

# All eyes on me: Predicting consumer intentions on social commerce platforms using eye-tracking data and ensemble learning

Patrick Mikalef<sup>a,b,\*</sup>, Kshitij Sharma<sup>a,c</sup>, Sheshadri Chatterjee<sup>d</sup>, Ranjan Chaudhuri<sup>e</sup>, Vinit Parida<sup>f,g,h</sup>, Shivam Gupta<sup>i</sup>

<sup>a</sup> Department of Computer Science, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, Norway

<sup>b</sup> Department of Technology Management, SINTEF Digital, Trondheim, Norway

<sup>c</sup> University of Science and Technology, Trondheim, Norway

<sup>d</sup> Department of Computer Science & Engineering, Indian Institute of Technology Kharagpur, India

<sup>e</sup> Indian Institute of Management Ranchi, Jharkhand, India

<sup>f</sup> Department of Social Sciences, Technology and Arts, Luleå University of Technology, Sweden

<sup>g</sup> Entrepreneurship and Innovation, Luleå University of Technology, Sweden

<sup>h</sup> School of Management, University of Vaasa, Vaasa, Finland

<sup>i</sup> Department of Information Systems, Supply Chain Management & Decision Support, NEOMA Business School, Reims, France

## ARTICLE INFO

### Keywords:

Eye-tracking  
Social commerce  
Ensemble learning  
Prediction  
Machine learning

## ABSTRACT

Understanding what information is important for consumers when making a purchase-related decision has been a key question for researchers and practitioners ever since the advent of empirical research in commerce. Nevertheless, our knowledge of what information is important has been formed primarily through post-purchase conscious capturing approaches, such as surveys and questionnaires. To overcome these limitations, we ground this research on an exploratory study that captures eye-tracking data during a decision-making task of product selection. Grounded on the dynamic attention theory, we utilize different information types and formats present on a popular social commerce platform, to identify elements which are important when deciding about online product purchase decision. Specifically, we employ a series of prediction algorithms and use an ensemble learning setup to predict the aspects that contribute to product selection by consumers. Our analysis highlights the most important informational cues to accurately predict product selection among alternatives. In addition, the results showcase how such elements shift in importance during the temporal sequence of comparing different product alternatives. Our results provide insight into how we can understand the journey of decision-making for social commerce customers when navigating through information to select a product. In addition, it opens the discussion about the shifts that eye-tracking in combination with machine learning can create for researchers and marketers.

## 1. Introduction

Ever since the emergence of online commerce, a primary question for research and practice has been to understand what informational cues trigger consumers to select certain products during their online purchases [32,74,75]. A key aspect in this quest has been to present consumers with sufficient information about a product to make ease the decision-making process [16,43]. Social commerce platforms which have become increasingly prevalent over the last decade have added to the complexity of this task, as they typically incorporate both marketer-

and consumer-generated information about a product [17,44,80]. In addition, such information is often presented in different formats, which facilitates consumers to select from a variety of informational cues that will ease the selection of products [6]. Yet, while there has been a significant advancement in the front of presenting information to consumers, we still have limited knowledge on how consumers utilize such information, and more importantly what type of information is critical for them during their decision-making process [59]. Furthermore, we have limited knowledge on how the importance of different informational cues evolves during the product selection process [49].

\* Corresponding author.

E-mail addresses: [patrick.mikalef@ntnu.no](mailto:patrick.mikalef@ntnu.no) (P. Mikalef), [kshitij.sharma@ntnu.no](mailto:kshitij.sharma@ntnu.no) (K. Sharma), [ranjan@nitie.ac.in](mailto:ranjan@nitie.ac.in) (R. Chaudhuri), [vinit.parida@ltu.se](mailto:vinit.parida@ltu.se) (V. Parida), [shivam.gupta@neoma-bs.fr](mailto:shivam.gupta@neoma-bs.fr) (S. Gupta).

<https://doi.org/10.1016/j.dss.2023.114039>

Received 2 December 2022; Received in revised form 30 May 2023; Accepted 14 June 2023

Available online 17 June 2023

0167-9236/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

A growing stream of research has sought to understand more about how users of social commerce platforms process different formats of information by capturing physiological data through eye-tracking approaches [7,10,23,71]. Eye movements and attributes of the pupil can help us understand the connection between eye movements and cognitive processes, and specifically how visual stimuli affect humans in decision making processes [4]. Specifically, the eye-mind hypothesis argues that the direction of our eye movements helps to assess what an individual is thinking and where the attention of the individual is focused [35]. The resulting data gathered from eye tracking methods can be analyzed statistically and graphically to provide evidence towards specific visual patterns and the sequence of gazing [11]. By examining the saccades along with other patterns of eye movements, researchers can ascertain the attractiveness of a given product, as well as the interaction of the user with different types and formats of information that are provided [45]. Such analyses provide valuable insight into what informational cues have the highest impact and which ones are ignored by users during their decision-making process [9]. Furthermore, such approaches can enable researchers and practitioners to understand how the attention of individuals dynamically changes over time during the decision-making process [58,76].

Prior studies have provided us with rich insight on the importance of customer satisfaction during the decision-making process of product selection [78] and its significance on prompting purchase intentions [33,36,37,40]. An important part in realizing customer satisfaction has to do with the quality information that is provided to users during their decision-making process [17,60,78,80]. Delving into this topic through eye-tracking methods, recent studies have found that users tend to complement different types and formats of information when deciding on which products to purchase [71]. Fei et al. [23] find that the type of information presented on online commerce platforms and the format or presentation have differentiating effects on consumer intentions. Building on an e-commerce platform using an eye-tracking approach, Brand and Reith [10] show that both the type and representation of information, have an impact on how users make decisions on credibility evaluations. Thus, there is a growing consensus that the type and representation format of information that users are provided with on online platforms, has an effect on their decision-making process [2,5,66,81].

While there is a growing understanding concerning the role of information cues on online platforms, we still have limited knowledge on how the utilization of the different types and formats of information influences a user's decision-making process during product selection. In addition, there is still a limited understanding concerning which informational cues are most important during the different phases of decision-making for users of social commerce platforms. From a methodological perspective, eye-tracking approaches provide several strengths in capturing detailed data on the gaze patterns of users, however, a challenge is in making inferences from such data. In this study we build on machine learning approaches to identify aspects of the product selection process that can accurately predict which product will be selected by consumers [25]. Doing so enables us to understand how users utilize different types of information present on social commerce platforms, and how the importance of such information changes during the decision-making process. We, therefore, seek to understand not only what type of insight eye-tracking can enable during the decision-making process of consumers, but also how such insight can be used to optimize the product selection experience. To address this research gap, the aim of this study is to address the following research questions (RQs).

**RQ1.** What informational cues are important for predicting consumer choices during the decision-making process on social commerce platforms?

**RQ2.** How does the importance of informational cues for consumers change over time during the decision-making process on social commerce platforms?

To answer these questions, we conducted two separate studies with participants in a lab setting. During these studies, participants were provided with several sets of products to select from on a popular social commerce platform. Using an eye-tracking device, we captured aspects of their visual processing patterns during product selection, as well as the types of information they used on the social commerce platform to reach a final decision. Using this data, we employed several machine learning and prediction algorithms to understand which features were the best predictors of product selection, as well as how these features shifted in importance over the duration of product selection. Our results highlight key elements that play an important part in decision-making for product selection. They also highlight the strengths of combining eye-tracking with advanced ways of analyzing data through machine learning. Based on these results, we draw several key research and practical implications.

## 2. Theoretical background

### 2.1. Dynamic attention theory

The Dynamic Human-Centered Communication System theory (referred to as Dynamic Attention theory) [41,73] has been used in this study to understand how the attention of a user in an online platform such as social commerce leads to certain outcomes, and how attention shifts over time. In this theory, attention is considered as a process involving interaction between humans as well as environment [41]. Dynamic attention theory helps to assess the intention of the consumers for the consumption of information and supports interpreting how the consumers make a decision of which product to select. Dynamic attention theory is founded on human-computer interaction (HCI) research which presumes that an active observer is supposed to use stable elements available in the environment which is also known as distributed cognition to facilitate decision-making [56].

In terms of dynamic attention theory, recording of eye tracking highlights that humans tend to acquire information from the available environment by using low-efforts gazing strategies compared to using high-efforts based on a strategy of memorization [42]. This human behavioral attribute has been the driver of design for many social commerce platforms and is the main driver by which user of such online platforms make a decision towards purchasing a product or a service. Understanding and predicting within the context of dynamic attention theory can be studied through the use of eye tracking approaches [70]. By tracking attention through an eye-tracking method, movements of eye occur in the direction of the object which receives more attention in terms of the dynamic attention theory, but "the resulting fixation and fixation durations may not be long enough for memory encoding" [73]. Hence, eye tracking methods can provide a more accurate and nuanced assessment of the implications of this theory.

Through the lens of dynamic attention theory, eye tracking is considered as a specific type of communication between the consumers and the product at hand, which is a complex dynamic system comprising of a human, a message, a congenial medium, and a location [41]. Thus, dynamic attention theory provides enables us to understand how the attention of a user in relation to a product evolves over time while a decision has been made. In the context of this study, dynamic attention theory is used to gain insight into the consumers' decision-making process and to try to understand the key elements that lead to the selection of a specific product. To gain such insight, the use of eye-tracking data is deemed as the most suitable as it is possible to record gaze patterns, saccades, and attention over a product decision-making activity.

### 2.2. Eye-tracking and decision-making

Applying eye tracking in studies of human behavior aids in capturing real-time information on individuals' fixations and visualization patterns as consumers; in turn granting researchers the ability to study the

behavioral-environmental processes more effectively behind transaction decisions [19,67]. Eye tracking is considered as a measurement of the movements of eye for determining the gaze location and to ascertain where the attraction is mainly focused [26]. Research using eye-trackers has built on a number of different measures, as for example fixation duration, an indicator of concentration and focus on a particular object or area [68]. There is not a particular study where the authors have provided a universally accepted definition of fixation in the context of eye tracking research because such definition differs between on-screen system of eye tracking and mobile eye tracking system [28]. Fixation may be interpreted as the duration during which the pupils of a person remained still on an object even the object is in movement [29]. Several studies have shown that there are mainly two types of systems that drive human visual attention which include stimuli-related attention known as bottom-up system, and observer related attention which is called top-down system [9].

In the context of eye tracking research, a seminal study by Ellis and Smith [22] highlights that the process of eye movement is either stratified random or completely random or may be statistically dependent stochastic. A study conducted in a physical store by Wästlund et al. [69] showed that when entering a shop, customers tend to typically start browsing the central areas of store shelves, and in sequence proceed to the peripheral areas on the horizontal axis. In this context, Chandon et al. [15] found that products which are placed on the horizontal central line as well as placed in the upper areas are found to have received better attention by the consumers in the context of on-screen self-test approach.

When considering online commerce, several studies have shown that the location of information that is depicted plays a critical role in purchase decision-making, where most tend to first observe the screen from the center and move towards the right side [64]. Researchers have demonstrated that since the central portion is considered as the optimal location containing maximum amount of preliminary information, people are used to have their attention centered to this part of the screen first [3]. Some research has also demonstrated that objects which are placed in the central area of the field of vision are attended by the people in a faster way for fostering initial saccades [13].

Nevertheless, the position of informational cues on the screen is not the only aspect that drives user attention and focus when making purchase decisions on online platforms. Several studies have shown that the type and format of information also play an important role in user attention [23,25,71]. Yang [77] found that positive and negative framing, or peripheral cues, can have an effect on purchase intentions by increasing attention on the cue message. Furthermore, on par with Yang's findings, Zhang and Benyoucef [80] found that there is a positive relationship between peripheral cues and purchase intention. During the decision phase, a consumer's mind is more heavily impacted by negative product reviews than positive ones [63]. As a result, product reviews may directly impact a consumer's decision on whether to purchase a product or service, depending on how they are framed and represented. The influence of information types is also noted in a study by Guerreiro et al. [27] who show that hedonic products present higher levels of fixation than utilitarian products.

While these studies provide valuable insight into the eye-tracking of individuals in relation to decision-making processes of consumers, they come under the limitation that they do not build on appropriate approaches of capturing the dynamic cognitive process that underpin such eye signals. In this study we leverage a series of prediction algorithmics and an ensemble learning setup to uncover how important eye-tracking information can provide us with real-time information about consumers preferences and decision-making processes. From a theoretical point of view, this approach facilitates a better understanding of how visual cues and information consumption influence decision making. From a methodological perspective, the combinations of methods and techniques allows to understand in real-time the types of informational sources that are important for consumers when making a purchase-

related decision and to predict with high accuracy the products that will be selected. These insights can help us understand how users of online social commerce platforms utilize different types and formats of information present on them, as well as uncover the temporal sequence of use of such information.

### 3. Research methodology

#### 3.1. Participants and procedure

We conducted two data collection studies to understand the relation between the customers' decision-making processes and their gaze-patterns. The choice of two studies was done on the basis of minimizing any potential bias due to the types and products that participants were offered to choose from as well as to differentiate the number of products that were needed to be compared to reach a decision. Thus, we followed relevant guidelines to mitigate the potential of bias in our tests by increasing the variance in a controlled manner [14,31]. The main idea for having two different studies (both in the terms of types of products and types of eye-trackers) was to have more variance in the model [8], which could lead to higher generalizability of the outcomes. Most of the measurements that we get from the two different eye-trackers is the same, for example, fixation duration, first fixations in the areas of interest, pupil diameter, saccade length and saccade velocity. Therefore, the two datasets were similar in nature. Moreover, having one head-mounted eye-tracker and one screen-based eye-tracker also allows us to study the consumer behavior in two different but ecologically valid settings. With a single study, the model would have been biased towards a specific setting for data collections and types of products. This would hinder the generalizability of our outcomes. On the other hand, we did not choose two different social commerce platforms to control the variability of the datasets.

In the first study, we recruited 30 participants (9 females, 22 males) with an average age of 24.84 years ( $SD = 6.58$  years). The participants were shown 12 products from different categories, and they chose six out of 12 products. These products appeared in pairs of two from which they needed to select one each time, and spanned different categories in order to ensure that there was no bias introduced due a specific type of product. For the second study we recruited 23 participants (10 females, 13 males) with an average age of 27.5 years ( $SD = 7.15$  years). The participants were shown three products and they chose one product out of three. We used a slight manipulation in the two studies on the number of products that the participants needed to compare among in order to reach a decision, the first being between two and the second being between three [31].

Participants for both studies were recruited through an open call with a brief description of the step and requirements of the study. We offered each participant a financial compensation equal to approximately \$20. The participants in both studies were familiar with Amazon as a company and had average experience with shopping on Amazon having purchased on the website more than once. During the studies, they participants signed an informed consent form that provided information concerning the purpose of the study and the way their data would be treated. They were notified that all data would remain anonymous, and the purpose of the study was solely for research purposes. Once they were briefed about the objective of the study, they were also informed that they needed to just select the product of their choice after assessing the relevant information, without the need to purchase it. We did however instruct them that they should treat the decision-making task as an actual product they would buy. Moreover, all the participants had  $20 \times 20$  vision (with or without correction). A threshold was defined and set for the calibration of eye-tracking, which is crucial to determine whether or not to re-calibrate and conduct the study. This threshold was set for the validation accuracy and was selected to be allowed a maximum of 2.50 degrees. The achieved validation accuracy was below 1.0 for all respondents, with a high of 0.98 and a low of 0.30.

Hence, no results were discarded due to an inability to produce a sufficient validation accuracy at this phase.

The eye-tracker's validation was conducted by repeating the same 5-point calibration scheme. Five points appear on screen, one after the other, and the participants are told to look at them for a few seconds. Then the in-built algorithm of the eye-tracker was used to validate the calibration of the eye-tracker. The validation threshold of 2.5 degrees corresponds to <10 pixels on the screen from a viewing distance of 50–70 cm. This is an acceptable error margin in eye-tracking studies [30].

### 3.2. Data collection

During the data collection participants were given as much time as needed to select each product among the pairs they were presented with. Throughout the entirety of the study, respondents were free to ask questions to the researcher regarding the study, however, they were encouraged to complete the study with minimal disruption and assistance. In order to prevent the end result of the research from being undermined by sub-optimal research scenarios, measures to improve the reliability of the eye tracking research were emphasized. Given that most eye tracking devices utilize infrared light reflecting from the pupil, alongside complex algorithms to track eye movements, the lighting in the study's environment must be stable. In other words, it is important to mitigate the amount of fluctuating infrared light and maintain consistent lighting levels to ensure high accuracy. Thus, we selected a controlled laboratory setting where incoming light and appropriate settings could be carefully calibrated. After participants signed a consent form, we collected non-sensitive demographic data, such as age and gender of the respondents through a quick demographic survey. In sequence, a brief introduction to the eye tracking system is given, in addition to a quick explanation of what will be tested during the ensuing study. Prior to the actual data collection, a calibration exercise was done in order to ensure the correct syncing of every individual participant's retinal movement to the recording equipment, which is an important part for the validity of measurements and results.

We utilized two different types of eye-trackers to control for possible movement of users during the studies and variations in sampling accuracy [20]. First with a stationary eye-tracker, we used Tobii Pro X3-120 and collected data at 120 HZ. The screen was a full HD 24-in. monitor with a display resolution of 1920 × 1080. Second, we used mobile eye-tracking glasses, where the participants were wearing the eye-tracking glasses and they were provided a laptop to watch the products' web-pages. We used SMI ETG (eye-tracking glasses) and Tobii Eye-tracking glasses to collected data at 60 Hz. The laptop screen was a full HD 17-in. with as display resolution of 1680 × 1050. For both the setups, we ensured that the laboratory was appropriately lit, mitigating effects from external elements of light and annoyances, and elements of obstruction and distraction were minimized before conducting the actual study. The use of two sampling methods was done to minimize the potential noise or error points in data collection, as well as to control for any differences between the two groups of users [14]. During the assignment of participants to tasks, there was a random allocation of the devices used. Our analysis showed that there were no significant differences between the participants that used the mobile and stable eye-trackers [31]. Moreover, the data exported from both types of eye-trackers has the same structure and we made sure the calibration and validation errors were similar in both cases.

### 3.3. Variables and measurements

#### 3.3.1. Dependent variable

In this contribution, we attempt to predict whether the given product was chosen to be bought by the participants or not, using the eye-tracking data. This is a binary variable with two non-ordinal categories: selected and rejected. This variable was measured through a

question which asked respondents to add their selected product in the purchase basket.

#### 3.3.2. Independent variables

Tobii's default algorithm (i.e., in-built function in the Tobii software for gaze data processing) was used to identify fixations and saccades (for details please see Olsen [47]). A filter (i.e., in-built function in the Tobii software) was used to remove the raw gaze points that were classified as blinks. Eye movement data provide the mean, variance, minimum, maximum and median of several parameters, such as pupil diameters, fixation details, saccade details, blink details, and event statistics. Table 1 provides an overview of the extracted features as well as the respective reference from the literature. All the Areas of interests (AOIs) were calculated as a proportion of the time students spend looking at the different areas of the screen.

Amazon has developed and implemented social commerce tools that have become familiar concepts to most of its users. Such tools include customer reviews and the product rating system, which function as our primary components comprising the *social popularity*, also known as peer influence. Customer reviews are peer reviews posted by other individuals who have purchased and used, or have experience with, the product or service in question. Amazon also possesses the ability to display time left of sale and the remaining quantity currently available for products when applicable. The latter comprises *scarcity*, present for the product if it is sufficiently low in stock, demonstrated by Amazon with red text at the right-hand side of the product display. The third category of AOIs is *product information*, which is readily available for all products on Amazon, based on what the vendor has provided as available description and information, e.g. product description, technical specifications, and product summary. Finally, the remainder of the available information on Amazon's web page for the product display is referred to as *distractions*. This comprises all visible elements that are not directly related to the product itself, but rather to Amazon's related product advertisements, based on their recommendation system. Examples of this include related products, recommended products, frequently bought together, and any other recommendation system appearances. The AOIs are selected to obtain substantial data on the relevant variables while attempting to balance sensitivity and selectivity for the targeted areas (Fig. 1). They are also chosen to represent and distinguish between the variables (product information, social popularity, scarcity, and distractions) as well as several additional unrelated informational sections presented by Amazon. Table 3 summarizes the AOIs for each product. It is important to point out that because the areas of the different AOI vary a lot in the webpage used, there was a normalization scheme used while computing the time spent on each individual AOI. The time on the different AOIs was normalized for the

**Table 1**  
Features extracted from the eye-tracking data.

Eye-tracking parameters	Features extracted
Diameter	Pupil [53] (mean, median, min, max, SD)
Fixation	Fixation duration [61] (mean, median, min, max, SD) Fixation dispersion [34] (mean, median, min, max, SD) Skewness of fixation duration histogram [55]
Saccade	Ratio of forward saccades to total saccades [39] (scanpath velocity) Ratio of global and local saccades [79] (threshold on sac. vel.) Skewness of saccade velocity histogram [52] Saccade velocity [57] (mean, median, min, max, SD) Saccade length [38] (mean, median, min, max, SD) Saccade amplitude [50] (mean, median, min, max, SD) Saccade duration [65] (mean, median, min, max, SD)
Events	Num. Fixations, Num Saccades, Fixation to saccade ratio
Others	Time spent on Areas Of Interest (AOIs) (see the Table 2 for the details about AOIs) Cognitive load (mean, SD, skewness) [21] Index of information Processing (mean, SD, skewness) [46]



**Table 2**

Description of AOI categories.

AOI category	AOI names
Product Information	Additional Details, Other Technical Details, Price, Product Description, Product Images Small, Product Main Image, Product Summary, Technical Details, Title, Zoomed Image
Social Popularity	4 Stars and Above, Customer Questions and Answers, Detailed 3 Star Reviews, Detailed All Reviews, Detailed Negative Reviews, Review Summary, Summary Reviews, Top Critical Review, Top Positive Review, Top Reviews
Scarcity	Quantity
Distractions	Compare Similar Products, Customers Also Viewed, Frequently Bought Together, Inspired By, Recommended Products, Related Products, Sponsored Products

**Table 3**

prediction results from using the full data length.

Metric	Training	Testing
Precision	88.12	83.33
Recall	87.25	86.96
F1-Score	87.68	85.11
Accuracy	87.56	85.42

area of the AOI by dividing the time spent on the AOI by the square root of the area of the given AOI.

### 3.4. Prediction algorithms

#### 3.4.1. Support vector machines

SVM maps an input  $x$  onto a multidimensional space using kernel functions (linear, radial or polynomial), and then any kind of regression can be used to model the input data in the new feature space. The quality of estimation is measured by the  $\epsilon$ -intensive loss function given by Cortes and Vapnik [18].

#### 3.4.2. Gaussian process models

GPM is similar to SVM; the only difference is that the mapping from the original space to a multidimensional space is governed by Gaussian latent variables that are parametrized using different kernel functions [12]. In this study, we set the kernel functions to take linear, radial, and polynomial forms.

#### 3.4.3. Random forest

Random Forests are ensembles of decision trees mostly used for classification and/or regression purposes. The training algorithm for RF applies the general technique of bagging repeatedly selects a random sample with replacement of the training set, fits trees to these samples, and uses these replicates as new testing sets. The random forest is able to permute the given feature set and compute the feature importance for each feature in a given dataset, by optimizing one of the modelling parameters, e.g., root mean squared error, proportion of variance explained; or in the case of classifications, precision and/or recall. Using the individual feature importance from RFs, one can put a threshold either on the number of features (10 in our case) or on the importance values of the features to select the required number of features.

### 3.5. Ensemble learning setup

One way of using the results from multiple models is to use a weighted average from all the prediction algorithms. The weights for individual prediction are considered based on their accuracy during the validation phase. There are 3 major advantages of these methods [24,48,54]: 1) We can compare the performance of the ensemble methods to the diversification of our models predicting cognitive performance. It is advised to keep a diverse set of models to reduce the

variability in the prediction and hence, to minimize the error rate. Similarly, the ensemble of models will yield better performance on the test case scenarios (unseen data), as compared to the individual models in most of the cases. 2) The aggregate result of multiple models always involves less noise than the individual models. This leads to model stability and robustness. 3) Ensemble models can be used to capture the linear, as well as the non-linear relationships in the data. This can be accomplished by using two different models and forming an ensemble of both.

The main reason for selecting the models was to introduce the three categories of prediction models was to make sure that we are using the benefits of the models and maximize the prediction performance. For example, SVM is a preferred model when the dataset is relatively small and have high dimensional feature set. Moreover, SVM is known to reduce the risk of over fitting [51]. One drawback of SVM is that it might be difficult to choose the kernel function (Amari & [1]) and therefore, we chose to use all three of them in this contribution. Considering the Gaussian process models, they also have similar advantages as SVMs with one difference. The major difference in GPM and SVM is the use of gaussian latent variables to map the original feature space onto the higher dimensional space. Once again, there is no clear indication, from the data, about which method of creating the higher dimensional space from the original space. Therefore, we decided to use both methods to optimize performance. Finally, we included Random Forest in our ensemble because it is known to be efficient while handling multidimensional data when the different dimensions have varied distributions and ranges. This is the case in our dataset as well, for example, the underlying distribution of the five major categories of features (diameter, fixation, saccades, events, and other) are different from each other.

### 3.6. Training validation and testing

We perform out-of-sampling testing (i.e., leave-one-participant-out), dividing all 3 first datasets into 3 subsets: (1) training, (2) validation, and (3) testing. We keep the testing set aside (15%). The datasets are split based on participant identifiers. All the models are trained and validated using the training and validation sets with a cross validation. The cross-validation is performed using leave-one-participant-out. We used the following metrics to evaluate the performance of the ensemble classifier:

1. Precision =  $TP / (TP + FP)$ ;
2. Recall =  $TP / (TP + FN)$ ;
3. Accuracy =  $(TP + TN) / (TP + TN + FP + FN)$ ;
4. F1 score =  $2TP / (2TP + FP + FN)$ .

Where,

TP = true positive;  
FP = false positive;  
TN = true negative;  
FN = false negative.

For evaluating the prediction quality the “chosen” class is the “positive” class. For the baseline prediction, we selected the “random prediction baseline”, due to the balanced nature of our dataset.

The validation in the prediction pipeline was done using leave-one-participant out scheme, this is the most suited method of validation for datasets with smaller sample size [72] as opposed to K-fold validation scheme that is suited for larger datasets, for example, as it is the case with Aribag & Schwartz [2]. In this process the training phase is repeated while leaving one participant out for validation every time. This process is repeated until all the participants in the training data have been left out once. Once the process is complete the validation accuracy is calculated as the mean of all the leave-one-participant-out iterations.

[illegible]





### About Amazon Renewed

Amazon Renewed is your trusted destination for replacement and refurbished products that are inspected and tested to work and look like new. A Renewed supplier uses Amazon-qualified performance managers, performs a full diagnostic test, replaces any defective parts, and thoroughly cleans the product. The products will have minimal to no signs of wear, no visible cosmetic imperfections when held 12 inches away, and may arrive in a box or be a better or like new with relevant accessories that may be present.

**This product is backed by the Amazon Renewed Guarantee**  
Your product is eligible for a replacement or refund within 90 days of receipt if it does not work as expected. Get quick support for claims and free troubleshooting via a single point of contact at Amazon. The guarantee is in conjunction with Amazon's standard return policy. [Learn more](#)

Shop for smartphones, computers, laptops, tablets, home and kitchen appliances, game consoles, office products, and more on [Amazon Renewed](#).

Product information		Technical Details	
Technical Details		Additional Information	
Summary		Additional Details	
Processor	3.2 GHz Intel Core i7	ASIN	B0F4XW6W6C
RAM	8 GB ddr4_sdrac	Customer Reviews	★★★★★ 1 rating
Hard Drive	1 TB 8mm_parallel_ssd_sdrac	Best Sellers Rank	#14,908 in Computers & Accessories (See Top 100 in Computers & Accessories)
Graphics Coprocessor	NVIDIA GeForce GTX 1070	ML478 in Power Computers	
Card Description	dedicated	Shipping Weight	36.9 pounds (See shipping rules and policies)
Graphics Card Ram Size	8 GB	Date First Available	October 22, 2018
Wireless Type	802.11ac		

Compare with similar items				
   				
<p>This item <b>Intel Core i7-8700 Gaming Desktop Intel Core i7-8700 16GB DDR4 Memory, 256GB SSD + 1TB SATA Hard Drive, 600 Watts GeForce GTX 1070, Windows 10 Home, Raven Blue (Renewed)</b></p> <p><b>Add to Cart</b></p>				
<p><b>Intel Gaming 5089 Gaming Desktop Intel 8th Gen. Core i7-8700 8-Core up to 4.60 GHz, 16GB DDR4 Memory, 256GB SSD + 1TB SATA Hard Drive, 600 Watts GeForce GTX 1070, Windows 10 Home, Raven Blue</b></p> <p><b>Add to Cart</b></p>				
<p><b>Del G1 Gaming Desktop, Intel Core i7-9700, NVIDIA GeForce RTX 2060 6GB, 600W, 512GB SSD Storage, 16GB RAM, (5050-71850G-PUS)</b></p> <p><b>Add to Cart</b></p>				
<p><b>Del i5680-7818BLU-PUS Intel Core i7-9700, NVIDIA GeForce RTX 2060, Intel Core i7-9700, 16GB DDR4 Memory, 128GB SSD+1TB SATA HDD, NVIDIA GeForce GTX 1060, Raven Blue, Windows 10 64-bit</b></p> <p><b>Add to Cart</b></p>				
Customer Rating	★★★★☆ (1)	★★★★☆ (1)	★★★★☆ (28)	★★★★☆ (118)
Price	\$1,299 <sup>90</sup>	\$1,529 <sup>90</sup>	\$1,455 <sup>90</sup>	\$1,419 <sup>90</sup>
Shipping	\$4.95	FREE Shipping on orders over \$25	FREE Shipping on orders over \$25	FREE Shipping on orders over \$25
Sold by	IT Surplus Liquidators	Techno Intelligence	Swan Bell	Rock eDigital
computer memory size	8 GB	8 GB	16 GB	16 GB
Processor (CPU) Model	Intel Core i7	Intel Core i7	Intel Core i7	Intel Core i7
Processor (CPU) Manufacturer	Intel	Intel	Intel	Intel
Processor (CPU) Speed	5.2 GHz	5.2 GHz	4.7 GHz	5.2 GHz
Graphics Coprocessor	NVIDIA GeForce GTX 1070	NVIDIA GeForce GTX 1070	NVIDIA GeForce RTX 2060	NVIDIA GeForce GTX 1060
Graphics Description	dedicated	dedicated	Dedicated	Integrated
Hard Disk Size	1 TB	1 TB	512 GB	120 GB
Hardware Platform	PC	PC	Windows	PC
Item Dimensions	17.28 x 8.5 x 18.1 in	17.28 x 8.5 x 18.1 in	—	17.23 x 8.5 x 18.07 in
Item Weight	—	—	—	30.27 lbs
Operating System	Windows 10	Windows 10 Home	Windows 10 Home	Windows 10 Home
Processor Count	6	6	8	6

## Compare Similar Products

### Product description

- High-powered processing:** 10th Intel Core i7 (8-Core) processors deliver exceptional performance and headroom for light-second responsiveness and powerful multitasking capabilities.
- Vibrantly-gaming:** Designed for performance and expandability, the chassis supports NVIDIA GeForce GTX 1070 graphics configurations that meet the stringent Ready for VR demands.
- Purposely designed:** The Intelligent Gaming Desktop is designed for optimal airflow and temperatures with a solid side panel with Polar Blue LED lighting.
- Dynamic audio:** Waves MaxxAudio Pro provides audio controls and profiles with a smart stage that adds immersion to games, music and videos. Equipped with Performance 7.1 Channel HD audio sound card.
- Quick operation:** The Intelligent thermal design, meticulous component placement and optimized graphics solutions all contribute to efficient heat control.
- Next-Gen:** Microsofts gaming machine comes with 600 Watt Power Supply Unit (PSU), one PCIe x16 expansion slot and up to 4 bays for future storage expansion (3x HDD and 2 5.25 bays).
- Dominate both screens:** An HDMI port allows you to connect to an external monitor to view content in up to 4K UHD/2160p resolution for stunning detail.

### Product Description

Videos

Page 1 of 3

Videos for related products

The World's Best  
**ACER DESKTOP COMPUTERS**  
4:48  
The 5 Best Acer Desktop Computers  
BestReviews

Gaming  
**PERSONAL COMPUTERS**  
4:51  
The 10 Best Gaming Personal Computers  
Ezold Wiki

The Best  
**DESKTOP COMPUTERS**  
4:12  
The 5 Best Desktop Computers  
BestReviews

Desktop, Hiss, and Alike-Gone  
**Computer Companion**  
1:58  
Amazon PC

[Upload your video](#)

Customer questions & answers

Customer Questions and Answers

Q How a question? Search for answers

Question:

Does this item include keyboard and mouse?

Answer:

Yes I received it with keyboard and mouse  
By Sammie - on November 28, 2018

Customer reviews Summary

Top Reviews

5 star rating

5 star

100%

4 star

(%)

3 star

(%)

2 star

(%)

1 star

(%)

How does Amazon calculate star ratings?

See all reviews from the United States +

Review this product

Share your thoughts with other customers

Write a customer review

Your Browsing History

View or edit your browsing history +

See personalized recommendations

Sign in

New customer? Start here.

**Fig. 1.** AOI example from one product.

### 3.7. Prediction with partial temporal data

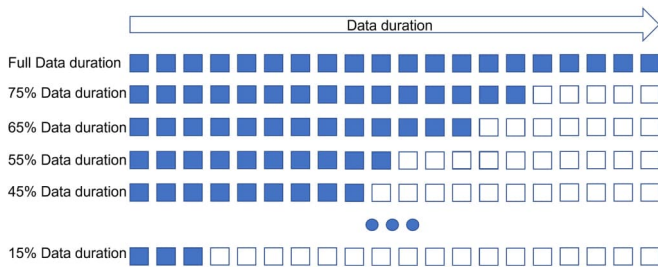
We also predicted the dependent variable using partial data, to test “*how quickly can we predict whether the customer will buy the product?*”. For this, we took 75% of the data (approx. 6–7 min of eye-tracking data) from each participant and used the methods described above to predict the dependent variable. Then, we keep removing 10% data based on the time up to 15% of the data (approx. 1–2 min of eye-tracking data) (Fig. 2). For each partial dataset we evaluate the prediction performance and the set of most important features. This set is chosen based on the features’ importance computed from the random forest classifier and has a value of  $>75$  (out of 100).

## 4. Results

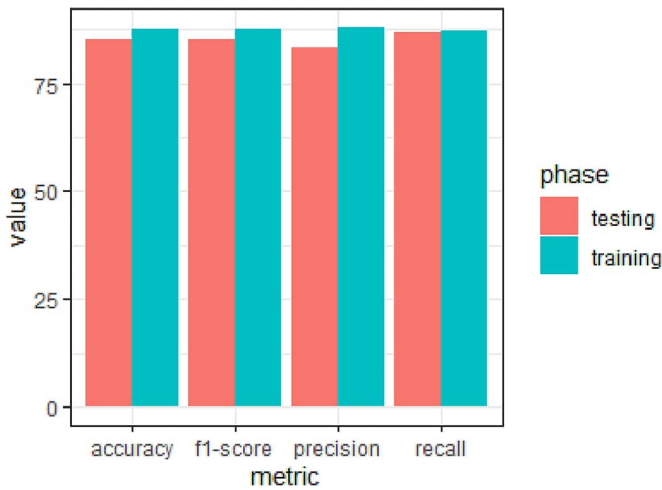
In this section, we will present the results from the two prediction setups. The first set of results is about the using the whole data duration for training, validation and testing along with the variable importance. Second, we will present the prediction results from using the partial data duration along with the changing variable importance for each of the most important predictors.

#### 4.1. Basic prediction results

First, we predicted whether the product was “chosen” based on the complete data and we obtain a good prediction quality as it can be seen in the Fig. 3 and in the Table 3. The most important variables are shown



**Fig. 2.** Schematic representation of the different data duration for the prediction of whether the customer selected or rejected a given product. Each block represents 5% of the data. Blue blocks show the data used for prediction and the white boxes show the remaining unused data in each iteration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

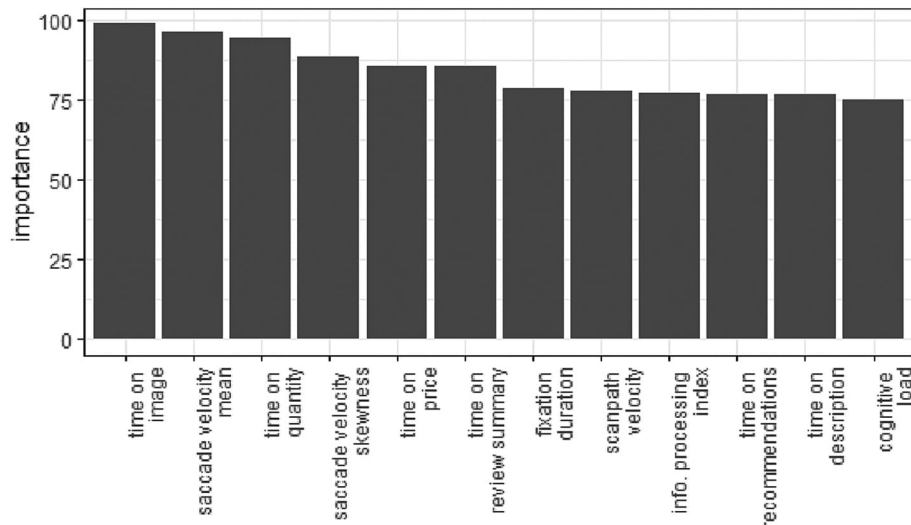


**Fig. 3.** Visual comparison of the prediction quality metrics from using the full data duration.

in the Fig. 4. We can observe that the precision (88.12) and recall (87.25) from the training phase are similar to those in the testing phase (precision = 83.33; recall = 86.96)., this shows that we successfully avoided over-fitting of the data, because the testing was done using an out of sample setup. In both these phases, we clearly improve the

random assignment baseline (precision = 0.50, recall = 0.50, f1-score = 0.50, accuracy = 0.50). Moreover, the most important features for predicting the dependent variables are cognitive load, information processing index, time on AOIs (price, image, recommendations, quantity, review summary, description), fixation duration, scanpath velocity, saccade velocity mean, saccade velocity skewness. The list of most important variables in the predict model show that there are clear patterns from the gaze data that can help us in classifying whether the customer would have selected the product or not. For example, we observe that apart from the AOI-based features, many other variables appear as the most important variable. This indicates that the importance of the visual information processing behavior is as important as the behavior induced by the information provided by the e-commerce platform. We performed single feature predictions to show how the consumer choice would be predicted with the features concerning the time on the different AOIs. We performed Single feature predictions (Appendix A, Table A3) to compare the most important feature with other AOI-based features that are in the top 12 most important features.

Moreover, we also showcase the effectiveness of the ensemble methods used in this contribution by two different methods. First, we compare the proposed method by replacing one of the methods (at a time) with logistic regression and artificial neural networks. The comparison results are shown in Appendix A (Tables A4 and A5). None of the two methods provide better prediction performance than the one used in the paper. Second, we also compare the ensembling of the three models using feature fusion with the ensembling using the weighted average. In the Appendix A (Table A6), we show the precision and recall when one of the five major categories of the features (Diameter, fixation, saccade, Events, others) are missing from one of the three major categories of the prediction algorithms groups (SVM, GPM, RF). We observe that none of the ensembles are better than the weighted average method of ensembling the models. One clear explanation for such results is the lack of information in each individual model which leads to the poor performance. Finally, we compare the ensembling of the three models using another scheme feature fusion with the ensembling using the weighted average. In the Appendix A (Table A7), we show the precision and recall when two of the five major categories of the features (Diameter, fixation, saccade, Events, others) are missing from one of the three major categories of the prediction algorithms groups (SVM, GPM, RF). We observe that none of the ensembles are better than the weighted average method of ensembling the models. One clear explanation for such results, is the lack of information in each model resulting in poor performance.



**Fig. 4.** The most important variables from the ensemble learning pipelines.

**Table 4**  
prediction quality from the partial data prediction.

Data	Training phase				Testing phase			
	Recall	Precision	F1-score	Accuracy	Recall	Precision	F1-score	Accuracy
75%	84.80	82.78	84.60	83.92	80.62	80.47	80.82	80.91
65%	83.34	82.44	83.25	82.30	80.01	80.06	80.60	78.72
55%	83.22	81.95	82.85	81.13	77.92	79.83	78.82	76.58
45%	82.32	79.61	80.13	79.20	77.79	78.18	78.18	76.31
35%	82.09	79.53	79.90	78.63	76.45	77.83	77.73	76.26
25%	78.56	79.45	79.71	78.58	76.41	76.33	76.16	76.01
15%	< 25	< 25	< 25	< 25	< 25	< 25	< 25	< 25

**Table 5**  
most important features from the partial data prediction.

Feature	Data partitions					
	75%	65%	55%	45%	35%	25%
Cognitive load	80.46	92.84	79.65	93.85	94.27	91.48
Info. processing index	75.76	81.83	87.78	81.24	89.04	83.99
Time on price	90.47	95.88	80.79	75.84	79.70	90.08
Time on image	91.96	86.74	93.44	88.17	91.77	82.94
Time on recommendations	98.29	92.79	95.89	90.88	75.17	97.18
Time on quantity	97.18	90.77	80.62	93.99	79.48	84.26
Time on review summary	80.16	76.11	89.20	95.85	77.33	82.43
Time on description	75.16	82.03	88.19	80.75	92.24	93.71
Fixation duration	97.43	92.89	92.20	95.97	94.65	84.31
Scanpath velocity	96.68	76.03	97.42	82.50	84.83	78.94
Saccade velocity mean	80.36	93.10	98.06	78.46	75.53	81.14
Saccade velocity Skewness	81.08	81.48	91.87	79.84	75.97	95.20

#### 4.2. Partial prediction results

Next, we present the results from predicting the dependent variable using shorter periods of the data. [Tables 4 and 5](#) present the prediction quality and the variable importance for the different data lengths, respectively. We observe from the [Table 5](#) that the set of the most important features remains the same as it was for the full data prediction, with changes in the importance order. We can also observe that the prediction quality does not deteriorate until the 35% data partition and at the 25% data partition the recall is the only metric that changes by a considerable difference. This shows that by using a little over two minutes of eye-tracking data, we can predict whether the customer is going to buy a given product. From [Table 5](#), we also observe that the most important feature set remains consistent across the different data slices (how much data we are using in the terms of duration). However, there are slight differences in the ranking of these most important variables but there is nothing that bursts out of order in the terms of feature importance. This is indicative of the fact that there is a considerable amount of information in the gaze data from the participants while they process e-commerce websites, and this information is also consistent over time.

We compare the proposed method with two other methods, which have similar data collection settings as in our studies. The first (Appendix A, Table A1) utilizes Hidden Markov Models (HMM, [76]) and the second (Appendix A, Table A2) utilizes ANOVA-based prediction [25].

#### 5. Discussion and conclusions

In this research we have sought to understand how consumers utilize different types of informational cues when purchasing online, and specifically, when attempting to make a choice between alternative products. Online commerce platforms nowadays incorporate different types and formats for presenting information to consumers, both marketer- and consumer-generated. Nevertheless, we still know very little about how consumers interact with such content and which types of

information they utilize during the process of making a purchase-related decision. Furthermore, there is a lack of understanding concerning how users during their decision-making process utilize different informational cues to reach a decision. Using an eye-tracking approach and building on ensemble learning methods of prediction, we uncover what aspects of information on social commerce sites enhance decision-making.

In relation to our first research question (RQ1), our analysis revealed that there are certain informational cues that decision-makers placed more focus on and where more important in explaining a purchase decision. Specifically, the most important cues included the time a user spent on the image, the saccade velocity, the time on quantity, saccade velocity skewness, time on price, and time on review summary. These outcomes demonstrate that certain informational cues are of high significance for decision-makers in the context of social commerce, since the process of deciding what product to purchase is a result of carefully examining the visual characteristics of the product at hand, as well as the expected cost, scarcity, and experiences of others. The saccades, which correspond to the eye movement from one point of fixation to another, also indicate that when there is increased speed and dynamics of fixation that is a strong indicator of consumer choices during the decision-making process. Taken together, these findings indicate that both different informational cues presented on screen as well as patterns of gazing behavior can capture consumer decision-making. Specifically, in regard to the informational cues, the results show that consumers are particularly drawn to the visual stimuli from product images, as well as the trade-off between price and scarcity, in relation to other consumer experiences, as presented by the relevant information.

When exploring the second research question (RQ2), our analysis reveals that there are indeed fluctuations in the importance of different aspects in relation to decision-making on social commerce platforms. More precisely, we find that by using different segments of activity the informational cues play an important role on decision-making shift. In fact, time on recommendations appears to be a very important determinant during the first quarter of the allocated time, while it later shifts to the price information, and in sequence to the quantity of the product. This temporal sequence of importance for the different information cues allows us to gain a better understanding of the decision-making process that users of social commerce platforms go through when deciding about what product to purchase. Within the context of social commerce product selection, it highlights that consumers initially rely on experiences provided by others, which may indicate that they perceive this type of information as more credible than that of marketers. In sequence, once trust has been established towards a product, informational cues that revolve around the assessments of the trade-off between price and its scarcity receive more importance. Finally, details about the product such as the visual appearance and a re-confirmation of opinions from other users conclude decision-making. Such insight provides us with a more nuanced understanding of decision-making as a process that is dynamic and changing, and where the importance of informational cues shifts over time.



### 5.1. Research implications

This study has built on a combination of data collection approaches and methods to analyze eye-tracking data which is gaining momentum in the domain of consumer decision-making. Our work contributes by extending the knowledge in this domain concerning how different informational cues influence consumer decision-making during product selection of social commerce websites, where there is a plethora of different information types and formats. In addition, we explore the dynamic nature of decision-making identifying the temporal sequence of user preferences when interacting with such information. In addition, the approaches used to analyze such types of data can be transferred to other application areas where there is a presence of such complex phenomena. Specifically, this study contributes to ongoing research in the following ways.

When it comes to understanding how consumers interact with information presented on online commerce platforms, research to date has identified several important aspects and information types that trigger consumer intentions [23,69,71]. To date, most studies have attempted to examine the value or importance of certain types of information (user- vs marketer-generated) when selecting product online, or on identifying for users perceive certain types of prompts when those are presented to them. Nevertheless, contemporary social commerce platforms manage to contain diverse types of information generated by marketers and consumers, as well as different formats of presenting such information. In our study we have defined several different areas of interest (AOIs) which correspond to some of the information users are exposed to during a purchase decision. Our findings indicate that users make use of certain types of information over others, and that the value of information presented on such websites does not have equal weight when it comes to ensuring that consumers make an informed decision. In our analysis we have included predictors of the layout, as well as on the gaze patterns of consumers. Specifically, the analysis pinpoints to the fact that certain types of information such as the time spent on the image, price, and on reviews have a strong effect on decision-making prediction. Thus, this finding indicates that we can infer more detailed information when can capture, or control for both.

Nevertheless, another key finding is that the importance of informational sources and gaze patterns dynamically changes during the decision-making process. As users interact with online interfaces the significance of certain types of information either increases or decreases in importance. This finding shows that there is a temporal significance of key aspects for users as they are making decisions. Several of the identified informational cues shift in importance for users of social commerce platforms as they decide which product to select. This finding provides some context to our first point, that the importance of certain types of information and how they are represented should be considered in the context of when exactly during the decision-making process they are utilized. This finding indicates that informational sources may have an ephemeral value in satisfying consumer requirements. Extending on such a reasoning entail that we need to develop a more nuanced understanding of how consumers of information utilize such sources over time, rather than in a snapshot in time. Thus, a fruitful domain for future studies is to adopt a more dynamic perspective to understanding how users interact with online commerce platforms, and conceptualize decision-making as an active process of interaction, between the user and the available information.

Adding to the above, our approach opens future research avenues for understanding designing dynamic interfaces that can satisfy user requirements. With the prevalence of smart glasses and more advanced eye-trackers in devices that are not intrusive, there is a renewed interest in such dynamic interfaces, and presenting users with the right type and format of information that is needed at the different stages of the decision-making process. Such interfaces that can capture data in real-time from eye-tracking inferences and dynamically alter the information that is presented to users or consumers is likely to be an important

area of research in the years to come. To date, we have been accustomed to static interfaces that only change after they have been prompted by user action. The same applies also in the case of social commerce websites where consumers need to initiate an action to receive information that may interest them, such as clicking on an option to get additional information. By capturing real-time data from eye-movements, there is a renewed interest in developing interfaces that can present important information automatically.

A key aspect in being able to analyze data such as eye movements and make accurate assessments of user intentions is utilizing advanced methods of prediction. In this study, we showcase how different prediction algorithms can be used on eye-tracking data in order to provide a more detailed understanding of aspects that contribute to product selection. While this study is not the first to combine such data and analysis approaches [4,23,62], it does showcase how it can yield interesting findings within the area of social commerce by focusing on key types of information presented on such online platforms. The combination of rich data from physiological data with advanced methods of analyzing such data can enable the examination of complex phenomena in emerging digital technologies (e.g., smart glasses, smart windshields etc.). We therefore argue that such approaches will likely be very useful in future research as a more in-depth lens of understanding user behavior when interacting with digital technologies.

### 5.2. Practical implications

Apart from several important research implications, the study also provides some interesting insights for practitioners, both on how to develop interfaces for social commerce platforms as well as for future advancements in the domain. In terms of designing and developing interfaces on social commerce platforms, our results provide some useful information to UX designers on aspects that are of increased importance to consumers when making a product selection. They also highlight the need to design based on principles of efficiency and use from the end-user's perspective. As marketers want to improve the efficiency of product selection and browsing by providing the right type of information to consumers, it is important that such input is utilized in the design process. Prior studies have shown that information overload can negatively affect product selection on such platforms, so a key take for designers is not to include as much information as possible on social commerce websites but rather to focus on the ones that are of importance to consumers.

In addition, our results highlight that there are certain aspects from the analysis that users find particularly important in their decision-making process. These can be leveraged from practitioners in the design of online platforms for optimizing the decision-making process of users. Furthermore, the temporal order of importance can provide insight into the structure of design and how to present important information in a way that eases the user's decision-making. For instance, time on recommendations was found to be an important aspect of information for the users of social commerce platforms during product selection which highlights the focus that should be placed on presenting such information in a clear and distinct way, without proximity to other aspects that might distract the user's attention and using appropriate font size and style to facilitate easy readability. Similarly, the time on image features as one of the most important aspects of prediction for purchase intention, which indicates that is a need for practitioners to design interfaces that can capture the intricacies of the product that is visualized. In other words, images of products on social commerce platforms must be of high quality and featured in a way that can make them easily accessible and interactive to users.

Furthermore, an interesting practical implication has to do with emergence of novel wearable devices and how they can be leveraged to create dynamic interfaces for information presentation. A growing number of users are now using smart glasses, or even laptops that have enhanced abilities of understanding face gestures and track eye

movements. This trend denotes a progression in terms of information that can be utilized by marketers to provide more personalized information to consumers and enhance their product selection when purchasing online. For instance, it has become common not to utilize location data to provide accurate advertisements, or third-party browsing history for product suggestions. With novel devices that can track eye-movement, we are likely to see new design paradigms where information that is presented to consumers is adjusted along the product-browsing journey. In addition, much of what is presented to users may be dynamically adjusted based on several factors such as individual preferences, type of product, or even based on real-time data like where the consumer is focusing their attention or if there is a detection of drowsiness, lack of attention, or cognitive overload. Such new waves of designing interfaces with the integration of real-time physiological data are likely to herald a new era for online purchasing.

### 5.3. Limitations and future research

Although in this study we have attempted to minimize the presence of any bias and to provide results that are generalizable, the outcomes do not come without limitations. First, we have based our analysis on a product selection study which was conducted in a lab and with a pre-selection of products for participants. While such a set-up approximates reality, it does not accurately capture the process of product selection of individuals. A more realistic approach would have been to allow individuals to search and find different products on their own, and possibly select between multiple different options. Nevertheless, we did not use this approach as we wanted to be able to compare similar conditions for all participants. A second limitation is that during the participant selection and data collection, we used individuals who live in Scandinavia. Such individuals are well accustomed to using social commerce platforms such as Amazon, so their use and consumption of information are likely to differ from users that use this platform for the first time or have limited experience. In addition, we did not perform a separate analysis to check how younger vs older users would reach a conclusion about which product to select, and what types of information or aspects were more important in determining their selection. Finally, each platform that is used for social commerce presents different types of information and in different formats. For instance, platforms such as Instagram which are primarily used to promote products and secondarily to enable purchase of the products via third-party websites will likely result in different aspects being important for users. Thus, it is interesting to identify how the different types of social media platforms trigger different types of information requirements from their users.

### CRedit authorship contribution statement

**Patrick Mikalef:** Conceptualization, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Kshitij Sharma:** Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft. **Sheshadri Chatterjee:** Investigation, Writing – review & editing. **Ranjan Chaudhuri:** Investigation, Writing – review & editing. **Vinit Parida:** Supervision, Resources, Project administration. **Shivam Gupta:** Supervision, Resources, Project administration.

### Declaration of Competing Interest

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

### Data availability

The authors do not have permission to share data.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.dss.2023.114039>.

## References

- [1] S.-I. Amari, S. Wu, Improving support vector machine classifiers by modifying kernel functions, *Neural Netw.* 12 (6) (1999) 783–789.
- [2] A. Aribarg, E.M. Schwartz, Native advertising in online news: trade-offs among clicks, brand recognition, and website trustworthiness, *J. Mark. Res.* 57 (1) (2020) 20–34.
- [3] A.S. Atalay, H.O. Bodur, D. Rasolofoarison, Shining in the center: central gaze cascade effect on product choice, *J. Consum. Res.* 39 (4) (2012) 848–866.
- [4] V. Bachurina, S. Sushchinskaya, M. Sharav, E. Burnaev, M. Arsalidou, A machine learning investigation of factors that contribute to predicting cognitive performance: difficulty level, reaction time and eye-movements, *Decis. Support. Syst.* 155 (2022), 113713.
- [5] D. Bacić, R. Henry, Advancing our understanding and assessment of cognitive effort in the cognitive fit theory and data visualization context: eye tracking-based approach, *Decis. Support. Syst.* 163 (2022), 113862.
- [6] S. Banerjee, S. Bhattacharyya, I. Bose, Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business, *Decis. Support. Syst.* 96 (2017) 17–26.
- [7] R. Batista Duarte, D. Silva da Silva, V. de Albuquerque Brito, C.S. Lopes, A systematic literature review on the usage of eye-tracking in understanding process models, *Bus. Process. Manag. J.* 27 (1) (2021) 346–367.
- [8] M. Belkin, D. Hsu, S. Ma, S. Mandal, Reconciling modern machine-learning practice and the classical bias–variance trade-off, *Proc. Natl. Acad. Sci.* 116 (32) (2019) 15849–15854.
- [9] R. Boardman, H. McCormick, Attention and behaviour on fashion retail websites: an eye-tracking study, *Inf. Technol. People* 35 (7) (2022) 2219–2240.
- [10] B.M. Brand, R. Reith, Cultural differences in the perception of credible online reviews—the influence of presentation format, *Decis. Support. Syst.* 154 (2022), 113710.
- [11] T.T. Brunyé, A.L. Gardony, Eye tracking measures of uncertainty during perceptual decision making, *Int. J. Psychophysiol.* 120 (2017) 60–68.
- [12] R. Calandra, J. Peters, C.E. Rasmussen, M.P. Deisenroth, Manifold Gaussian Processes for Regression. 2016 International Joint Conference on Neural Networks (IJCNN), 2016.
- [13] D. Camors, Y. Trotter, P. Pouget, S. Gilardeau, J.-B. Durand, Visual straight-ahead preference in saccadic eye movements, *Sci. Rep.* 6 (1) (2016) 1–9.
- [14] B.T. Carter, S.G. Luke, Best practices in eye tracking research, *Int. J. Psychophysiol.* 155 (2020) 49–62.
- [15] P. Chandon, J.W. Hutchinson, E.T. Bradlow, S.H. Young, Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase, *J. Mark.* 73 (6) (2009) 1–17.
- [16] N. Chaudhuri, G. Gupta, V. Vamsi, I. Bose, On the platform but will they buy? Predicting customers' purchase behavior using deep learning, *Decis. Support. Syst.* 149 (2021), 113622.
- [17] J.V. Chen, B.-c. Su, A.E. Widjaja, Facebook C2C social commerce: a study of online impulse buying, *Decis. Support. Syst.* 83 (2016) 57–69.
- [18] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.* 20 (1995) 273–297.
- [19] R.-F. Day, Examining the validity of the Needleman–Wunsch algorithm in identifying decision strategy with eye-movement data, *Decis. Support. Syst.* 49 (4) (2010) 396–403.
- [20] S. Dowiasch, P. Wolf, F. Bremmer, Quantitative comparison of a mobile and a stationary video-based eye-tracker, *Behav. Res. Methods* 52 (2020) 667–680.
- [21] A.T. Duchowski, K. Krejtz, I. Krejtz, C. Biele, A. Niedzielska, P. Kiefer, M. Raubal, I. Giannopoulos, The index of pupillary activity: Measuring cognitive load vis-à-vis task difficulty with pupil oscillation, in: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018.
- [22] S.R. Ellis, J.D. Smith, Patterns of statistical dependency in visual scanning, in: *Eye Movements and Human Information Processing*, 1985, pp. 221–238.
- [23] M. Fei, H. Tan, X. Peng, Q. Wang, L. Wang, Promoting or attenuating? An eye-tracking study on the role of social cues in e-commerce livestreaming, *Decis. Support. Syst.* 142 (2021), 113466.
- [24] V.V. Gavrishchaka, M.E. Koepke, O.N. Ulyanova, Ensemble learning frameworks for the discovery of multi-component quantitative models in biomedical applications, in: *2010 Second International Conference on Computer Modeling and Simulation*, 2010.
- [25] S. Goyal, K.P. Miyapuram, U. Lahiri, Predicting consumer's behavior using eye tracking data, in: *2015 Second International Conference on Soft Computing and Machine Intelligence (ISCMI)*, 2015.
- [26] D.J. Graham, J.L. Orquin, V.H. Viesschers, Eye tracking and nutrition label use: a review of the literature and recommendations for label enhancement, *Food Policy* 37 (4) (2012) 378–382.
- [27] J. Guerreiro, P. Rita, D. Trigueiros, Attention, emotions and cause-related marketing effectiveness, *Eur. J. Mark.* 49 (11/12) (2015) 1728–1750.
- [28] J. Gwizdzka, Exploring eye-tracking data for detection of mind-wandering on web tasks, in: *Information Systems and Neuroscience: NeuroIS Retreat 2018*, 2019.
- [29] R.S. Hessels, D.C. Niehorster, M. Nyström, R. Andersson, I.T. Hooge, Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers, *R. Soc. Open Sci.* 5 (8) (2018), 180502.

- [30] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, J. Van de Weijer, *Eye Tracking: A Comprehensive Guide to Methods and Measures*, OUP Oxford, 2011.
- [31] K. Holmqvist, S.L. Örbom, I.T. Hooge, D.C. Niehorster, R.G. Alexander, R. Andersson, J.S. Benjamins, P. Blignaut, A.-M. Brouwer, L.L. Chuang, Eye tracking: empirical foundations for a minimal reporting guideline, *Behav. Res. Methods* 55 (1) (2023) 364–416.
- [32] R.E. Hostler, V.Y. Yoon, Z. Guo, T. Guimaraes, G. Forgionne, Assessing the impact of recommender agents on on-line consumer unplanned purchase behavior, *Inf. Manag.* 48 (8) (2011) 336–343.
- [33] C.-L. Hsu, K.-C. Chang, M.-C. Chen, The impact of website quality on customer satisfaction and purchase intention: perceived playfulness and perceived flow as mediators, *IseB* 10 (2012) 549–570.
- [34] T. Jaarsma, H. Jarodzka, M. Nap, J.J. van Merriënboer, H.P. Boshuizen, Expertise under the microscope: processing histopathological slides, *Med. Educ.* 48 (3) (2014) 292–300.
- [35] L. Jenke, K. Bansak, J. Hainmueller, D. Hangartner, Using eye-tracking to understand decision-making in conjoint experiments, *Polit. Anal.* 29 (1) (2021) 75–101.
- [36] Y. Jiang, B.W. Ritchie, M.L. Verreynne, Building tourism organizational resilience to crises and disasters: a dynamic capabilities view, *Int. J. Tour. Res.* 21 (6) (2019) 882–900.
- [37] Z. Jiang, J. Chan, B.C. Tan, W.S. Chua, Effects of interactivity on website involvement and purchase intention, *J. Assoc. Inf. Syst.* 11 (1) (2010) 34–59.
- [38] E.M. Kok, H. Jarodzka, Before your very eyes: the value and limitations of eye tracking in medical education, *Med. Educ.* 51 (1) (2017) 114–122.
- [39] C. Krischer, W.H. Zangemeister, Scanpaths in reading and picture viewing: computer-assisted optimization of display conditions, *Comput. Biol. Med.* 37 (7) (2007) 947–956.
- [40] Y.-F. Kuo, C.-M. Wu, W.-J. Deng, The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services, *Comput. Hum. Behav.* 25 (4) (2009) 887–896.
- [41] A. Lang, Dynamic human-centered communication systems theory, *Inf. Soc.* 30 (1) (2014) 60–70.
- [42] A. Lang, R.L. Bailey, Understanding information selection and encoding from a dynamic, energy saving, evolved, embodied, embedded perspective, *Hum. Commun. Res.* 41 (1) (2015) 1–20.
- [43] T. Mavlanova, R. Benbunan-Fich, G. Lang, The role of external and internal signals in E-commerce, *Decis. Support. Syst.* 87 (2016) 59–68.
- [44] N. Meilatinova, Social commerce: factors affecting customer repurchase and word-of-mouth intentions, *Int. J. Inf. Manag.* 57 (2021), 102300.
- [45] R.V. Menon, V. Sigurdsson, N.M. Larsen, A. Fagerström, G.R. Foxall, Consumer attention to price in social commerce: eye tracking patterns in retail clothing, *J. Bus. Res.* 69 (11) (2016) 5008–5013.
- [46] P. Mikalef, K. Sharma, I.O. Pappas, M. Giannakos, Seeking information on social commerce: an examination of the impact of user- and marketer-generated content through an eye-tracking study, *Inf. Syst. Front.* 23 (2021) 1273–1286.
- [47] A. Olsen, The Tobii iVT fixation filter, *Tobii Technol.* 21 (2012) 4–19.
- [48] M. Papoukova, P. Hajek, Two-stage consumer credit risk modelling using heterogeneous ensemble learning, *Decis. Support. Syst.* 118 (2019) 33–45.
- [49] L. Peng, W. Zhang, X. Wang, S. Liang, Moderating effects of time pressure on the relationship between perceived value and purchase intention in social E-commerce sales promotion: considering the impact of product involvement, *Inf. Manag.* 56 (2) (2019) 317–328.
- [50] M.H. Phillips, J.A. Edelman, The dependence of visual scanning performance on saccade, fixation, and perceptual metrics, *Vis. Res.* 48 (7) (2008) 926–936.
- [51] D.A. Pissner, D.M. Schnyer, Support vector machine, in: *Machine Learning*, Elsevier, 2020, pp. 101–121.
- [52] S. Prasad, S.L. Galetta, Eye movement abnormalities in multiple sclerosis, *Neurol. Clin.* 28 (3) (2010) 641–655.
- [53] L.P. Prieto, K. Sharma, L. Kidzinski, M.J. Rodríguez-Triana, P. Dillenbourg, Multimodal teaching analytics: automated extraction of orchestration graphs from wearable sensor data, *J. Comput. Assist. Learn.* 34 (2) (2018) 193–203.
- [54] X. Qiu, L. Zhang, Y. Ren, P.N. Suganthan, G. Amarantunga, Ensemble deep learning for regression and time series forecasting, in: *2014 IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL)*, 2014.
- [55] E.M. Reingold, E.D. Reichle, M.G. Glaholt, H. Sheridan, Direct lexical control of eye movements in reading: evidence from a survival analysis of fixation durations, *Cogn. Psychol.* 65 (2) (2012) 177–206.
- [56] Y. Rogers, H. Sharp, J. Preece, *Interaction Design: Beyond Human-Computer Interaction*, John Wiley & Sons, 2023.
- [57] M. Russo, M. Thomas, D. Thorne, H. Sing, D. Redmond, L. Rowland, D. Johnson, S. Hall, J. Krichmar, T. Balkin, Oculomotor impairment during chronic partial sleep deprivation, *Clin. Neurophysiol.* 114 (4) (2003) 723–736.
- [58] M. Ryan, N. Krucien, F. Hermens, The eyes have it: using eye tracking to inform information processing strategies in multi-attributes choices, *Health Econ.* 27 (4) (2018) 709–721.
- [59] M. Salehan, D.J. Kim, Predicting the performance of online consumer reviews: a sentiment mining approach to big data analytics, *Decis. Support. Syst.* 81 (2016) 30–40.
- [60] A. Savoy, G. Salvendy, Factors for customer information satisfaction: user approved and empirically evaluated, *Int. J. Human-Comput. Interact.* 32 (9) (2016) 695–707.
- [61] S. Schroeder, J. Hyönä, S.P. Liversedge, Developmental eye-tracking research in reading: introduction to the special issue, *J. Cogn. Psychol.* 27 (5) (2015) 500–510.
- [62] M. Shojaeizadeh, S. Djasasbi, R.C. Paffenroth, A.C. Trapp, Detecting task demand via an eye tracking machine learning system, *Decis. Support. Syst.* 116 (2019) 91–101.
- [63] K. Tzaflikou, N. Protopoulos, Diagnosing user perception and acceptance using eye tracking in web-based end-user development, *Comput. Hum. Behav.* 72 (2017) 23–37.
- [64] B.T. Vincent, R. Baddeley, A. Corneani, T. Troscianko, U. Leonards, Do we look at lights? Using mixture modelling to distinguish between low- and high-level factors in natural image viewing, *Vis. Cogn.* 17 (6–7) (2009) 856–879.
- [65] T. Vuori, M. Olkkonen, M. Pölönen, A. Siren, J. Häkkinen, Can eye movements be quantitatively applied to image quality studies?, in: *Proceedings of the Third Nordic Conference on Human-Computer Interaction*, 2004.
- [66] D. Wang, X.R. Luo, Y. Hua, J. Benitez, Big arena, small potatoes: a mixed-methods investigation of atmospheric cues in live-streaming e-commerce, *Decis. Support. Syst.* 158 (2022), 113801.
- [67] Q. Wang, S. Yang, M. Liu, Z. Cao, Q. Ma, An eye-tracking study of website complexity from cognitive load perspective, *Decis. Support. Syst.* 62 (2014) 1–10.
- [68] S.V. Wass, T.J. Smith, M.H. Johnson, Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults, *Behav. Res. Methods* 45 (2013) 229–250.
- [69] E. Wästlund, P. Shams, M. Löfgren, L. Witell, A. Gustafsson, Consumer perception at point of purchase: evaluating proposed package designs in an eye-tracking lab, *J. Bus. Retail Manag. Res.* 5 (1) (2010) 42–51.
- [70] M. Wedel, R. Pieters, A review of eye-tracking research in marketing, *Rev. Mark. Res.* (2017) 123–147.
- [71] J. Willems, C.J. Waldner, J.C. Ronquillo, Reputation star society: are star ratings consulted as substitute or complementary information? *Decis. Support. Syst.* 124 (2019), 113080.
- [72] T.-T. Wong, Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation, *Pattern Recogn.* 48 (9) (2015) 2839–2846.
- [73] B. Wooley, S. Bellman, N. Hartnett, A. Rask, D. Varan, Influence of dynamic content on visual attention during video advertisements, *Eur. J. Mark.* 56 (13) (2022) 137–166.
- [74] Y. Wu, E.W. Ngai, P. Wu, C. Wu, Fake online reviews: literature review, synthesis, and directions for future research, *Decis. Support. Syst.* 132 (2020), 113280.
- [75] J. Yang, R. Sarathy, J. Lee, The effect of product review balance and volume on online shoppers' risk perception and purchase intention, *Decis. Support. Syst.* 89 (2016) 66–76.
- [76] L. Yang, O. Toubia, M.G. De Jong, A bounded rationality model of information search and choice in preference measurement, *J. Mark. Res.* 52 (2) (2015) 166–183.
- [77] S.-F. Yang, An eye-tracking study of the elaboration likelihood model in online shopping, *Electron. Commer. Res. Appl.* 14 (4) (2015) 233–240.
- [78] V.Y. Yoon, R.E. Hostler, Z. Guo, T. Guimaraes, Assessing the moderating effect of consumer product knowledge and online shopping experience on using recommendation agents for customer loyalty, *Decis. Support. Syst.* 55 (4) (2013) 883–893.
- [79] W.H. Zangemeister, T. Liman, Foveal versus parafoveal scanpaths of visual imagery in virtual hemianopic subjects, *Comput. Biol. Med.* 37 (7) (2007) 975–982.
- [80] K.Z. Zhang, M. Benyoucef, Consumer behavior in social commerce: a literature review, *Decis. Support. Syst.* 86 (2016) 95–108.
- [81] K. Zhao, P. Zhang, H.-M. Lee, Understanding the impacts of user- and marketer-generated content on free digital content consumption, *Decis. Support. Syst.* 154 (2022), 113684.

**Patrick Mikalef** is a Professor in Data Science and Information Systems at the Department of Computer Science. In the past, he has been a Marie Skłodowska-Curie post-doctoral research fellow working on the research project “Competitive Advantage for the Data-driven Enterprise” (CADENT). He received his B.Sc. in Informatics from the Ionian University, his M.Sc. in Business Informatics from Utrecht University, and his Ph.D. in IT Strategy from the Ionian University. His research interests focus on the strategic use of information systems and IT-business value in turbulent environments. He has published his work in over 150 international conferences and peer-reviewed journals including the *Journal of the Association for Information Systems*, *European Journal of Information Systems*, *Journal of Business Research*, *British Journal of Management*, *Information and Management*, *Industrial Management & Data Systems*, and *Information Systems and e-Business Management*. He serves as a senior editor of the *European Journal of Information Systems* and as editor of *Information of Management*. Patrick Mikalef is also a distinguished member of the Association for Information Systems (AIS).

**Kshitij Sharma** is an Associate Professor in Human-Computer Interaction and Collaborative/cooperative learning. His doctoral work was in the area of using multimodal data (EEG, eye-tracking, facial expressions, audio, dialogues, blood pressure, skin conductance, heart rate) to explain the differences between and predict, experts and novice groups; good and poor students; functional and non-functional groups. The main context for the application of his research has been education. His research interests are primarily in the area of Applied Machine Learning, Artificial Intelligence, and Human-Computer Interaction (HCI) with a heavy emphasis on groups' behavior and physiological data such as eye-tracking, EEG, facial expressions (theoretical and practical methods in digital interaction). He seeks to understand relations between users' data (EEG, eye-tracking, system log data, users' actions) and the profile of the user (expertise, motivation, strategy, performance) based on empirical experimentation (controlled experiments) and mixed methods analysis (utilizing a multitude of digital technologies). The knowledge gained from these studies is then used to provide feedback to the group or adapt to the needs of the group in a proactive manner. For this effort, in his studies, he has combined eye-tracking and users' actions to

provide more comprehensive results through data science, statistics, and machine learning practices.

**Sheshadri Chatterjee** is a post-doctoral research scholar at the Indian Institute of Technology Kharagpur, India. He has completed a PhD from the Indian Institute of Technology Delhi, India. He is having work experience in different multinational organizations such as Microsoft Corporation, Hewlett Packard Company, IBM, and so on. Sheshadri has published research articles in several reputed journals such as Government Information Quarterly, Information Technology & People, Journal of Digital Policy, Regulation and Governance, and so on. Sheshadri is also a certified project management professional, PMP from Project Management Institute (PMI), USA, and completed PRINCE2, OGC, UK, and ITIL v3 UK.

**Ranjan Chaudhuri** was a Fulbright Fellow to the University of Alabama in Huntsville in 2012. He is an Associate Professor in Marketing Sciences at the National Institute of Industrial Engineering, Mumbai. In the recent past, Dr. Chaudhuri also served as a Faculty at the Indian Institute of Technology Kharagpur and the Indian Institute of Technology Delhi. Dr. Chaudhuri has over twenty years of industrial, teaching and research experience. Dr. Chaudhuri holds an MBA in Marketing and a PhD in Management Sciences. Dr. Chaudhuri authored/coauthored >120 publications in referred National and International Journals and Conference Proceedings and contributed chapters in seven books and authored one monograph published by a leading press in European Union.

**Vinit Parida** is a Chaired Professor for Entrepreneurship and Innovation, Associate Editor for Journal of Business Research, Member of Swedish ministry's high-level group on digital transformation of Swedish industry, Scientific Leader for NorrlandsNavet- A Kamprad Family Foundation Center for SMEs Growth and Innovation in Northern Sweden,

and Board member for RE:Source Strategic Innovation Program (SIP) Vinnova. His research results have been published in 200+ leading international peer-reviewed journals, conferences, book chapters and industry/popular publications. Such as Academy of Management Journal, Journal of Management, Strategic Management Journal, Journal of Management Studies, Entrepreneurship Theory and Practice, Journal of Product Innovation Management, MIT Sloan Management Review, California Management Review, Long Range Planning, Industrial Marketing Management, Journal of Business Research, International Journal of Production Economics, Production and Operation Management, International Journal of Operations & Production Management, Strategic Entrepreneurship Journal, Entrepreneurship and Regional Development, Journal of Small Business Management, and Journal of Cleaner Production. He is active within different academic communities and has presented research results in well-known international conferences, such as Babson College Enterprise Research Conference, Research in Entrepreneurship and Small Business (RENT), The Annual ICSB World Conference, International Conference on Management of Technology (IAMOT), International Product Development Management Conference, and CRIP IPSS conference.

**Shivam Gupta** is a Professor at NEOMA Business School, France with a demonstrated history of working in the higher education industry. Skilled in Statistics, Cloud Computing, Big Data Analytics, Artificial Intelligence and Sustainability. Strong education professional with a Doctor of Philosophy (PhD) focussed in Cloud Computing and Operations Management from Indian Institute of Technology (IIT) Kanpur. Followed by PhD, postdoctoral research was pursued at Freie Universität Berlin and SUSTech, China. He has completed HDR from University of Montpellier, France. He has published several research papers in reputed journals and has been the recipient of the International Young Scientist Award by the National Natural Science Foundation of China (NSFC) in 2017 and winner of the 2017 Emerald South Asia LIS award.