alternative with respect to each sub-criterion are presented. The scores were obtained from six apparel e-commerce expert interviews.

Table 3. Experts' scores (weights) of alternatives with respect to sub-criteria

Sub-criteria	**Local Weights**			
	A1	**A2**	**A3**	**A4**
Reliability of material information	0.140	0.204	0.299	0.358
Reliability of color information	0.217	0.217	0.228	0.338
Reliability of fit & size information	0.093	0.186	0.362	0.359
Reliability of style information	0.106	0.144	0.359	0.392
The way of data collection through technology	0.424	0.325	0.143	0.108
Data handling by online apparel retailer	0.503	0.278	0.123	0.096
Responsiveness	0.438	0.339	0.101	0.122
Search time	0.274	0.332	0.171	0.223
Availability	0.475	0.307	0.096	0.122
Attractiveness	0.104	0.144	0.316	0.436
Required preparatory work time	0.304	0.361	0.151	0.185

Looking at Table 3, the result obtained from the expert interviews show that A3 scores the best with respect to the sub-indicator 'reliability of fit & size information', however closely followed by A4. According to Expert 2, 40% of all apparel returns in the company are indeed a cause of fit & size issues (e.g., the size chart that is not accurate enough so that apparel does not fit) and 40% of all returns also stem from apparel items not being as expected (disconfirmation driven). According to Expert 1, the amount of apparel returns stemming from fit & size issues are even higher, nearly 52% whilst for material, color and style it is 6% for each attribute. According to Expert 4, in total 37% of all apparel is returned as a result of fit & size issues (18 % too small and 19% too big), whilst style is 31% (e.g., the style, when worn, does not look as good as expected) and for material and color the number of returns is a combined 2% (e.g., other hue, or unclear pictures of apparel items). Based on this, the observation can be made that the identified apparel return reasons from literature used in this research are indeed valid, since the literature study has shown that most returns stem from fit & size issues.

The results presented in Table 3 show that A4 scores the best with respect to the indicators reliability of material, color and style information. The main reason why A4 is still perceived as the best with respect to these three sub-indicators stems from its ability to try-on apparel items on the virtual appearance of the individuals' own body image, which gives a better perception and feel of the apparel style and color according to the interviewed experts. Furthermore, the dynamic movement which can be created gives a better feel and perception of the material quality, which makes it a superior alternative, is very effective to evaluate the personal match of apparel items with online. Furthermore, A4 also scores the best with respect to the sub-indicator 'attractiveness'. A4 was perceived as the most attractive alternative, due to its ability to try-on apparel items on one's own mirrored image. In addition, the dynamic movement where apparel moves with the individuals' body movements, makes it more exiting, playful and visually appealing for customers to use.

Looking at the same table, A1 scores the best with respect to all two sub-indicators belonging to the main indicator 'information gathering and handling'. This implies that customers are perceived to have the least to no privacy and security concerns when using A1, since no data is collected in order to be able to use the instruments to evaluate apparel items with. Furthermore, based on all six experts, A1 scores the best with respect to responsiveness and availability since it is perceived as the least technically complex requiring the least amount of storage capacity as only pictures and a size table need to be uploaded along with the mix-and-match function. As such, the loading time will be the least negatively affected using A1. Since A1 requires the least amount of computational power, it can easily be made available on any device compared to the other alternatives.

A2 scores the best with respect to the sub-indicators search time and required preparatory work time. A2 is perceived as the best alternative with respect to search time, since the number of clicks/efforts needed to acquire the necessary apparel attribute information to evaluate apparel items with is considered to be the lowest for A2. Customers only have to fill in a form wherein they indicate their body-measurements information, which according to the experts is relatively easy and quick to do. In terms of 'required preparatory work time' A2 scores the best, since most of the experts perceived A2 to require the least amount of work customers need to do before they can start using a technology to evaluate the overall personal clothing match/fit.

4.4. Ranking the Technological Alternatives

Table 4 provides an overview of the obtained scores from expert interviews along with the weights obtained through the customer survey. In the first column, the sub-indicators (i.e., sub-criteria) are indicated. The subsequent four columns indicate the assigned scores of each alternative with respect to each criterion, obtained from six expert interviews. In the last column, the global weights are indicated. Using the additive value function (2), the final scores were obtained, and the alternatives were ranked based on preference.

Table 4. Ranking of the technological alternatives

Sub-criteria	**Scores of Technological Alternatives**				**Global weights**
	A1	**A2**	**A3**	**A4**	
Reliability of material information	0.140	0.204	0.299	0.358	0.107
Reliability of color information	0.217	0.217	0.228	0.338	0.110
Reliability of fit & size information	0.093	0.186	0.362	0.359	0.140
Reliability of style information	0.106	0.144	0.359	0.392	0.084
The way of data collection through technology	0.424	0.325	0.143	0.108	0.101
Data handling by online apparel retailer	0.503	0.278	0.123	0.096	0.133
Responsiveness	0.438	0.339	0.101	0.122	0.067
Search time	0.274	0.332	0.171	0.223	0.072
Availability	0.475	0.307	0.096	0.122	0.062
Attractiveness	0.104	0.144	0.316	0.436	0.058
Required preparatory work time	0.304	0.361	0.151	0.185	0.065
Total Score	0.2748	0.2517	0.2221	0.2516	
Ranking	1	2	4	3	

Based on the obtained criteria-weights through the survey and the scores from online apparel retail experts, it can be observed that A1 has the highest chance of reaching users' technology acceptance. A2 is perceived as the second best, closely followed by A4. A3 is perceived to have the lowest chance of reaching technology acceptance.

Managerial Implications

In order to establish what the perceived employment possibility of the technological alternatives is in companies, at the end of each expert interview, the same six online apparel experts were asked to score all four technological alternatives with respect to the criterion 'implementation possibility in company', again using BWM. Through this, insight was gained about factors which can encourage or inhibit the adoption of each technological alternative in online apparel retailing.

The results obtained from the expert interviews, as indicated in Table 5, show that when it comes to the practical implementation of the alternatives in companies, the same ranking is obtained as the ranking regarding the customers' acceptance of the alternatives (indicated in Table 4). Consequently, A1 is perceived to have the least number of managerial implications for online apparel retailers, since out of all four alternatives A1 is for the most part already employed, aside from the mix-and-match function to evaluate the entire outfit with. A2 is perceived as the second best, since it is perceived as the most technically and financially feasible for the companies, after A1. Looking at the state-of-the-art-technologies, A4 is perceived as the third best alternative, closely followed by A3 which is perceived to have the greatest number of managerial implications. A2 is much better than A3 and A4 when it comes to implementation in the company, as based on the experts results for both A3 and A4 experts have to be hired as the current developers do not have the knowledge to operationalize the technologies, which costs more money and time. When using A3 and A4, the whole chain needs to be aligned to the digital way of working which is required to operationalize A3 and A4, which according to the expert interviews does not seem feasible for mature multi-brand stores. In addition, testing the technologies and gathering customer opinions also takes much more time, effort and money compared to A1 and A2.

Table 5. Ranking of inherent managerial implications per alternative for online apparel retailers

Alternatives	Implementation possibility in company	Ranking
A1	0.456	1
A2	0.342	2
A3	0.099	4
A4	0.104	3

5. Discussion

The analysis shows that reliable fit & size information is the most important sub-indicator contributing to the customers' technology acceptance, which according to studies such as Hidellaarachchi et al. (2018), Shen et al. (2019), Misra et al. (2018), Saarijärvi et al. (2017), and Peng and Al-Sayegh (2014) and the approached online apparel retailers is indeed perceived as the apparel attribute with the most apparel returns. This proves that the Bayesian BWM is indeed a valid method to predict the importance of criteria. Furthermore, it seems that currently, A1 has the highest chance of reaching technology acceptance.

The results have indicated that the technology which has the highest probability of customers' acceptance is also the one which is currently the most employed by online apparel retailers in the Netherlands. This shows that the Bayesian BWM method is indeed an effective method to predict technology acceptance.

The reason to why A3 and A4 are the least preferred, might be due to the fact that they are relatively state-of-the-art. To find experts with sufficient expertise especially about A3 and A4 to participate in the BWM was rather difficult. Since only six experts stemming from four companies were approached, the individual influence of the assigned scores is higher, which also impacts the end results. However, when looking at the experts' data, most experts shared the same arguments and opinions implying that data saturation was reached.

Although all six interviewed experts shared the same opinions regarding the perceived usefulness, trust and ease of use of the technological alternatives, the nature of technology development of A3 and A4 is perceived as a reason for the discrepancy between the alternatives (A1 and A2 are inferior to A3 and A4 when it comes to technology superiority). This can mostly be seen by looking at the alternatives' weights with respect to the second and third most important sub-indicators affecting the users' technology preference which are 'the way of data collection through technology' and 'information handling by online clothing retailers' (see Table 4), implying that privacy and security concerns can occur when using A4 and A3. According to studies conducted by Hidellaarachchi et al. (2018) and Apeagyei (2010) customer might perceive discomfort or privacy and security concerns regarding sharing personal body-measurements data and the way body-measurements data can be obtained and used by online apparel retailers. This is also the case for A4 and A3 with nearly all sub-indicators related to the perceived user-friendliness which are responsiveness, search time, availability and required preparatory work time, implying that these two state-of-the-art alternatives are perceived as the least user-friendly.

Since the technological alternatives build upon each other in terms of functionality, the level of perceived technological complexity and data required increases. As a result of this, Table 4 shows that whilst the reliability of information provision regarding material, color, fit & size and style increases per subsequent alternative, the privacy and security concerns increase and the perceived user-friendliness (ease of use) of the technology decreases. Table 5 also indicates that per subsequent alternative, the managerial implications increase due to an increase in technical complexity and required data.

However, whilst it was true that online apparel retailers were asked to approach the scoring of the alternatives with respect to each criterion from the customers (users) perspective, it is still possible that the scores obtained from the expert interviews are (slightly) biased. Variables such as experience with functionalities of A1 and A2, low trust in new technologies, time of adoption, low level of innovativeness could all be underlying reasons explaining their assigned scores.

6. Conclusion and Recommendations

The purpose of this paper was to predict the customers' acceptance regarding various technological alternatives designed to increase customers online purchase successes and reduce unnecessary apparel returns. The Bayesian BWM was applied to operationalize TAM, which involves identifying various indicators (criteria), quantifying the importance of each indicator through the assigned preference and determining which indicator has the highest impact on technology acceptance through the assigned weight. Furthermore, by applying the Bayesian BWM, the decision-makers' preference of a criterion could explicitly be confirmed with a certain confidence level.

Within this research, 11 sub-indicators for the customers' technology acceptance and four technological alternatives have been analyzed. The analysis shows that reliable fit & size information is the most important sub-indicator contributing to the customers' technology acceptance. Furthermore, it seems that currently, A1 has the highest chance of reaching technology acceptance.

Based on the outcome of the research, more mature companies, especially multi-brand stores, are advised to focus on A2, since compared to A1, A2 requires the least amount of effort (time, money, expertise) to implement. Since (i) the survey results have shown that reliable fit & size information is perceived as the most important indicator for technology acceptance and (ii) based on the experts interviews the fit & size recommendation function of A2 can provide more reliable information compared to the static height/size chart of A1, there is suggested to gradually move from A1 to A2. Since out of the four apparel attributes, style information is perceived as the least important, there is suggested to first focus on the other apparel attributes especially fit & size information (the most important indicator). In order to prevent apparel returns, new companies entering the market are advised to focus on the sub-indicators with the highest weight and the alternatives which score the best with respect to the criteria with the highest weights, as this might help them to increase the number of successful sales and prevent unnecessary apparel returns.

Although all six interviewed experts mostly shared the same opinions regarding the perceived usefulness, trust and ease of use of the technological alternatives, there is still advised to continue this research by interviewing more experts who are more active in the field of product IT development, to explore the two newer technologies (A3 and A4) better, since the nature of technology development of A3 and A4 is perceived as the main reason for this discrepancy (A1 and A2 are inferior to A3 and A4 when it comes to technology superiority).

Although the results have indicated that A1 is the most preferred alternative, this cannot be guaranteed with full certainty. It could be that the four technologies could co-exist in practice, since in time the current technological superiority of A1 and A2 over A3 and A4 might change. Table 4 shows that A4 scores very similar to A2 (the second-best alternative). Furthermore, A3 scores the best with respect to the sub-indicator 'reliability of fit & size information' (the main reason for apparel returns). However, A4 is not far behind. Since the obtained expert results might be (somewhat) biased, the effectiveness of A3 over A4 with respect to fit & size cannot be fully guaranteed. Variables such as experience with functionalities of A3 or A4 could be underlying reasons explaining their assigned scores. For these reasons, further research is required regarding the technical applications of these technologies. Since A3 and A4 score the best with respect to providing reliable apparel attribute information (material, color, fit & size and style information), which is required to make better online purchase decisions and refrain customers' from returning items, future research could examine if they can co-exist.

In addition, future research could also examine what the customers' technology preference will be amongst different customer segments, by first identifying different clusters (groups) based on characteristics such as age, gender, and shopping experience. Through this, more in-depth insight might be gained regarding the estimated use of the different technological alternatives and the possibility of co-existence.

Future steps should also take into consideration that this research is limited to the apparel e-commerce sector in the Netherlands, implying that the obtained weights which have led to this ranking of the customers technology preference might be different based on other contextual variables and empirical setting. Furthermore, this research did not examine various measurement approaches which can be used to obtain customers body-measurements along with different applications of the included fit & size recommendation tools. As a result, further research can explore the various approaches and different applications in further detail along with their imposed benefits and costs. Given the uncertainty factor of online customer reviews, as the provided information is based upon customers opinion, online customer reviews along with customers hotline instruments were not included in this research. Therefore, the way in which these instruments can contribute to the reduction of online apparel returns can be further explored.

References

Ajzen, I. (1985) From intentions to actions: a theory of planned behavior. In *Action-Control: From Cognition to Behavior* (pp. 227–253), New York, US: Springer Verlag.

Algharabat, R. S., and Shatnawi, T. (2014) 'The effect of 3D product quality (3D-Q) on perceived risk and purchase intentions: the case of apparel online retailers', *International Journal of Electronic Business*, 11(3), 256-273.

Apeagyei, P. R. (2010) 'Application of 3D body scanning technology to human measurement for clothing Fit', *International Journal of Digital Content Technology and its Applications*, 4(7), 58-68.

Brooks, A. L., and Brooks, E. (2014) 'Towards an inclusive virtual dressing room for wheelchair-bound customers', *International Conference on Collaboration Technologies and Systems (CTS)*, Atlanta, US, 582-589.

Chuttur, M. Y. (2009) 'Overview of the technology acceptance model: Origins, developments and future directions', *Working Papers on Information Systems*, 9(37), 9-37.

Davis, F. D. (1985) '*A technology acceptance model for empirically testing new end-user information systems: Theory and results*', US: Massachusetts Institute of Technology.

de Leeuw, S., Minguela-Rata, B., Sabet, E., Boter, J., and Sigurðardóttir, R. (2016) 'Trade-offs in managing commercial consumer returns for online apparel retail', *International Journal of Operations & Production Management*, 36(6), 710-731.

de Prieëlle, F., de Reuver, M., and Rezaei, J. (2020) 'The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry', *IEEE Transactions on Engineering Management* (in press).

Difrancesco, R. M., Huchzermeier, A., and Schröder, D. (2018) 'Optimizing the return window for online fashion retailers with closed-loop refurbishment', *Omega*, 78, 205-221.

Edwards, J. B., McKinnon, A. C., and Cullinane, S. L. (2010) 'Comparative analysis of the carbon footprints of conventional and online retailing', *International Journal of Physical Distribution & Logistics Management*, 40(1), 103-123.

Feitelson, E., and Salomon, I. (2004) The political economy of transport innovations. In *Transport Developments and Innovations in an Evolving World* (pp. 11-26), Berlin, Germany: Springer.

Fishbein, M., and Ajzen, I. (1977) *Belief, attitude, intention, and behavior: An introduction to theory and research*, Boston, US: Addison-Wesley.

Gallino, S., and Moreno, A. (2018) 'The value of fit information in online retail: Evidence from a randomized field experiment', *Manufacturing & Service Operations Management*, 20(4), 767-787.

Geels, F. W. (2004) 'From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory', *Research Policy*, 33(6-7), 897-920.

Griffis, S. E., Rao, S., Goldsby, T. J., and Niranjan, T. T. (2012) 'The customer consequences of returns in online retailing: An empirical analysis', *Operations Management*, 30(4), 282-294.

Gu, Z. J., and Tayi, G. K. (2015) 'Consumer mending and online retailer fit-uncertainty mitigating strategies', *Quantitative Marketing and Economics*, 13(3), 251-282.

Gupta, H., Kusi-Sarpong, S., and Rezaei, J. (2020) 'Barriers and overcoming strategies to supply chain sustainability innovation', *Resources, Conservation and Recycling*, 161, 104819.

Hidellaarachchi, D., Gunatilake, H., Perera, S., Sandaruwan, D., and Weerasinghe, M. (2018) 'Towards a virtual garment fitting model for male upper body for online marketplace', *18th International Conference on Advances in ICT for Emerging Regions (ICTer)*, Colombo, Sri Lanka, 322-331.

Hox, J., and Bechger, T. (1999) 'An Introduction to Structural Equation Modeling', *Family Science Review*, 11, 354-373.

Kaushik, V., Kumar, A., Gupta, H., and Dixit, G. (2020) 'A hybrid decision model for supplier selection in Online Fashion Retail (OFR) ', *International Journal of Logistics Research and Applications*, In press.

Keeney, R. L. and H. Raiffa (1976) *Decisions with Multiple Objectives*, New York, US: Wiley.

Kristensen, K., Borum, N., Christensen, L. G., Jepsen, H. W., Lam, J., Brooks, A. L., and Brooks, E. P. (2013) 'Towards a next generation universally accessible 'online shopping-for-apparel' system', *International Conference on Human-Computer Interaction*, Las Vegas, US, 418-427.

Legris, P., Ingham, J., and Collerette, P. (2003) 'Review of the technology acceptance model', *Information and Management*, 40(3), 191-204.

Li, L., Wang, X., and Rezaei, J. (2020) 'A Bayesian best-worst method-based multicriteria competence analysis of crowdsourcing delivery personnel', *Complexity*, 4250417.

Liang, F., Brunelli, M., and Rezaei, J. (2020) 'Consistency issues in the best worst method: Measurements and thresholds', *Omega*, 96, 102175.

Liu, X., and Lei, M. (2008) 'Optimal uniform pricing and return policies for channel strategies in E-commerce age. *IEEE International Conference on Service Operations and Logistics, and Informatics*, Beijing, China, 1487-1492.

Marangunić, N., and Granić, A. (2015) 'Technology acceptance model: a literature review from 1986 to 2013', *Universal Access in the Information Society*, 14(1), 81-95.

Minnema, A., Bijmolt, T. H., Gensler, S., and Wiesel, T. (2016) 'To keep or not to keep: Effects of online customer reviews on product returns', *Journal of Retailing*, 92(3), 253-267.

Misra, R., Wan, M., and McAuley, J. (2018) 'Decomposing fit semantics for product size recommendation in metric spaces', *12th ACM Conference on Recommender Systems*, Vancouver, Canada, 422-426.

Mohammadi, M., and Rezaei, J. (2020) 'Bayesian best-worst method: A probabilistic group decision making model', *Omega*, 96, 102075.

Moktadir, M. A., Kumar, A., Ali, S. M., Paul, S. K., Sultana, R., and Rezaei, J. (2020) 'Critical success factors for a circular economy: Implications for business strategy and the environment', *Business Strategy and the Environment*, 29(8), 3611-3635.

Nachtigall, C., Kroehne, U., Funke, F., and Steyer, R. (2003) 'Pros and cons of structural equation modeling', *Methods Psychological Research Online*, 8(2), 1-22.

Nasibov, E., Vahaplar, A., Demir, M., and Okur, B. (2016) 'A fuzzy logic Approach to predict the best fitted apparel size in online marketing', *IEEE 10th International Conference on Application of Information and Communication Technologies*, Phoenix Park, Korea, 1-4.

Peng, F., and Al-Sayegh, M. (2014) 'Personalised virtual fitting for fashion', *International Journal of Industrial Engineering and Management*, 5(4), 233–240.

Rezaei, J. (2015) 'Best-worst multi-criteria decision-making method', *Omega*, 53, 49-57.

Rezaei, J. (2020) 'A Concentration Ratio for Nonlinear Best Worst Method', *International Journal of Information Technology & Decision Making*, 19(3), 891-907.

Saarijärvi, H., Sutinen, U. M., and Harris, L. C. (2017) 'Uncovering consumers' returning behaviour: a study of fashion e-commerce', *The International Review of Retail, Distribution and Consumer Research*, 27(3), 284-299.

Saaty, T. L. (2004) 'Fundamentals of the analytic network process-multiple networks with benefits, costs, opportunities and risks', *Journal of Systems Science and Systems Engineering*, 13(3), 348-379.

Seewald, A. K., Wernbacher, T., Pfeiffer, A., Denk, N., Platzer, M., Berger, M., and Winter, T. (2019) 'Towards Minimizing e-Commerce Returns for Clothing', *11th International Conference on Agents and Artificial Intelligence (ICAART)*, Prague, Czech Republic, 801-808.

Shen, J., Shang, X., and Dai, J. (2019) 'Garment E-commerce Return and Exchange 4PL Innovation', *International Symposium on Education and Humanities Sciences*, Wuhan, China, 309-317.

Vogelsang, K., Steinhüser, M., and Hoppe, U. (2013) 'A qualitative approach to examine technology acceptance', *The 34th International Conference on Information Systems (ICIS)*, Milano, Italy, 15-18.

Walsh, G., and Möhring, M. (2017) 'Effectiveness of product return-prevention instruments: Empirical evidence', *Electronic Markets*, 27(4), 341-350.

Walsh, G., Möhring, M., Koot, C., and Schaarschmidt, M. (2014) 'Preventive product returns management systems-A review and model', *22nd European Conference on Information Systems (ECIS)*, Tel Aviv, Israel, 9-11.

Wang, Y., Ramachandran, V., and Sheng, O. (2016) 'The causal impact of fit valence and fit reference on online product returns', *37th International Conference on Information Systems (ICIS)*, Dublin, Ireland, 11-14.

Zhang, Y. (2018) *Enhancing e-Commerce Performance: Product Return and Online Customer Review Perspectives*, East Lansing, US: Michigan State University.

Appendix A – Identified return reasons regarding apparel attributes

Description of category	Reason for returns (Return factors)	References
Disconfirmation driven: The quality of apparel differs or is not what was expected on the basis of the information provided on the website regarding apparel attributes such as material and color.	A different material quality than what was expected.	Saarijärvi et al. (2017), Algharabat and Shatnawi (2014), Gallino and Moreno (2018), Brooks and Brooks (2014)
	The color hue differs from what was expected.	Saarijärvi et al. (2017); Algharabat and Shatnawi (2014); Zhang (2018)
	Misleading information (apparel description, apparel images).	Saarijärvi et al. (2017)
	An unexpected negative aspect that was not visible in the apparel images (e.g., rips or tears).	Saarijärvi et al. (2017)
Size (chart) driven: The apparel size is not correct, although the customer exactly choses his or her size (e.g., small, medium, large) Size is your actual measurement (think waist, inseam, neck, etc.).	Size variations, inconsistencies or mismatches: The size of apparel items is too big or too small.	Saarijärvi et al. (2017), Hidellaarachchi et al. (2018), Seewald et al. (2019), Misra et al. (2018), Nasibov et al. (2017), Algharabat and Shatnawi (2014), Kristensen et al. (2013), Apeagyei (2010), de Leeuw et al. (2016), Wang et al. (2016), Peng, and Al-Sayegh (2014); Brooks and Brooks (2014); Shen et al. (2019)
Feeling driven: When actually trying on the apparel item, the customer does not feel 'right'.	The apparel item does not match the customers' style.	Saarijärvi et al. (2017), de Leeuw et al. (2016)
	The feeling of the apparel item is not right.	Saarijärvi et al. (2017), Brooks and Brooks (2014); Liu and Lei (2008)
	Customers' misjudgement/misperception of the right fit.	Saarijärvi et al. (2017), Seewald et al. (2019), Gu and Tayi (2015), Algharabat and Shatnawi (2014), Kristensen et al. (2013), Apeagyei (2010), Gallino and Moreno (2018), Wang et al. (2016), Peng and Al-Sayegh (2014); Zhang (2018)
Benefit Maximization driven: The customer orders multiple apparel items with the aim to keep only one or few of the items.	Ordering multiple sizes of the same apparel item, in order to keep only one.	Saarijärvi et al. (2017), Gallino and Moreno (2018); de Leeuw et al. (2016), Brooks and Brooks (2014)
	Ordering the same apparel item in multiple color, in order to keep only one.	Saarijärvi et al. (2017), de Leeuw et al. (2016), Brooks and Brooks (2014)
	Ordering alternative apparel items (e.g., different styles) for the same need, in order to keep only one.	Saarijärvi et al. (2017), de Leeuw et al. (2016)

Appendix B – Technological alternatives with their relevant components

Apparel attributes	A1: The bare minimum *	A2: The bare minimum with a fit & size recommendation instrument *	A3: Avatar (digital computer-based twin) *	A4: Virtual Dressing Room (VDR) *
Material quality information	Alternative product pictures, zoom function	Alternative product pictures, zoom function	Alternative product pictures displayed on avatar, zoom function, draping technology	Alternative product dynamic images displayed on individuals' real mirrored self (using e.g., ICT tools such as augmented reality), zoom function
Color information	Alternative product pictures, zoom function	Alternative product pictures, zoom function	Alternative product pictures displayed on avatar, zoom function	Alternative product dynamic images displayed on individuals' real mirrored self, zoom function
Fit & size information	Static height/size chart	Size recommendation application	Virtual try- on experience through avatar (personalized or retail-specified), mix and match function, size recommendation application	Virtual try-on experience on the individuals' real mirrored self, using camera sensors and contemporary ICT such as augmented reality, mix and match function, size recommendation application
Style information	Alternative product pictures, mix and match function, zoom function	Alternative product pictures, mix and match function, zoom function	Virtual try- on experience through avatar (personalized or retail-specified), mix and match function, zoom function	Virtual try-on experience on the individuals' real mirrored self, using camera-based sensors and ICT tools such as augmented reality, mix and match function, zoom function

* A photo of the apparel item, information about the apparel item in text form such as size, material and style are also indicated and the ability for color selection exists, as this information and functionality is already provided and employed by most online apparel retailers in the Netherlands. This is the case for all four alternatives.