MEASURING THE EFFECTIVENESS OF ONLINE SALES BY CONDUCTING A/B TESTING

MJERENJE UČINKOVITOSTI ONLINE PRODAJE PROVOĐENJEM A/B TESTIRANJA

MARKET TRŽIŠTE

Miroslav Mandić[a], Iva Gregurec[b], Uglješa Vujović[c]

[a] University of Zagreb, Faculty of Economics & Business, Trg J. F. Kennedyja 6, 10000 Zagreb, CROATIA, e-mail: mmandic@efzg.hr

[b] University of Zagreb, Faculty of Organization and Informatics, Pavlinska 2, 42000 Varaždin, CROATIA, e-mail: iva.gregurec@foi.unizg.hr

[c] dm-drogerie markt Croatia, Kovinska ulica 5a, 10090 Zagreb, CROATIA, e-mail: ugljesa.vujovic@dm.hr

Abstract

Purpose – The aim of this paper is to demonstrate the application of A/B testing for measuring the effectiveness of online sales in order to determine which changes to a website's user interface have the greatest effect on the improvement of key performance indicators (KPIs) of online sales.

Design/Methodology/Approach – A total of five A/B tests were conducted, four of which concerned the manipulation of user interface elements whereas one examined the difference in KPIs depending on the quality of the search engine used. Testing was conducted from January to July 2021 on a sample of a minimum of 7,000 visitors of the website of a company operating on the Croatian market.

Findings and Implications – The conducted tests show that sometimes, as can be seen from the results of the first A/B test, the existing version of the website should be kept. However, as shown by the second, fourth, and fifth A/B test, changes to a website's user interface can be of significant help in improving KPIs. The third A/B test highlighted the need for multiple tests of the same user interface element in order to achieve the full potential of an online business.

Sažetak

Svrha – Cilj je rada prikazati primjenu A/B testiranja mjerenja učinkovitosti online prodaje kako bi se utvrdilo koje promjene korisničkog sučelja na web-stranici u najvećoj mjeri utječu na poboljšanje ključnih pokazatelja uspješnosti online prodaje.

Metodološki pristup – Ukupno je provedeno pet A/B testova, od kojih su se četiri odnosila na manipuliranje elementima korisničkog sučelja, dok je jedan ispitivao razliku u ključnim pokazateljima uspješnosti ovisno o kvaliteti tražilice. Testiranje je provedeno od siječnja do srpnja 2021. godine na uzorku od minimalno sedam tisuća posjetitelja web-stranice poduzeća koje posluje na hrvatskom tržištu.

Rezultati i implikacije – Rezultati provedenog testiranja pokazuju kako se ponekad, kao što je vidljivo iz rezultata prvog A/B testa, treba zadržati postojeća verzija web-stranice. No, isto tako, kao što je to slučaj u provedenom drugom, četvrtom i petom A/B testu promjene korisničkog sučelja, web-stranice mogu znatno pripomoći u poboljšanju ključnih pokazatelja uspješnosti. Treći A/B test prikazuje potrebu višestrukih testiranja istog elementa korisničkog sučelja kako bi se postigao puni potencijal online poslovanja.

Limitations – The general limitations of A/B testing, such as its focus on short-term goals, the lack of insights into real consumer behavior, and the use of CMS systems all affect the results of conducted research.

Originality – As the analysis of existing research papers established the absence of a unified division of KPIs, their summarization and use for the purposes of the present research is the main contribution of this study.

Keywords – A/B testing, effectiveness, key performance indicators (KPIs), website, online sales

Ograničenja – Opća ograničenja A/B testiranja, poput usmjerenosti na kratkoročne ciljeve, nemogućnosti uvida u stvarno ponašanje potrošača te korištenje CMS sustava također utječu i na provedeno istraživanje.

Doprinos – Analizom dosadašnjih radova, ustanovljeno je kako ne postoji jedinstvena podjela ključnih pokazatelja uspješnosti. Stoga je njihovo sumiranje i korištenje za potrebe provedenog istraživanja osnovni doprinos rada.

Ključne riječi – A/B testiranje, učinkovitost, ključni pokazatelji uspješnosti, web-stranica, online prodaja

1. INTRODUCTION

Today, when companies – especially those conducting their sales in an online environment – have to create or maintain a competitive advantage, measuring the success of their marketing campaigns is of essence. This can be very challenging for a number of reasons. By studying key performance indicators (KPIs) such as conversion rate, number of orders, cart abandonment rate, bounce rate, time spent on the website, number of unique visitors, and other indicators, companies gather important information about the features of online sales as well as about the behavior of their customers (Esteller-Cucala, Fernandez & Villuendas, 2019). Based on this, companies can gain insights on whether the decisions of marketers have positive or negative impacts on their online business. One way to obtain such data is through A/B testing. A/B testing has been a widely used and increasingly popular technique of controlled experimentation performed on live users (Tamburrelli & Margara, 2014). Nowadays, this method of testing plays a significant role in evaluating the impact that new changes on a website have on its users (Esteller-Cucala et al., 2019). The concept of A/B testing is based on an online experiment, conducted by showing different variations of a website to different users and measuring the collected data to find out which variation is the most effective for turning them into customers (Siroker & Koomen, 2015; King, Churchill & Tan, 2017; Kohavi & Longbotham, 2017). By conducting A/B testing, in addition to storing and analyzing raw data, clear answers to questions about how to improve the efficiency of an entire company and find the optimal combination of website elements for its consumers can be obtained; this, in turn, provides not only a better user experience but also instills trust in a particular brand or company (Grundy, 2016). The aim of this study is to conduct A/B testing to measure the effectiveness of online sales so that companies can determine which changes to the user interface of its website have the greatest effect on the KPI improvement.

The literature search revealed no previous studies encompassing all the concepts of this research. Additional searches were conducted in all combinations, as well as individually, for the following terms: A/B testing, performance measurement and KPIs, online sales, measurement of online sales effectiveness, and effectiveness of transactional websites. Therefore, the present study fills that gap by integrating these concepts and conducting primary research aimed at yielding useful and applicable results.

In most cases, the research papers analyzed each of the mentioned concepts individually but no paper was found to utilize A/B testing to measure the effectiveness of online sales and determine which changes to a website's user interface have the greatest effect on the improvement of KPIs of online sales.

The research questions in this paper are:

R1: Which changes to a website's user interface have the greatest effect on the improvement of online sales KPIs?

R2: What is the importance of personalizing a website's user interface for the improvement of online sales KPIs?

2. THEORETICAL BACKGROUND

2.1. A/B test

As instruments of quantitative research, controlled experiments best show the connections between any changes that may be made and their impact on user behavior (Kohavi, Longbotham, Sommerfield & Henne, 2009). Such experiments generate a large amount of data that can be analyzed using various data mining techniques to gain a deeper understanding of factors that affect the goals of a particular company. Companies that use controlled experiments, applying clear evaluation criteria, can improve their systems with automated optimization and real-time analysis (Gallo, 2017).

According to Box, Hunter, and Hunter (2005), the controlled experiment theory is approximately 100 years old and dates back to Fisher's experiments at the Rothamsted Agricultural Experimental Station in England in the 1920s. With the topic of offline experiments being well developed in statistics (taken from Kohavi & Longbotham, 2017), marketers adopted it during the 1960s and 1970s (Gallo, 2017) to perform tests in order to evaluate different forms of direct marketing. A/B testing in its current form began to be applied in the 1990s with the growth of the Internet (Kohavi & Longbotham, 2017) and is often associated with websites and applications.

A/B tests are also referred to as split testing, bucket testing, randomized experiments, control/treatment tests, and online field experiments (Kohavi & Longbotham, 2017; Esteller-Cucala et al., 2019). They are defined as a random experimentation process in which two or more versions of a variable (web page, web page element, email, etc.) are simultaneously displayed to different segments of website visitors to determine which version performs better and which improves the observed KPIs of the website (Mullin, 2022). This is a standard way of evaluating consumer involvement or satisfaction with a new feature, service, or product on a website.

A/B testing takes several forms: classic A/B tests, split tests or redirect tests, and multivariate or "MVT" tests. Classic A/B tests usually present two versions of a website at the same URL to online users, thus enabling the comparison of two or more versions of the same element on the same web page. Split tests involve changing the content or design at the same URL. They redirect the website owner's traffic to several different URL addresses and each of the offered variations is completely redesigned in the website code (Brebion, 2015). Multivariate or MVT testing is an experiment in which multiple elements of a website are changed and tested to determine which combination of elements of a website leads to the maximum positive impact on conversion (Ahluwalia, 2020).

The A/B testing process starts with a decision on what to test, followed by that on how to evaluate the performance of the first decision. To run the test, two groups of visitors, that is live users visiting a website, are randomly assigned one of two variants or two different versions of a website – variant A (the control group), which is usually the current version, and variant B (the treatment group), which is usually a newer version with minor differences, of the system being evaluated. The two variants are compared based on certain metrics of interest related to user preferences to gain insights into which version has a greater impact on the KPIs being observed. After the testing is conducted, the variant that shows a statistically significant improvement is retained, while the other is rejected (Kohavi et al., 2009; Tamburrelli & Margara, 2014).

Certain papers in the literature have compared A/B testing with other tests (Zhou, Yuan, Huang, Zhang & Kaner, 2022; Quin & Weyns, 2022). Additionally, papers that indicate limitations of A/B testing and propose various improvements were also identified. Primarily, it has been highlighted that A/B testing requires a significant amount of time to conduct (several weeks), which affects the cost of the test (Gilotte et al., 2018; NeCamp et al., 2019; Masoero et al., 2023). The sequential randomized trial (SRT) has been proposed as a solution because data collection can be terminated following a smaller number of observations, allowing decisions to be made earlier (Shi et al., 2022; Bondarenko, 2019; Quin & Weyns, 2022; Xie & Liu, 2018; Sheng & Wang, 2023; Niculescu et al., 2021; Firmenich et al., 2019). Esteller-Cucala et al. (2020) pointed out the experimental pitfalls of implementing A/B testing, with the fact that most companies do not have a clear procedure for selecting their evaluation metrics as the basis for A/B testing being the main pitfall. Zhou, Kroehl, Meier and Kaizer (2023) mentioned the need for large sample sizes (ranging from 500 to 250,000 per arm) when conducting A/B testing. Martín et al. (2021) propose A/B testing based on multi-armed bandit (MAB) algorithms to improve A/B testing.

Moreover, there is also a group of papers pointing to the merits of A/B testing as a key tool for testing web interfaces. Kohavi et al. (2009) outlined A/B testing as a powerful technique for evaluating the success of a specific design element, particularly when it comes to live websites. Li, Zhou, Luo, and Dong (2022) noted on the crucial importance of A/B testing in the software industry and its use in combination with machine learning to improve user experience. Hagar and Diakopoulos (2019) highlighted A/B testing as a key tool in content optimization, with click-through rate being a key KPI.

Regarding the studies that use A/B testing in analyzing web interfaces, special mention should be made of papers that emphasize the importance of web interfaces: Esteller-Cucala et al. (2020), Jylhä and Hamari (2020), Jankowski, Hamari, and Wątróbski (2019), Lestari, Muslim, and Moch (2019), Chen and Zhai (2023), and Tounekti, Ruiz-Martínez, and Skarmeta-Gómez (2021). The latter authors (Tounekti et al., 2021) highlighted the ease of use, usefulness, security, confidentiality, privacy, payment method preferences, visual interface design, and credibility as the key factors for accepting a new web interface.

However, the cited studies failed to identify changes in the web interface that have the most significant impact on the key performance indicators of online sales. This leads us to the first research question:

R1: Which changes to a website's user interface have the greatest effect on the improvement of online sales KPIs?

Lestari et al. (2019) define the user interface as the visual component of a computer application or operating system where the user interacts with the computer or software. Most user interfaces are designed with a focus on usability and efficiency. Parboteeah, Valacich, and Wells (2009) argued that the user interface serves as a stimulus that influences the cognitive and affective responses of online users when interacting with a website, ultimately affecting their impulsive buying intentions. Purwati (2011)

noted that the user interface has become a key success factor as users engage with sellers in cyberspace through the seller's website. The design elements being explored include navigation, home link, search, cart viewing, help, account management, catalog, and personalization functions. Designing the user interface for e-commerce websites is highly challenging. E-commerce sites must cater to a wide range of users, involve significant user interactivity, and remain user-friendly. In their study, Lestari et al. (2019) suggested that, when designing the user interface for e-commerce, greater attention should be given to the use of color, appearance, structure, and information quality. Vila, González, Vila, and Brea (2021) pointed to web design as a crucial factor in ensuring a good interface that can meet consumer needs. According to Yang (2019), with retailers are increasingly focusing on providing relevant content to individual customers' shopping paths, they should place customers at the center of their website design process, understanding that customers and their behavior should be the cornerstone of the design process.

Sukendar and Susena (2022) stressed that creating a good user interface begins with understanding user characteristics, goals, skills, and preferences. Factors to consider when creating the user interface include a user-friendly design that does not require excessive design elements, consistent design to make users feel comfortable, consideration of item placement to attract attention and improve readability, use of non-contrasting colors to avoid excessive attention, the importance of typography in creating hierarchy and clarity, consideration of font type, size, and layout to enhance readability, ensuring the system can communicate what is happening, notifications to users of errors, status changes, and actions (Sela, Lavie, Inbar, Oppenheim & Meyer, 2015; Miraz, Excell & Ali, 2016). Furthermore, Swasty & Adriyanto (2017) emphasized the importance in the first phase on creating a website that is attractive, easy to use, and capable of attracting visitors. In addition to being user-friendly, another advantage

of a website consists in its ability to communicate effectively with an audience and provide a unique online experience. Bortko, Jankowski, Bartków, Pazura, and Śmiałkowska (2020) suggested that, before designing, careful consideration should be given to the arrangement and intensity of graphic elements in order to avoid being too conspicuous while still gently capturing people's attention. Chen and Zhai (2023) highlighted the hover function as an essential interaction method in the user interface for shopping, which is closely linked to users' perceived ease of use and usefulness. They suggested studying the impact of dynamic signs on users of shopping websites. Usability is crucial in user interface design, indicating how user-friendly and appealing a website is to visitors (Alqudsi & Alkhaledi, 2015). Usability for any website includes copy, creating a visual flow, call to action, color, and conventions. Usability issues in interface design as well as searching and browsing are important factors to consider. Bartoš and Habarta (2019) noted that clickable elements must retain sufficient cues to suggest clickability. Signaling the ability to click using cues such as text, color, size, shape, borders, and positioning can give interactive components a proper appearance. The most crucial factor in attracting clicks is the quality of the link text. Link text should be unique and descriptive, start with keywords, and contain a call-to-action text. The color blue is the safest choice for the link color, but other colors work as well, provided that the links are visible in the body text. If there is no particular reason to prefer another color, it is recommended to have blue link text as the safest choice. Optimizing clickable elements can increase the conversion rate, which is crucial for website owners and businesses seeking to profit from their online operations. Swasty and Adriyanto (2017) highlighted five design principles in creating a compelling web design: clarity, good layout, good visual hierarchy, effective use of color, and easy navigation. On e-commerce websites, important elements include context, logo, text, and background color, while motion graphics influence emotions and mood when consumers are browsing the website.

Khlaisang (2017) also mentioned five criteria that are important in creating a new website: web technology, content, design/presentation, interaction, and creativity. Key components of interfaces include the following: logo, search tool, language bar, login tool, navigation bar, banner, modal tab, and accordion panel (Li et al., 2022). González-Mena, Del-Valle-Soto, Corona, and Rodríguez (2022) outline the importance of a website's dynamism, stating that elements such as size, color, page position, location, and the amount of information are crucial for user experience and evoking emotions. In the context of user experience, Cai, He, Dai, and Zhu (2018) pointed to the importance of the interface flow design. A complete shopping website should have at least two basic functions: an external service function and an internal management function. These two fundamental functions are indispensable and work together to ensure the normal operation of business activities. The external service function is primarily oriented towards consumers. Functional website design aims to enrich these two basic functions so that the entire website can meet user requirements. Chen and Soo (2018) stress that interface design should provide users with a sense of ease, while its logic and interaction should align with user thinking (Lu, Sebe, Hytnen & Tian, 2011). As noted by Zhou et al. (2022), a number of studies in the literature have investigated the visual aspects of user interaction by measuring users' eye movements. The arrangement of information, interface color, and the size and placement of login icons all affect how quickly users complete an activity (Foris, Tecau, Hartescu & Foris, 2020; Reyes, Rodríguez & Muñoz, 2020). Designing an appropriate user interface for complex website interfaces with a large amount of data describing user interactions should be recognized as a challenging task (Więckowska & Rudnicka, 2023).

Some recent research studies, including those conducted by Martín, Jiménez-Martín, Mateos,

and Hernández (2021); Kern, Hoek, and Hienert (2023); and Bleier and Eisenbeiss (2015), emphasize the importance of personalization of the web interface in e-commerce and online advertising. However, there is a gap in the literature regarding the significance of personalizing a website's user interface in the improvement of online sales KPIs, leading to the following research question:

R2: What is the importance of personalizing a website's user interface for the improvement of online sales KPIs?

Closely related to the user interface is personalization, which can be defined as the customization of web content and appearance to deliver the right content to the right person in the right format at the right time (Sundar & Marathe, 2010; Ghorab, Zhou, O'Connor & Wade, 2013; Krouska, Kabassi, Troussas & Sgouropoulou, 2022; Ismail, Hussein, Harous & Khalil, 2023). Karat, Brodie, Karat, Vergo, and Alpert (2003) and Toch, Wang, and Cranor (2012) used the phrase "personalizing websites" to refer to the use of personal information about individuals to tailor their website experience. Košir, Kononenko, and Bosnić (2014) argue that user profiling is a crucial initial step in personalizing web services and building recommendation systems. Sundar and Marathe (2010) suggest that personalization involves manipulating content and the interface, allowing online companies to leverage this opportunity to cater to their users' needs. Desai (2021) has pointed to personalization being used to meet individual user needs, enhance competitiveness in the market, and effectively manage information. Effectively designed personalization is crucial for improving performance (perceived usefulness, ease of use, pleasure, satisfaction, and control) and its impact on users' intention to revisit a personalized website. A personalized/adaptive website contains different information architecture, interface design, and navigation based on the type of user categorized in user modeling and business rules (Bunt, Conati & McGrenere, 2009). Xu and Wang (2022) discuss the interac-

tive design of a personalized website's search interface based on visual communication, which includes design of the overall architecture, navigation module design, search module design, link module design, interactive layout module design, and visual rendering module design. Recommendation systems are closely related to personalization and play a vital role in providing personalized recommendations for relevant products, according to Sulikowski and Zdziebko (2021). Their research indicates that the appearance of recommendation interfaces on e-commerce websites is essential for their attractiveness and, consequently, their sales. Earlier Heinz, Linxen, Tuch, Fraßeck, and Opwis (2017) highlighted various techniques such as using advertising banners on websites, wish lists, or other persuasive sales strategies, or allowing users to customize the website, to increase product sales.

E-commerce companies have widely adopted web personalization techniques, including recommendation systems, to influence user behavior and achieve customer retention (Krishnaraju, Mathew & Sugumaran, 2016; Jankowski et al., 2019). Research by Ferretti, Mirri, Prandi, and Salomoni (2017) heavily relied on personalization techniques for adapting web content to user habits, mainly with the aim of offering suitable products and services. Personalization is based on user preferences and needs (Jylhä & Hamari, 2020), with Chung, Wedel, and Rust (2016) suggesting that adaptive personalization may work better than customer customization (de Bellis, Hildebrand, Ito, Herrmann & Schmitt, 2019) in autonomously tailoring a product. Adaptive personalization has several key characteristics (Chung et al., 2016): it is performed automatically using algorithms, does not require proactive effort from users, observes customer behavior and adjusts the product over time. It is assumed that adaptive personalization will yield better results over time as it learns more about the user. Both Paliouras (2012) and Ramadan (2023) draw attention to the fact that, when users communicate online, they are co-creators of content and active participants

MARKET
TRŽIŠTE

in building digital interpersonal relationships in online communities. This shift from multichannel to omnichannel marketing is seen as a key evolution. Personalization technologies offer powerful tools to enhance user experience in a wide range of systems, but they also raise new privacy concerns (Sulikowski & Zdziebko, 2021; Dalal et al., 2022; Raber & Krüger, 2022).

2.2. Online sales

Online sales as a concept refers to a process implying a change of ownership or the right to use goods and services through electronically connected devices that communicate interactively with each other (VanHoose, 2011). In recent years, online sales have played an increasingly important role in retail sales. According to available data, at almost 5 trillion US dollars they accounted for almost 19% of retail sales in the world in 2021 (Global E-Commerce Share of Retail Sales 2026, 2022). In 2023, global online sales are expected to make up 20.8% of total global retail sales and amount to 6.3 trillion US dollars (Global Ecommerce Sales Growth Report for 2021-2026, 2022), and the assumption is that these figures will continue to grow, making online sales an increasingly lucrative option for businesses. The biggest expansion of online sales took place during the Covid-19 pandemic. The emergence of the pandemic at the global level led to numerous restrictive measures aimed at limiting social contact (Gregurec, Tomičić Furjan & Tomičić-Pupek, 2021). The response of end consumers to the reduction of working hours or complete closure of the stores whose products or services are not necessary to satisfy physiological needs has been a significant increase in the use of the internet to purchase such products and services (Gregurec, Tomičić-Pupek & Kutnjak, 2021). All this led to the emergence of an increasing number of online stores and market saturation and creation of as much value for end consumers as possible. Namely, the more value companies can deliver, the greater are the chances of their success.

2.3. Measuring the effectiveness of online sales

While there are numerous definitions of performance measurement, in general, they are focused on the process of collecting, analyzing, and reporting data on the performance of a company in achieving its goals. According to Franceschini, Galetto, and Maisano (2018), performance measurement is "the ongoing monitoring and reporting of program accomplishments, particularly progress towards pre-established goals." Baki (2020) argued that performance measurement is to be considered "the practice of quantifying the efficiency and effectiveness of an action." Therefore, performance measurement is necessary for decision-making at all levels: it can help improve the understanding of a business model and user data, allowing for a precise allocation of the website content and better product placement, as well as targeting of promotions and discounts, increasing sales, offering cross-selling options, etc. (Keenan, 2022; Ecommerce Metrics + KPIs, n.d.).

Performance metrics provide a real representation of important business activities and enable greater precision in their execution. Metrics can be used in two ways. First, they can be used to assess internal performance, which can be the basis for the strategies and actions of the company itself. While this application of metrics is internally oriented, another option is to use metrics as benchmarks from which comparisons can be made within the industry and against other norms (Fink, 2006). Therefore, measures and performance metrics are needed to test and discover the viability of business plans, and it would be utterly impossible to ensure a clear path for progress and the achievement of goals without them (Baki, 2020).

KPIs are widely adopted performance measurement tools, representing a set of metrics or measures that focus on critical aspects of a company's performance that are most important for ensuring its current and future success (Muntean, Tarnaveanu & Ion, 2016). Therefore, they are seen as a quantitative index that reveals the key

factors of a company's success. KPIs should be quantified and aligned with company goals. Given that KPIs clarify the main responsibilities and serve as a basis for identifying the performance measurement indicators of different departments within the company, performance evaluation can be established on a quantitative basis. The fact that the establishment of clear and achievable KPIs is key to effective performance management should also be mentioned (Tsai & Cheng, 2012).

By tracking KPIs, online businesses can identify sales progress or improve their customer-facing services. The initial step in the process of using the right indicators is setting goals and understanding the areas of business that affect those goals. For each goal – be it is an increase in sales, optimization of marketing processes, or improvement of services – different indicators should be employed to successfully reach that goal. By analyzing these factors, a company can select indicators to increase sales, increase traffic on its website, or reduce calls to the company's customer support. The main goal of any company engaged in online sales is to increase sales, but this goal in most cases implies a whole series of actions related to the required number of active website visitors, the price of products sold, competitive prices, and even the rate of cart abandonment (Ahmed et al., 2017).

As KPIs can be specific to a company's website, by monitoring and observing metrics that include KPIs specific to websites, companies can eventually gain the ability to determine the relative success or failure of their company, as well as the effectiveness of their website (Turner, 2010). The website of any business, whether transactional or non-transactional, serves as a key tool for ensuring a solid online position and visibility (Roy & Sharma, 2021). Given that this paper focuses on transactional websites engaged in sales, the main goals are most often an increase in online sales revenue (aimed at selling products and services offered by the company by converting website visitors into customers) and an increase in online sales traffic (aimed

at attracting visitors to the company's website) (Ecommerce Metrics + KPIs, n.d.).

According to Sultan (2022), KPIs are a set of metrics used to measure and quantify the success of specific business components. Gräve (2019) said that marketing professionals primarily rely on quantitative metrics. The commonly accepted KPIs (Sultan, 2022) for multi-channel companies include Sales Volume, Average Unit Price, Gross Profit Margin, Operating Expenses to Sales Ratio, Inventory Turnover Ratio, Days Inventory Outstanding, Accounts Receivable Turnover Ratio, and Total Assets Turnover Ratio.

Furthermore, Jansen, Jung, and Salminen (2022) compared four standard analytical metrics from Google Analytics with SimilarWeb using one-year average monthly data for 86 websites from 26 countries and 19 industry verticals. They utilized the following KPIs: total visits, unique visitors, bounce rates, and average session duration. Bressolles and Lang (2020) believed e-fulfillment, defined as the fulfillment of orders placed through the internet, including warehousing, picking and order preparation, distribution and delivery, and returns, to be a key KPI. Bazett, Bowden, Love, Street, and Wilson (2005) said the main driver of profitability in the retail sector was the share of sales of high-margin products. They highlighted the potential of new channels (such as the web) to increase this share through higher-margin top-up shops on the Internet.

2.4. KPIs used for measuring the effectiveness of online sales

In order to analyze existing research studies, relevant databases such as Scopus, WoS, Emerald, Science Direct, EbcsoHost, and Proquest were searched. It focused on the studies that deal with the main topics of this paper, including A/B testing, performance measurement and KPIs, online sales, measurement of online sales effectiveness, and effectiveness of transactional websites in general.

When it comes to the measurement of effectiveness and online sales, all available databases

were searched, but the papers found were related to advertising, specifically experiments on advertising effectiveness, website effectiveness analysis, web interface, and user experience (UX). Their connection to A/B testing was found to be methodological, as these papers employed some form of experimentation. Unfortunately, no papers encompassing all the concepts relevant to the research of this study were found. Therefore, the following section provides an overview of papers grouped according to the concepts they cover.

Experiments on advertising effectiveness

Chae, Bruno, and Feinberg (2019) emphasized the importance of profiling individual consumers in terms of their response to ad repetition. Goldfarb and Tucker (2015) investigated the relationship between standardization and creativity on ad effectiveness. Johnson et al. (2017) tested ghost ads, where experimental ads are visible to the ad platform and experimenter but invisible to the control group of consumers.

Website evaluation

Li and Wang (2011) conducted a study on website evaluation, including dimensions such as use, information quality, responsiveness, visual appearance, security/privacy, interactivity, trust, fulfillment, personalization, advertising/persuasion, playfulness, and technology integration. Bhat, Bevans, and Sengupta (2002) organized web metrics around key objectives, which include assessing (1) website or ad popularity or exposure, (2) stickiness (the ability of a website to attract and retain user attention) and customer relationship quality, (3) message relevance or usefulness to users, (4) co-marketing opportunities, and (5) efficacy of user targeting. These authors emphasize that there is no single best way to measure web effectiveness. It is advisable for web advertisers to carefully consider their goals and not rely on a single metric, but rather use multiple metrics when possible, to gain better insights. It may also be beneficial to integrate information from other methods. Arshad and

Ahmad (2020) mentioned that measuring website effectiveness involves analyzing web traffic, website speed, webpage load time, first view on different devices, the average time visitors spend on the website, and visitor behavior. Lewis and Reiley (2014), through their experiment, found a negative correlation between ad exposure and baseline purchasing, which deviates from the usual practice. They explained these results by saying that their sample included only those users that the advertiser intended to target with the advertising campaign.

User experience (UX)

Wilson and Djamasbi (2019) claimed that user experience (UX) research largely focuses on how users subjectively evaluate digital products, services, and software. Sánchez-Adame, Urquiza-Yllescas, and Mendoza (2020) noted that social networks are currently key commodities that allow people to share various content and opinions. According to Atoum, Almalki, Alshahrani, and Shehri (2021), user experience provides insights into users' perception of a product during its use or intended use. Lestari et al. (2019) aimed to assess the user interface of an official store by taking a user experience approach employing performance metrics, self-reported metrics, behavioral metrics, and issue-based metrics that evaluate performance, perception, behavior, and issues perceived by the user when interacting with the website.

The results of this analysis show that different authors propose different metrics for measuring website quality and KPIs. No research study was found that would correspond to all the KPIs that were considered by the A/B testing carried out for the purposes of this paper. Therefore, the summarization of KPIs can be considered a scientific contribution of this paper.

The KPIs of online sales identified in previous research and observed in this paper are:

- Shopping cart abandonment rate (Muntean et al., 2016; Ahmed et al., 2017; Priya, Subha & Balamurugan, 2017; Saleem, Uddin,

Habib-ur-Rehman, Saleem & Aslam, 2019; Moreno, 2020; Keenan, 2022),

• Average Revenue per Visitor (ARV) (Siroker & Koomen, 2015; Muntean et al., 2016; Kohavi & Longbotham, 2017),

• Conversion Rate (CR) (Muntean et al., 2016; Miikkulainen et al., 2018; Ahmed et al., 2017; Saleem et al., 2019; Keenan, 2022; Siroker & Koomen, 2015),

• Average number of Products in an Order (APO) (Muntean et al., 2016),

• Average Order Value (AOV) (Saleem et al., 2019),

• and three key indicators related to website content: Bounce rate (Keenan, 2022; Roy & Sharma, 2021; Moral, Gonzalez & Plaza, 2014; Turner, 2010; Plaza, 2011), Time spent on page or Average time on site (Boswell, 2011; Rano & Sungkur, 2019; Moral et al., 2014; Turner, 2010; Plaza, 2011), and Pageviews or a total number of times a page is viewed (Luo, Ngi, Li & Tian, 2021; Ahmed et al., 2017; Plaza, 2011).

• Web Analytics and Digital Marketing KPIs (Saura, Palos-Sánchez and Cerdá Suárez, 2017) are traffic, unique visitors, page views, conversion rate, goals, cost per lead (CPL), leads, and surveys.

• Search engine KPIs (Saura et al., 2017): type of traffic, keywords, time on site, CTR, ROI, type of users, traffic sources, and user-friendliness. For Search Engine Optimization (SEO).

• KPIs (Saura et al., 2017): PPC (Pay Per Click), users, conversion, search traffic, and ROI.

• Digital Marketing ROI KPIs (Saura et al., 2017): traffic, unique users, leads, conversion, A/B testing, conversion rate, goals conversion rate, new visitors, and returning visitors.

• Type of advertising contracting models (Saura et al., 2017): CPI (Cost per impression) or CPM (Cost per thousand impressions),

PPC (Pay Per Click), CPC (Pay Per Click), CPL (Cost per Lead), and CPA (Cost per Action).

• KPIs for digital marketing (Saura et al., 2017): Conversion Rate, Goals Conversion Rate, Type of Users, Type of Sources, Keywords/Traffic of Non-branded Keywords, and Keyword Ranking.

Before discussing the results of the conducted A/B testing, it is important to mention that e-commerce activities are becoming increasingly important and extensive and that the customer experience when visiting a company's website is a very important component of the buyer-seller relationship. Therefore, the experience of a website visitor requires a lot of attention, with the goal of making the website a more productive source of user site visits, leads, engagement, purchase-related activities, sales and profits, brand equity, customer retention and satisfaction, CRM activities, and ultimately customer lifetime value (Dale Wilson, 2010).

3. RESEARCH

3.1. Research methodology

The goal of the conducted research was to investigate the application of A/B testing in measuring the effectiveness of online sales on the example of a website of a company that operates in the Croatian market.

The research questions are as follows:

R1: Which changes to a website's user interface have the greatest effect on the improvement of online sales KPIs?

R2: What is the importance of personalizing a website's user interface for the improvement of online sales KPIs?

Before conducting the A/B testing, changes to the website's user interface were precisely determined, which to the greatest extent affected the improvement of the KPIs related to online sales success. For the purposes of this study, quantitative research was conducted using a

controlled experiment in the form of A/B testing on a sample of a minimum of seven thousand visitors divided into treatment (experimental) and control groups. The experiment was conducted from January to July 2021 on a company operating in Croatia. The paper presents the results of five manipulations of the user interface of the company's website.

The first A/B test was related to the manipulation of the appearance of the menu containing assortment categories. It was carried out over a period of seven days, from 1 to 8 March 2021. In that period, the website was visited by more than 7,000 users, and the traffic was evenly divided into two groups of visitors who were presented with two different versions of the menu of assortment categories. During the A/B testing, group A was shown a version of the homepage with the existing horizontal menu, while group B was presented with a version of the homepage with a drop-down menu, which was displayed by pressing the so-called "hamburger" button located in the upper right corner of the homepage.

The second A/B test was conducted over a period of 14 days, from 9 to 22 March 2021. That A/B test included changes to the user interface on the product details page, where customer reviews (testimonials) of the product were implemented under the field "Check product availability in the store." Thus, the influence of the appearance of the homepage slider with novelties was examined in relation to the observed KPIs. In that period, more than sixteen thousand users visited the website. The visitors were again evenly divided into two groups, control and experimental. During the period of observation, the control group was shown a version of the homepage with a slider which contained novelties from the assortment that were manually selected by the marketing department of the company. At the same time, the experimental group was shown a personalized slider with novelties from the assortment, which were displayed based on previous purchases made and product reviews left by these visitors.

The third A/B test, or rather A/B/C test, was carried out over a period of 14 days, from 23 March to April 6, 2021. Over that time, more than 18,000 users visited the website, with the traffic divided equally between one control group and two experimental groups. The control group was shown a version of the product details page that contained only a "Selected by others" slider at the bottom. The first experimental group was shown a "Similar products" slider with the same placement on the product details page, and a "Selected by others" slider below it. The second experimental group was shown the "Selected by others" slider in the top position and the "Similar products" slider for related purchases below it.

The fourth A/B test was conducted over a period of 35 days, from 7 April to 12 May 2021. During that period, the website was visited by forty-two thousand users, and again the traffic was evenly divided between two groups of visitors. The first one was the control group being shown, during the observed period, a version of the product details page without prominently displayed reviews (testimonials) of the customers placed under the category "Check product availability in physical store." The second, experimental group was shown versions of the product details page with a prominent review of the product in question during the same period.

The fifth A/B test was carried out over a period of seven days, from 11 to 18 June 2021. Over that time, the website was visited by 19,000 users. Traffic was divided evenly into two groups. The control group was provided with an existing solution, the Hybris search engine. For the same purpose, the use of the Algolia search engine was provided to the experimental group. This A/B test aimed to determine the influence of the search engine on the observed KPIs.

3.2. Research results

In the first A/B test, the appearance of the menu containing assortment categories was manipulated. The test was carried out over a period of seven days, from 1 to 8 March 2021. In that period, the website was visited by more

than seven thousand users, and the traffic was evenly divided into two groups of visitors who were presented with two different versions of the menu of assortment categories. During the A/B testing, group A was shown a version of the homepage with the existing horizontal menu, while group B was presented with a version of the homepage with a drop-down menu, which was displayed by pressing the so-called "hamburger" button located in the upper right corner of the homepage. Table 1 shows the results obtained by this A/B test.

TABLE 1: Results of A/B test depending on the form of the menu of assortment categories

KPIs	Horizontal menu	Drop-down menu	Percentage of changes +/- (%)
Number of page views per visitor	2.17	2.114	-2.6%
Percentage of visitors who added a product to the cart (%)	96.38%	87.03%	-9.07%
Bounce rate (%)	43%	44.33%	+3.1%
Average revenue per visitor (EUR)	3.48	3.15	-9.6%

Source: research

The table shows that visitors who were presented with a horizontal menu of assortment categories on the homepage viewed an average of 2.17 pages with product details. A total of 96.38% of these visitors added at least one product to their online shopping cart. Of all visitors who were shown a horizontal menu, 43% of them viewed only the homepage and left the site, so the bounce rate was 43%. The average revenue per visitor who was shown the horizontal menu of assortment categories was EUR 3.48. On the other hand, visitors who were shown a drop-down menu viewed an average of 2.114 product detail pages during the observed period. 87.03% of these visitors added at least one product to their online shopping cart. Of all visitors who were shown a drop-down menu, 44.33% of them did not continue their visit beyond the homepage of the website, i.e. the bounce rate was 44.33%. The average revenue per visitor who was shown an alternative version of the assortment menu was EUR 3.15. The user problem that was aimed to be solved by the first A/B test is the desire of users to reach the relevant products as quickly as possible. From the presented results, it is evident that the version of the homepage with a drop-down menu of assortment categories, compared to the version of the homepage with a horizontal menu, recorded strong declines in all relevant observed KPIs. Visitors who were shown a drop-down menu viewed an average of 2.6% fewer product detail pages compared to the visitors who were shown a horizontal menu. The percentage of visitors who added a product to their online shopping cart is 9.07% higher in favor of the visitors who were shown a horizontal menu in the observed period. The bounce rate is 3.1% higher for the visitors who were shown a drop-down menu in the observed period. The last observed KPI in this A/B test was average revenue per visitor. As shown in Table 1, visitors who were presented with a drop-down menu generated 9.6% less revenue compared to visitors who were shown a horizontal menu. Considering that the version of the homepage with a drop-down menu recorded strong declines in all KPIs compared to the version with the horizontal menu of assortment categories, the idea of changing the appearance of the menu can be rejected and the existing solution of the horizontal menu kept.

The second A/B test was conducted over a period of 14 days, from 9 to 22 March 2021. The

A/B test included changes to the user interface on the product details page, where customer reviews (testimonials) of the product were implemented under the field "Check product availability in the store." Thus, the influence of the appearance of the homepage slider with novelties was examined in relation to the observed KPIs. In that period, more than 16,000 users visited the website. The visitors were again evenly divided into two groups, control and experimental. During that period, the control group was shown a version of the homepage with a slider which contained novelties from the assortment that were manually selected by the marketing department of the company. At the same time, the experimental group was shown a personalized slider with novelties from the assortment, which were displayed based on previous purchases made and product reviews left by these visitors. In the cases where there was no information about the visitors' past purchases, the novelty bar displayed the novelties sorted by the ratings other visitors had left. The results of this A/B test are shown in Table 2.

TABLE 2: Results of A/B test depending on the appearance of a novelty slider

KPIs	Manually selected novelties	Novelties displayed based on past purchases	Percentage of changes +/- (%)
Average revenue on novelties per visitor (all visitors)	EUR 4.10	EUR 4.29	+4.68%
Average revenue on novelties per visitor (registered visitors)	EUR 4.39	EUR 4.74	+7.96%
Average number of products added to the online shopping cart (all visitors)	0.0019	0.0079	+340%

Source: research

As can be seen from Table 2, visitors who were shown novelties manually selected by the marketing department of the company on the homepage generated an average revenue of EUR 4.10 per visitor. Within that total group, a subgroup of visitors who registered on the website generated an average revenue of EUR 4.39 per visitor. The average number of products added to the online shopping cart was 0.0019 for the control group. The second group, which included visitors who were shown a slider on the homepage with novelties from the assortment selected based on past purchases and product reviews, generated an average revenue of EUR 4.29 per visitor. Registered visitors from this (experimental) group generated an average revenue of EUR 4.74 per visitor. The average number of products added to the online shopping cart was 0.0079. The user problem that was attempted to be solved by the second A/B test is as follows: "As a customer of XY company, I want to find out as much information as possible about new products that might be of interest to me, but I find this process quite difficult." The idea of this A/B test was to investigate whether the revenue increases in a situation where visitors who start their session on the homepage are offered personalized novelty suggestions. Table 2 shows the positive impact of the personalized slider with novelties on all observed KPIs. Therefore, it can be concluded that visitors who were shown a personalized slider with novelties in the observed period generated 4.68% more revenue per visitor on average than visitors who were shown a slider with manually selected novelties. Registered visitors from this experimental group generated 7.96% more revenue per customer than registered visitors from the

control group. The average number of product additions to the online shopping cart and the total number of sessions on the homepage was 340% higher in the observed period in favor of the slider with personalized novelties. Considering the presented results, the sliding bar with personalized novelties has a positive effect on all observed KPIs and should be applied as such for all visitors.

The third A/B test, or rather the A/B/C test, was carried out over a period of 14 days, from 23 March to 6 April 2021. During this period, more than 18,000 users visited the website, and the traffic was divided equally between one control group and two experimental groups. The control group was shown a version of the product details page that contained only a "Selected by others" slider at the bottom. The first experimental group was shown a "Similar products2 slider with the same placement on the product details page, and a "Selected by others" slider below it. The second experimental group was presented with the "Selected by others" slider in the top position and the "Similar products" slider for related purchases below it. The results of this A/B/C test are shown in Table 3.

TABLE 3: Results of A/B/C test with "Similar products" and "Selected by others" slider

KPIs	Only a "Selected by others" slider	"Similar products" slider in the top position	"Selected by others" slider in the top position
Add-to-cart rate (%)	1.76%	2.07%	2.02%
Slider interactions per session	2.23	2.26	2.24
Average revenue generated based on recommended products (EUR)	EUR 4.30	EUR 5.17	€4.53
Average number of products per order	3.17	3.04	3.08
Price of purchased recommended products (EUR)	EUR 3.02	€3.47	€3.28

Source: research

Based on the analysis of this A/B/C testing, 1.76% of visitors from the control group, who were shown only a "Selected by others" slider on the product details page, added at least one product from the slider to their online shopping cart. The average revenue generated by this group was EUR 4.30 per visitor. Visitors from the control group ordered an average of 3.17 products per order. The average price of the purchased products that visitors from this group selected from the slider of recommended products was EUR 3.02. On the other hand, 2.07% of visitors from the first experimental group, which during the observed period was shown a "Similar products" slider in the top position of the product details page, added at least one product to their online shopping cart. Those visitors had 2.17 interactions per session with the slider featuring recommended products. The average revenue generated by the first experimental group was EUR 5.17 per visitor. Visitors from this group ordered an average of 3.04 products per purchase. The average price of purchased products from the slider of recommended products chosen by this group was EUR 3.47. Of all visitors who were shown a "Selected by others" slider at the top of the product details page, 2.02% of them added at least one product to their online shopping cart. This experimental group of visitors achieved an

average of 2.24 interactions with the slider per session in the observed period. The average revenue was EUR 4.53. Visitors from the second experimental group ordered an average of 3.08 products per purchase, while the average price of purchased products from the slider of recommended products chosen by this group was EUR 3.28. The user problem that the third A/B/C test tried to solve is the following: "As a buyer of XY company products, I find it difficult to find products on the site that match my buying plan." This test is based on the idea that the customer achieves more interactions and adds more products to their online shopping cart if they are shown better suggestions of similar products and products that other customers have purchased. To find the answer, the A/B/C test was conducted, with changes in the KPIs and the significance of these changes shown in Table 4.

TABLE 4: Results of A/B/C test depending on the position of the slider with recommended products

KPIs	A-B change	A-C change	Significance of A-B change	Significance of A-C change
Add-to-cart rate (%)	+0.31	+0.26	✓	✓
Slider interactions per session	+0.03	+0.01	✓	✗
Average revenue generated based on recommended products (EUR)	+EUR 0.87	+EUR 0.23	✓	✗
Average number of products per order	-0.13	-0.09	✓	✗
Price of purchased recommended products (EUR)	+EUR 0.45	+ EUR 0.27	✓	✓

Source: research

The letter A in Table 4 represents the version of the product details page that only displayed a "Selected by others" slider. The letter B stands for the version of the product details page that had a "Similar products" slider at the top, and the letter C for the version of that page that had a "Selected by others" slider at the top. Table 4 shows that both alternative options exhibited growth in all observed KPIs, but the significance of these changes is quite different. In the case that only one A/B test with a version of the product details page that has a "Selected by others" slider at the top is conducted, the company could choose to implement that version of the product details page. However, it is evident from Table 4 that the company would not reach its full potential of online sales by doing so. Considering the thesis set out by the marketing department of the company and the results obtained by A/B/C testing, a decision was made to implement a version of the product details page that displays a "Similar products" slider at the top and a "Selected by others" slider at the bottom. Table 4 shows that visitors from this group buy fewer products on average but achieve higher revenue per visitor by EUR 0.87 compared to the control group and allocate EUR 0.45 more for the purchase of the recommended product compared to the control group. From these results, it is evident that customers are ready to allocate a larger amount of their income for purchases if they are offered the right recommendations.

The fourth A/B test was carried out over a period of 35 days, from 7 April to 12 May 2021. During that time, the website was visited by 42,000 users, and again the traffic was evenly divided between two groups of visitors. The first one was

the control group, which was shown during the observed period a version of the product details page without prominently displayed reviews (testimonials) of the customers placed under the category "Check product availability in physical store." The second, experimental group was shown versions of the product details page with a prominent review of the product in question during the same period. The results of this A/B test are shown in Table 5.

TABLE 5: Results of A/B test depending on customer reviews

KPIs	Group that was not shown a featured product review	Group that was shown a featured product review	Percentage of changes +/- (%)
Mobile Conversion Rate (%)	2.93%	3.098%	+5.74%
Desktop Conversion Rate (%)	22.26%	22.103%	-2.2%
Add-to-cart rate (%)	82.48%	83.255%	+0.94%

Source: research

As can be seen from Table 5, the conversion rate achieved by using mobile phones was 2.93% for the control group. The same group achieved a conversion rate of 22.26% using the computer. Of the total number of visitors from this control group, 82.48% added at least one product to their online shopping cart. The experimental group of visitors who were shown a featured product review on a product details page achieved a 3.098% conversion rate when using mobile phones. The same group achieved a 22.103% desktop conversion rate. In addition, 83.26% of visitors from the experimental group added at least one product to their online shopping cart. The user problem that was addressed by the fourth A/B test is as follows: "As a customer, I want to make a quick and reliable purchase decision. In that decision, I am supported by the experiences of other customers." The idea behind this test is that visitors achieve a higher add-to-cart rate and a higher conversion rate if testimonials from previous customers are highlighted on the product details page. As shown in Table 5, the experimental group of visitors using mobile phones achieved a 5.74% higher conver-

sion rate than the control group. Visitors from the experimental group who used a computer achieved a 2.2% lower conversion rate than the control group. On average, the experimental group added 0.94% more products to their online shopping cart than the control group. Table 5 also shows that the testimonials of others have a positive influence on visitors who use a mobile phone, but a negative influence on those who use a computer. Considering that the stated idea is only partially confirmed, it is up to the company to decide which user interface to use. At the same time, applying additional tests to gather more data is advisable.

Finally, the last A/B test was carried out over a period of seven days, from 11 to 18 June 2021. In the observed period, the website was visited by 19,000 users. Traffic was divided evenly into two groups. The control group was provided with an existing solution, the Hybris search engine while the experimental group was able to search through the Algolia search engine for the same purpose. This A/B test aimed to determine the influence of the search engine on the observed KPIs.

TABLE 6: Results of A/B test based on used search engine

KPIs	Hybris search engine	Algolia search engine	Percentage of changes +/- (%)
Average revenue per visitor (EUR)	EUR 0.87	EUR 0.88	+1.76%
Add-to-cart rate (%)	89.7%	91.61%	+2.13%
Conversion rate (%)	6.36%	6.49%	+1.97%
Interaction rate (%)	34.3%	35.13%	+2.43%

Source: research

Table 6 shows the average revenue per visitor from the control group at EUR 0.87. Also, 89.7% of all visitors from this group added at least one product to their online shopping cart after using the Hybris search engine. Meanwhile, 6.36% of visitors ended their session with a conversion, and 34.3% of them interacted with one of the search results. The experimental group, which used the Algolia search engine, generated an average revenue of EUR 0.88, with 91.61% of visitors from this group adding at least one product to their online shopping cart. The conversion rate for this experimental group was 6.49%. Of the total number of visitors from this group, 35.13% interacted with one of the search results. The user problem that the fifth A/B test tried to solve is: "As a user of the website search engine, I have trouble finding products that are relevant to me." The idea that was put forward for the implementation of this A/B test is that, if products relevant to the customer are displayed in the first search results, the rate of interaction with the search engine and the rate of adding products to the online shopping cart will increase. To conduct the A/B test, the control group was provided with the use of the Hybris search engine whereas the experimental group used the Algolia search engine. Table 6 shows how searches using the Algolia search engine led to an increase in all observed KPIs. Therefore, companies would be advised to implement the Algolia search engine as a basic solution for all visitors.

3.3. Summary of results

Based on the results presented in Table 1, it is evident that the version of the homepage with a drop-down menu, compared to the version of the homepage with a horizontal menu of assortment categories, recorded strong declines in all relevant observed KPIs. The results in Table 2 show a positive impact of the personalized slider with novelties on all observed KPIs, thus implying that the sliding bar with personalized novelties should be applied as such for all visitors. As evident from the results in Table 4, customers are prepared to allocate a larger amount of their income for purchases if they are offered the right recommendations. Table 5 indicates that the testimonials of others have a positive influence on visitors who use a mobile phone but a negative influence on visitors who use a computer. Table 6 shows how searches using the Algolia search engine increased all observed KPIs.

3.4. Research limitations and recommendations for future research

The basic limitations of this research study are related to the general limitations of A/B testing, namely its focus on short-term goals and lack of insights into the actual behavior of website visitors. By conducting A/B tests, raw statistical data are obtained on limited KPIs tested on a group of users over a period of time, without contextual, behavioral, or motivational confirming or disconfirming the data concerned. Also, a more technical limitation of A/B testing refers to the use of a content editing system instead of

a framework solution. Although content management systems (CMS) are very easy to use and make changes in the user interface, unlike framework solutions, they lack the ability to freely create web pages.

Recommendations for future research include implementing A/B/C testing wherever possible. The reason for testing at least two alternative user interface options is evident from Table 4. It was only with the third version of the website that the company found the strongest impact of the manipulated interface on KPIs. Failure to test with the help of a third user interface option would result in a loss of revenue for the company. In the future, it would also be worthwhile to conduct "MVT" testing, which is particularly useful when a company wants to test the impact of significant changes on a website or else when it wants to simultaneously test several elements of the user interface, compared to testing one specific element of the user interface for the purpose of deciding on improvements to the way companies do their business more quickly. Using any form of A/B testing is a good solution in the short term; however, combining it with more human-centered methods and techniques provides deeper insights into contextual, behavioral, or motivational data.

3.5. Discussion

R1: Which changes to a website's user interface have the greatest effect on the improvement of online sales KPIs?

A horizontal menu of assortment categories (Table 1), a sliding bar with personalized novelties (Table 2), and demonstrated suggestions of similar products and products that other customers have purchased (Table 3 and 4) are the changes to the website's user interface which have the greatest effect on the improvement of online sales KPIs. As previously mentioned, due to the lack of studies that integrate these concepts it is impossible to directly compare the results of this study to those of previous research. However, what can be compared with previous studies are the results in Table 6, which indicate

the importance of simple solutions. This is also supported by the following authors who emphasized the importance of simplicity: Raka and Setyohadi (2021) and Jankowski et al. (2019) advocated moderation of visual intensity (excessive visual intensity negatively affects user satisfaction and experience), Tounekti et al. (2021) emphasized the ease of use and usefulness, while Lestari et al. (2019) highlighted simplicity of appearance (Lu et al., 2011; Swasty & Adriyanto, 2017; Chen & Soo, 2018; Vila et al., 2021; Sukendar & Susena, 2022; Chen & Zhai, 2023).

R2: What is the importance of personalizing a website's user interface for the improvement of online sales KPIs?

The results of this research have confirmed that personalization has a crucial impact on the improvement of online sales KPIs (Table 2 and 4). Due to the aforementioned reasons (lack of identical studies), a direct comparison of research results is not possible. However, it is possible to connect these results with those of previous studies that emphasized the importance of personalization, such as those conducted by Košir et al. (2014), Chung et al. (2016), Sulikowski and Zdziebko (2021), Quin and Weyns (2022) who highlighted the importance of personalized recommendations; Esteller-Cucala et al. (2020) and Desai (2021) emphasized the importance of website personalization, while Martín et al. (2021) noted the importance of personalized services, and Bleier & Eisenbeiss (2015) highlighted the importance of personalized banners (Heinz et al., 2017).

When it comes to research in the field of KPIs, due to the lack of relevant studies, only a general alignment with the KPIs used in the research by Sultan (2022) and Saura et al. (2017) can be confirmed.

3.6. Theoretical and managerial implications

This study fills the research gap by integrating the following concepts: A/B testing, performance measurement and KPIs, online sales, measurement of online sales effectiveness, effectiveness

of transactional websites, website user interface, personalization, and conducting primary research that yields useful and applicable results. Five theoretical implications emerge from this paper. Firstly, A/B testing is one of the key tools for testing web interfaces (Goldstein, 2019; Mattos, Bosch, Olsson, Maryam Korshani & Lantz, 2020; Hasan, Chatterji & Koning, 2019; Kohavi et al., 2009) and can deliver more realistic results in comparison to eye tracking (Bartoš & Habarta, 2019; Sulikowski & Zdziebko, 2021; González-Mena et al., 2022; Zhou et al., 2022), especially when measuring online sales effectiveness. Secondly, A/B testing is a powerful technique for evaluating the success of a specific design element, particularly when it comes to a live website (Cruz-Benito et al., 2018; Li et al., 2022; Fabijan, Dmitriev, Holmstrom Olsson & Bosch, 2018; Hagar & Diakopoulos, 2019). No papers addressing the measurement of effectiveness and online sales were found to exist. Those identified are related to advertising, specifically experiments on advertising effectiveness, website effectiveness analysis, web interface, and user experience (UX), with A/B testing found among as one of the tools along with other similar forms of experimentation. Thus, the present study identified a significant gap in the literature with regard to the integration of the mentioned key concepts, specifically the KPIs of online sales identified in previous research. As e-commerce activities are becoming increasingly important and extensive, the customer experience when visiting a company's website is a very important component of the buyer-seller relationship (Dale Wilson, 2010).

The third implication is that designing an appropriate user interface is a complex and challenging task (Więckowska & Rudnicka, 2023), with most user interfaces focusing on usability and efficiency (Parboteeah et al., 2009; Purwati, 2011; Lestari et al., 2019; Vila et al., 2021). When designing a user interface, the key factors are the following: a user-friendly design that does not require excessive design elements, consistent design to make users feel comfortable, consideration of item placement to attract attention and improve readability, using non-contrasting colors to avoid excessive attention, the importance of typography in creating hierarchy and clarity, consideration of font type, size, and layout to enhance readability, ensuring the system can communicate what is happening, notifications to users of errors, status changes, and actions (Sela et al., 2015; Miraz et al., 2016; Swasty & Adriyanto, 2017; Sukendar & Susena, 2022). Usability is crucial in user interface design and for any website includes copy, creating a visual flow, call to action, color, and conventions (Bartoš & Habarta, 2019). The fourth implication, that is, how quickly users will complete the activity, is directly connected with the arrangement of information, interface color, and the size and placement of login icons (Foris et al., 2020; Reyes et al., 2020). The fifth implication, personalization (Chung et al., 2016), may work better than customer customization (de Bellis et al., 2019), because it is performed automatically using algorithms and does not require proactive effort from users.

For practitioners, it is important to emphasize the regular monitoring of all metrics and mandatory testing of changes on the web interface. This paper has six implications for practitioners. First, when creating a user interface, a horizontal menu should be preferred over a drop-down one. Second, the user interface on the product details page has a positive impact on the observed KPIs when combined with customer reviews (testimonials). Third, a personalized slider with novelties from the assortment has a positive impact on all observed KPIs. Fourth, customers are ready to allocate a larger amount of their income for purchases if they are offered the right (personalized) recommendations. Fifth, practitioners should be aware of the importance of analyzing the devices from which users browse the web interface and make purchases (mobile phone users achieved a higher conversion rate than computer users). In this context, the KPIs to analyze include mobile conversion rate (%), desktop conversion rate (%), and add-to-cart rate (%). Sixth, testimonials have a positive influence on visitors using a mobile phone, but a negative influence on visitors who use a computer.

4. CONCLUSION

The last few years have witnessed a global expansion of online sales, leading to a saturation of the market with online stores. It was this saturation that helped companies see the need to base their business on creating value for end consumers, and thus potential competitive advantage. By applying A/B testing, companies can gather data that will provide clear answers to questions about how to increase their efficiency. The aim of this study was to demonstrate how companies can measure the effectiveness of online sales using A/B testing to determine which changes to their website's user interface have the greatest effect on the improvement of online sales KPIs. The paper summarizes various metrics for measuring website quality and KPIs, which is considered to be its greatest scientific contribution. The conducted research shows how changing some elements of the user interface affects the observed KPIs. The first A/B test showed how decisions to change a certain element of the user interface without prior testing can have a detrimental impact on business. If the company were to implement a new version of the user interface without conducting A/B tests, this would lead to lower revenue per visitor and a lower rate of products added to the online shopping cart. The second A/B test showed that website visitors respond positively to a personalized offer of new products using a slider, resulting in more traffic per visitor. The third A/B test showed the need for multiple tests of the same user interface element to achieve its full potential. By conducting the fourth A/B test, the influence of other customers' reviews on the purchase decision was tested. However, that test failed to provide a definite answer to the best possible solution for the user interface. The incomplete results of that test opened the possibility for further tests and ideas about changes to the user interface. The last A/B test to have been conducted showed the impact of the quality of background processes on the observed KPIs. According to its results, customers are willing to pay a higher price if they can easily find what they are looking for. To conclude, although A/B testing has some limitations, it is the simplest form of controlled online experiment. Companies that embrace controlled experiments with clear evaluation criteria can improve their business with real-time analytics.

REFERENCES

1. Ahluwalia, S. (2020, September 4). *What is Multivariate Testing? When and How to Use MVT?*. Retrieved from: https://vwo.com/blog/multivariate-testing/
2. Ahmed, H., Jilani, D. T. A., Haider, W., Abbasi, M. A., Nand, S., & Kamran, S. (2017). Establishing Standard Rules for Choosing Best KPIs for an E-Commerce Business based on Google Analytics and Machine Learning Technique. *International Journal of Advanced Computer Science and Applications, 8*(5).
3. Alqudsi, T. M., & Alkhaledi, R. M. (2015). Website usability: The case of Kuwaiti middle school students. *Electronic Library, 33*(3), 557-572.
4. Arshad, M., & Ahmad, P. (2020). Measuring effectiveness of library websites with user trafffific analytics: A review. *Library Philosophy and Practice*, 1-16.
5. Atoum, I., Almalki, J., Alshahrani, S. M., & Shehri, W. A. (2021). Towards measuring user experience based on software requirements. *International Journal of Advanced Computer Science & Applications, 12*(11), 25-31.
6. Baki, R. (2020). Evaluating hotel websites through the use of fuzzy AHP and fuzzy TOPSIS. International Journal of Contemporary Hospitality Management, 13(12), 3747-3765.

7. Bartoš, P., & Habarta, F. (2019). Does the form of text play a role in click through the button on websites?. *Trendy v Podnikání, 9*(1), 3-8.

8. Bazett, M., Bowden, I., Love, J., Street, R., & Wilson, H. (2005). Measuring multichannel effectiveness using the balanced scorecard. *Journal of Direct, Data and Digital Marketing Practice, 6*(3), 224.

9. Bhat, S., Bevans, M., & Sengupta, S. (2002). Measuring users' web activity to evaluate and enhance advertising effectiveness. *Journal of Advertising, 31*(3), 97-106.

10. Bleier, A., & Eisenbeiss, M. (2015). Personalized online advertising effectiveness: The interplay of what, when, and where. *Marketing Science, 34*(5), 669-688.

11. Bondarenko, Y. (2019). *Sequential A/B testing*. Paper published at 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT), 488-491.

12. Bortko, K., Jankowski, J., Bartków, P., Pazura, P., & Śmiałkowska, B. (2020). Attracting user attention to visual elements within website with the use of Fitts's law and flickering effect. *Procedia Computer Science, 176*, 2756-2763.

13. Boswell, P. (2011, May). *Google Analytics: Measuring Content Use and Engagement*. Proceedings of 58th Annual Conference Society for Technical Communication. Sacramento, CA: Society for Technical Communication.

14. Box, G. E. P., Hunter J. S., & Hunter W. G. (2005). *Statistics for experimenters: design, innovation, and discovery*. Hoboken: Wiley.

15. Brebion, A. (2015, September 1). *How to Choose Between A/B Test, Split Test or Multivariate Test? AB Tasty*. Retrieved from: https://www.abtasty.com/blog/ab-vs-multivariate-test/

16. Bressolles, G., & Lang, G. (2020). KPIs for performance measurement of e-fulfillment systems in multi-channel retailing. *International Journal of Retail & Distribution Management, 48*(1), 35-52.

17. Bunt, A., Conati, C., & McGrenere, J. (2009). Mixed-Initiative interface personalization as a case study in usable AI. *AI Magazine, 30*(4), 58-64.

18. Cai, L., He, X., Dai, Y., & Zhu, K. (2018). Research on B2B2C E-commerce website design based on user experience. *Journal of Physics: Conference Series, 1087*(6), 62043.

19. Chae, I., Bruno, H. A., & Feinberg, F. M. (2019). Wearout or weariness? measuring potential negative consequences of online ad volume and placement on website visits. *Journal of Marketing Research, 56*(1), 57-75.

20. Chen, C., & Soo, K. C. (2018). The research on human engineering-based design of website interface of smart cellphone. *Journal of Physics: Conference Series, 1098*(1), 12012.

21. Chen, C., & Zhai, W. (2023). The effects of dynamic prompt and background transparency of hover feedback design on the user interface of shopping websites. *Asia Pacific Journal of Marketing and Logistics, 35*(4), 809-827.

22. Chung, T. S., Wedel, M., & Rust, R. T. (2016). Adaptive personalization using social networks. *Journal of the Academy of Marketing Science, 44*(1), 66-87.

23. Cruz-Benito, J., Vazquez-Ingelmo, A., Sanchez-Prieto, J. C., Theron, R., Garcia-Penalvo, F. J., & Martin-Gonzalez, M. (2018). Enabling adaptability in web forms based on user characteristics detection through A/B testing and machine learning. *IEEE Access, 6*, 2251-2265.

24. Dalal, R. S., Howard, D. J., Bennett, R. J., Posey, C., Zaccaro, S. J., & Brummel, B. J. (2022). Organizational science and cybersecurity: Abundant opportunities for research at the interface. *Journal of Business and Psychology, 37*(1), 1-29.

25. Dale Wilson, R. (2010). Using clickstream data to enhance business-to-business web site performance. Journal of Business & Industrial Marketing, 25(3), 177-187.

26. de Bellis, E., Hildebrand, C., Ito, K., Herrmann, A., & Schmitt, B. (2019). Personalizing the customization experience: A matching theory of mass customization interfaces and cultural information processing. *Journal of Marketing Research, 56*(6), 1050-1065.

27. Desai, D. (2021). Website personalization: Strategy for user experience design & development. *Turkish Journal of Computer and Mathematics Education, 12*(12), 3516-3523.

28. Ecommerce Metrics + KPIs: Measure Your Ecommerce Success. (n.d.). *BigCommerce*. Retrieved from: https://www.bigcommerce.com/articles/ecommerce/ecommerce-metrics/

29. Esteller-Cucala, M., Fernandez, V., & Villuendas, D. (2019). *Experimentation Pitfalls to Avoid in A/B Testing for Online Personalization*. 27th Conference on User Modeling, Adaptation and Personalization, 153-159.

30. Fabijan, A., Dmitriev, P., Holmstrom Olsson, H., & Bosch, J. (2018). *Online controlled experimentation at scale: An empirical survey on the current state of A/B testing*. Paper published at 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT), 68-72.

31. Ferretti, S., Mirri, S., Prandi, C., & Salomoni, P. (2017). On personalizing web content through reinforcement learning. *Universal Access in the Information Society, 16*(2), 395-410.

32. Fink, D. (2006). Value decomposition of e-commerce performance. *Benchmarking: An International Journal, 13*(1/2), 81-92.

33. Firmenich, S., Garrido, A., Grigera, J., Rivero, J. M., & Rossi, G. (2019). Usability improvement through A/B testing and refactoring. *Software Quality Journal, 27*(1), 203-240.

34. Foris, D., Tecau, A. S., Hartescu, M., & Foris, T. (2020). Relevance of the features regarding the performance of booking websites. *Tourism Economics: The Business and Finance of Tourism and Recreation, 26*(6), 1021-1041.

35. Franceschini, F., Galetto, M., & Maisano, D. (2018). *Designing Performance Measurement Systems: Theory and Practice of Key Performance Indicators*. Boston, MA: Springer.

36. Gallo, A. (2017). *A Refresher on A/B Testing. Harvard Business Review Home*. Retrieved from: https://hbr.org/2017/06/a-refresher-on-ab-testing

37. Ghorab, M. R., Zhou, D., O'Connor, A., & Wade, V. (2013). Personalised information retrieval: Survey and classification. *User Modeling and User-Adapted Interaction, 23*(4), 381-443.

38. Gilotte, A., Calauzènes, C., Nedelec, T., Abraham, A., & Dollé, S. (2018). *Offline A/B testing for recommender systems*. Ithaca: Cornell University Library.

39. Global Ecommerce Sales Growth Report for 2021-2026. (2022, April 26). *Shopify*. Retrieved from: https://www.shopify.com/blog/global-ecommerce-sales

40. Global e-commerce share of retail sales 2026. (2022, July). *Statista*. Retrieved from: https://www.statista.com/statistics/534123/e-commerce-share-of-retail-sales-worldwide/

41. Goldfarb, A., & Tucker, C. E. (2015). Standardization and the effectiveness of online advertising. *Management Science, 61*(11), 2707-2719.

42. Goldstein, S. (2019). *A novel technique for A/B testing using static prototypes*. Weave, 2(1), 35-47.

43. González-Mena, G., Del-Valle-Soto, C., Corona, V., & Rodríguez, J. (2022). Neuromarketing in the digital age: The direct relation between facial expressions and website design. *Applied Sciences, 12*(16), 8186.

44. Gräve, J. (2019). What KPIs are key? evaluating performance metrics for social media influencers. *Social Media + Society, 5*(3).

45. Gregurec, I., Tomičić Furjan, M., & Tomičić-Pupek, K. (2021). The Impact of COVID-19 on Sustainable Business Models in SMEs. *Sustainability, 13*(3).

46. Gregurec, I., Tomičić-Pupek, K., & Kutnjak, A. (2021). *Disruption caused business transformations in tourism-related industries*. 2021 IEEE Technology & Engineering Management Conference – Europe, 1-6.

47. Grundy, J. (2016). *10 Benefits of A/B testing*. Retrieved from: https://hbr.org/2017/06/a-refresher-on-ab-testing

48. Hagar, N., & Diakopoulos, N. (2019). Optimizing content with A/B headline testing: Changing newsroom practices. *Media and Communication, 7*(1), 117-127.

49. Hasan, S., Chatterji, A., & Koning, R. M. (2019). *Experimentation and startup performance: Evidence from A/B testing.* Cambridge: National Bureau of Economic Research.

50. Heinz, S., Linxen, S., Tuch, A. N., Fraßeck, L., & Opwis, K. (2017). Is it still where I expect it? Users' current expectations of interface elements on the most frequent types of websites. *Interacting with Computers, 29*(3), 325-344.

51. Ismail, H., Hussein, N., Harous, S., & Khalil, A. (2023). Survey of personalized learning software systems: A taxonomy of environments, learning content, and user models. *Education Sciences, 13*(7), 741.

52. Jankowski, J., Hamari, J., & Wątróbski, J. (2019). A gradual approach for maximising user conversion without compromising experience with high visual intensity website elements. *Internet Research, 29*(1), 194-217.

53. Jansen, B. J., Jung, S., & Salminen, J. (2022). Measuring user interactions with websites: A comparison of two industry standard analytics approaches using data of 86 websites. *PloS One, 17*(5), e0268212-e0268212.

54. Jylhä, H., & Hamari, J. (2020). Development of measurement instrument for visual qualities of graphical user interface elements (VISQUAL): A test in the context of mobile game icons. *User Modeling and User-Adapted Interaction, 30*(5), 949-982.

55. Karat, C. M., Brodie, C., Karat, J., Vergo, J., & Alpert, S. R. (2003). Personalizing the user experience on ibm.com. *IBM Systems Journal, 42*(4), 686-701.

56. Keenan, M. (2022, March 10). *15 Critical Ecommerce Metrics You Must Track in 2022.* Shopify. Retrieved from: https://www.shopify.com/blog/basic-ecommerce-metrics

57. Kern, D., Hoek, W., & Hienert, D. (2023). *Evaluation of a search interface for preference-based ranking -- measuring user satisfaction and system performance.* Ithaca: Cornell University Library.

58. Khlaisang, J. (2017). Proposing a new pedagogy-based website design: A usability test with lifelong learners. *Education and Information Technologies, 22*(4), 1713-1735.

59. King, R., Churchill, E. F., & Tan, C. (2017). *Designing with Data: Improving the User Experience with A/B Testing.* London: O'Reilly Media.

60. Kohavi, R., & Longbotham, R. (2017). *Online Controlled Experiments and A/B Testing. In C. Sammut & G. I. Webb (eds.), Encyclopedia of Machine Learning and Data Mining* (pp. 922-929). Boston, MA: Springer.

61. Kohavi, R., Longbotham, R., Sommerfield, D., & Henne, R. M. (2009). Controlled experiments on the web: Survey and practical guide. *Data Mining and Knowledge Discovery, 18*(1), 140-181.

62. Košir, D., Kononenko, I., & Bosnić, Z. (2014). Web user profiles with time-decay and prototyping. *Applied Intelligence, 41*(4), 1081-1096.

63. Krishnaraju, V., Mathew, S. K., & Sugumaran, V. (2016). Web personalization for user acceptance of technology: An empirical investigation of E-government services. *Information Systems Frontiers, 18*(3), 579-595.

64. Krouska, A., Kabassi, K., Troussas, C., & Sgouropoulou, C. (2022). Personalizing environmental awareness through smartphones using AHP and PROMETHEE II. *Future Internet, 14*(2), 66.

65. Lestari, R. A., Muslim, E., & Moch, B. N. (2019). User interface evaluation of official store for FMCG (fast moving consumer goods) products in e-commerce website using user experience approach. *IOP Conference Series: Materials Science and Engineering, 505*(1), 12079.

66. Lewis, R. A., & Reiley, D. H. (2014). Online ads and offline sales: Measuring the effect of retail advertising via a controlled experiment on yahoo. *Quantitative Marketing and Economics, 12*(3), 235-266.

67. Li, W., Zhou, Y., Luo, S., & Dong, Y. (2022). Design factors to improve the consistency and sustainable user experience of responsive interface design. *Sustainability, 14*(15), 9131.

68. Li, X., & Wang, Y. (2011). Measuring the effectiveness of US official state tourism websites. *Journal of Vacation Marketing, 17*(4), 287-302.

69. Lu, Y., Sebe, N., Hytnen, R., & Tian, Q. (2011). Personalization in multimedia retrieval: A survey. *Multimedia Tools and Applications, 51*(1), 247-277.

70. Luo, P., Ngai, E. W. T., Li, Y., & Tian, X. (2021). Effects of visit behavior on online store sales performance: Personal computer (PC) versus mobile channels. *Internet Research, 32*(3), 728-767.

71. Martín, M., Jiménez-Martín, A., Mateos, A., & Hernández, J. Z. (2021). Improving A/B testing on the basis of possibilistic reward methods: A numerical analysis. *Symmetry, 13*(11), 2175.

72. Masoero, L., Hains, D., & McQueen, J. (2023). *Leveraging covariate adjustments at scale in online A/B testing*. Ithaca: Cornell University Library.

73. Mattos, D. I., Bosch, J., Olsson, H. H., Maryam Korshani, A., & Lantz, J. (2020). *Automotive A/B testing: Challenges and lessons learned from practice*. Paper published at 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT), 101-109.

74. Miikkulainen, R., Iscoe, N., Shagrin, A., Rapp, R., Nazari, S., McGrath, P., Schoolland, C., Achkar, E., Brundage, M., Miller, J., Epstein, J., & Lamba, G. (2018). *Sentient ascend: AI-based massively multivariate conversion rate optimization*. Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence, 7696-7703.

75. Miraz, M. H., Excell, P. S., & Ali, M. (2016). User interface (UI) design issues for multilingual users: A case study. *Universal Access in the Information Society, 15*(3), 431-444.

76. Moral, P., Gonzalez, P., & Plaza, B. (2014). Methodologies for monitoring website performance: Assessing the effectiveness of AdWords campaigns on a tourist SME website. *Online Information Review, 38*(4), 575-588.

77. Moreno, Ó. (2020, June 17). *What is Cart Abandonment Rate and How Much is too Much?*. Retrieved from: https://www.instasent.com/blog/en/what-is-cart-abandonment-rate

78. Mullin, S. (2022, November). *A/B Testing: The Complete Guide with Expert Tips*. Retrieved from: https://www.shopify.com/blog/the-complete-guide-to-ab-testing

79. Muntean, M. I., Tarnaveanu, D., & Ion, A. R. (2016). E-Commerce Performance. Shopping Cart Key Performance Indicators. *Informatica Economica, 20*(1), 62-71.

80. NeCamp, T., Gardner, J., & Brooks, C. (2019). *Beyond A/B testing: Sequential randomization for developing interventions in scaled digital learning environments*. Ithaca: Cornell University Library.

81. Niculescu, I., Hu, H. M., Gee, C., Chong, C., Dubey, S., & Li, P. L. (2021). *Towards inclusive software engineering through A/B testing: A case-study at windows*. Proceedings of the 2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering Education and Training Track (ICSE-SEET), 180-187.

82. Paliouras, G. (2012). Discovery of web user communities and their role in personalization. *User Modeling and User-Adapted Interaction, 22*(1/2), 151-175.

83. Parboteeah, D. V., Valacich, J. S., & Wells, J. D. (2009). The influence of website characteristics on a consumer's urge to buy impulsively. *Information Systems Research, 20*(1), 60-78.

84. Plaza, B. (2011). Google Analytics for measuring website performance. *Tourism Management, 32*(3), 477-481.

85. Priya, V., Subha, S., & Balamurugan, B. (2017). Analysis of performance measures to handle medical E-commerce shopping cart abandonment in cloud. *Informatics in Medicine Unlocked, 8*, 32-41.

86. Purwati, Y. (2011). standard features of e-commerce user interface for the web. *Researchers World – Journal of Arts Science and Commerce, 2*(3), 77.