AR E-COMMERCE EXPERIENCE: CASE STUDY BRAND MAYBELLINE AT TOKOPEDIA

Lia Herna

Received January. 15, 2024; Revised April. 24, 2024; Accepted May. 27, 2024.

Abstract: Online shopping culture is one of the popular cultures that is currently developing. This popular culture develops along with the rapid development of new media. E-commerce is starting to utilize Augmented Reality as part of its features. Tokopedia launched an AR feature for consumers to try various makeup products. As AR applications in the consumer market are still emerging, studies on AR are insufficient. In this study, the authors analyzed the level of user acceptance of AR implementation in Tokopedia e-commerce by testing the user experience (UX) dimension using a mixed method. Qualitative methods with 5 respondents and quantitative interviews with a system usability scale (SUS) and a user experience questionnaire (UEQ) were conducted on 23 respondents. The results obtained, SUS got a value of 77.61 and UEQ with the value of each category being greater than the impression value. The AR feature has a very positive reception overall, indicating its potential for future use in all industries.

Keywords: augmented reality; user experience; maybelline; tokopedia

Introduction

Shopping is a cultural practice that individuals or groups engage in to fulfill their daily necessities. According to Hud- dleston and Minahan (2011), shopping activities involve evaluating a product or service, finding a store that offers the best product or service, locating the desired product or service in the store, and deciding to purchase it. Since the Covid-19 pandemic spread in Indonesia, the number of internet users and online shops has increased. By 2023, 77% of Indonesians have used the internet and 62.6% have bought products or services online (Kemp, 2023). Online shopping culture is a form of popular culture that is currently emerging. This popular culture evolves along with the advancement of communication technology, media development and the rapid growth of new media (Utamanyu and Darmastuti, 2022).

However, after the pandemic, consumers have resumed their old habits of shopping for their necessities by visiting physical stores and trying out the products offered there. SimilarWeb data shows a trend of declining visits to E-commerce sites in Indonesia. This phenomenon has affected some of the E-commerce sites with the largest traffic in Indonesia, such as Shopee, Tokopedia, Lazada and Blibli. Doni P. Joewono, the Deputy Governor of BI, estimated that the transaction value fell short of the target due to the relaxation of Covid-19 related activity restrictions, which encouraged people to go back to shopping at conventional stores (Ahdiat, 2023). Physical stores also enhance themselves...
by creating a different store atmosphere, which can provide a pleasant purchasing environment, attract consumers to visit and feel comfortable shopping and add value to the products sold (Nofiawati and Yuliandi, 2014). As presented by Sephora, its store lighting, luxurious interior, various testers of its cosmetic products that allow consumers to feel the difference of each colour, shape and fragrance, and its free eyebrow grooming service. Buying cosmetics, women often face difficulties in buying cosmetic products with the wrong colour several times before they can find the colour that suits their preference, so women tend to buy cosmetics offline.

Based on the buying culture that occurs in the cosmetic industry, e-commerce must prepare strategies that can accommodate and adapt to this situation. E-commerce has started to utilize Augmented Reality as part of its features. AR is defined as a technology (in mobile phones or supported devices) that enables virtual 3D objects to be presented in the real-world environment (Nofiawati and Yuliandi, 2014). This technology can enhance the online shopping experience in e-commerce in various ways. With AR, consumers can visualize how the product looks on them (Kim, Jiyeon and Forsythe, 2008) AR helps consumers to choose and see if the makeup they are going to buy matches their needs and tastes without having to go to a physical store. AR technology can also make us more confident and satisfied with our shopping decisions because we can see the results directly. AR technology in e-commerce is one of the trends that deserve attention in the future because it can provide a more realistic and enjoyable shopping experience.

One of the adopters is an e-commerce company made in Indonesia, Tokopedia, which launched an AR feature for consumers to try various makeup products, such as lipstick, foundation and others virtually before buying. To try this AR feature, simply type “Virtual Try On”, users can enter the search column on the Tokopedia application page, and then on that page, users can immediately try various products with AR that says “Try the items, let’s!” Users can directly choose up to 4 colour variants of products simultaneously and compare them directly. This AR feature can be enjoyed by Android and iOS users with the latest version of the Tokopedia application (Ubaidillah, 2019).

The author observes the growing trend of technology and online shopping culture in Indonesia, particularly AR technology. Website, marketplace such as Shopee, Blibli, Tokopedia have begun to employ this AR technology for buying cosmetic products. The user interface and user experience provided by the website and application are well-received by users and give suggestions to enhance quality and user satisfaction. It is vital for the industry to understand the extent of the impact of the AR feature in Tokopedia e-commerce on user experience, specifically in buying cosmetic products and staying competitive in the market. Mclean and Wilson (2019) stated that AR applications in consumer markets are still in the growth phase so studies on AR are still lacking. In this study, the authors analyze AR E-Commerce Experience: Case Study Brand Maybelline At Tokopedia according to the level of user acceptance of AR implementation in Tokopedia e-commerce by testing on user experience (UX) dimensions using system usability scale (SUS) by analyzing three categories namely acceptability ranges, grade scale and adjective ratings (Yulianto, Hartanto and Santosa, 2021), as well as user experience questionnaire (UEQ) measuring and evaluating two factors: usability aspects, which include user efficiency in completing tasks without additional effort; perspicuity, which refers to how easy a system is for users to use on their first try; dependability, which refers to the user’s sense of control over the system; and user
experience aspects, which also include the system’s attractiveness, stimulation, and novelty, as well as whether or not the user finds it to be interesting to use, innovative, and motivating overall. UEQ is a simple technique for measuring user experience that yields heuristic results for a system or digital product (Tanjungan, 2022).

Previous research, conducted by K.M. Kristi and N. Kusumawati (2008) regarding on the level of consumer acceptance and perception of Augmented Reality (AR) technology in Indonesia’s cosmetics Industry in this research they use Maybelline’s three augmented reality platforms: Shopee, social media, and their website by analyzing three categories of AR namely interactivity, vividness, and novelty, towards perceived easy to use, perceived usefulness and enjoyment and analyzed using PLS-SEM and Kruskall-Wallis. The results show that while vividness is not found to affect perceived ease of use, all characteristics of augmented reality (AR) influence indices of technology acceptability. Another study was conducted by Yulianto, Haryanto, and Santosa (2021) regarding the Evaluation of Interactive Books Based on Augmented Reality Using System Usability Scale and User Experience Questionnaire which evaluates books based on AR technology as a learning media for the Cirebon mask art. The overall test shows that the evaluation results of books based on AR technology have very good acceptance for use as a learning media for Cirebon mask dance art.

**Literature Review**

Augmented Reality (AR) is a technology that combines virtual objects with real environments, involving the overlay of computer graphics in the real world in real-time. A real-time direct or indirect picture of the physical environment that has been improved or supplemented by the addition of virtual, computer-generated data is another definition of augmented reality. Augmented reality integrates virtual and real-world items and is interactive in three dimensions. AR can display information in the form of text, images, video, sound, or 3D objects that can interact with the user. Through the transfer of virtual data to both their immediate surroundings and any indirect view of the real-world environment, such as live video streaming, augmented reality seeks to make the user’s life easier. (Mangtani, Bajpai, Sahasrabudhe and Wasule, 2022).

In online retail, augmented reality has a feature called AR virtual try-on, which combines the user’s real-time video with a digital overlay, creating a new experience where users can try on products as many times as possible before making their final decision. Users of the AR virtual try-on may experience a sense of actual product use. (Vieira, Rafael, and Agnihotri, 2022). A camera is needed to take images for AR virtual try-on applications. These images will then be integrated with the information you wish to present and displayed simultaneously on the user’s screen. (Hung, Chang and Ma, 2021). AR virtual try-on is an innovation used by e-commerce to display its products online (Plotkina and Saure, 2019). Virtual makeup applications with AR capabilities allow users to see products with just their eyes. By using AR, products such as clothing, glasses, lipstick, eye shadow, nail colour, and jewelry can be selected. Buyers do not need to visit the dressing room repeatedly to try on different clothes. This can reduce the possibility of infection spreading (Kim, Jieon and Forsythe, 2008).

Usability is defined as “the degree to which a system, product, or service can be used by specified users to achieve specific goals with effectiveness, efficiency, and satisfaction in a specified context of use” by the International Organization for Standardization (ISO 8241-11, 2018). Users are requested to complete a ques-
tionnaire about their experience with the current system or application after using it in order to rate its usability. (Upendi, Kurniawan and Panjaitan, 2019).

Methodology

This research used mixed methods. For the qualitative method, the author conducted in-depth interviews with 5 female respondents with an age range of 17 to 42 years old residing in Indonesia, who had bought makeup online and had used the e-commerce Tokopedia. This method was done online via Gmeet for 30 minutes. In the qualitative method, the author formulates some inquiries, namely in Table 1.

| Table 1. Interview discussion guide  
(Source: Personal documentation) |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interview section</strong></td>
</tr>
<tr>
<td>General background</td>
</tr>
<tr>
<td>General experience with online and offline shopping</td>
</tr>
<tr>
<td>Experience with AR as virtual try-on feature</td>
</tr>
<tr>
<td>AR virtual technology try-on trial</td>
</tr>
<tr>
<td>AR virtual try-on experience</td>
</tr>
</tbody>
</table>

The characteristics of the respondents from the quantitative method were the same as the qualitative method. The respondents were asked to perform usability testing by first accessing a provided link, which was the link of the Maybelline Brand on the Tokopedia page that had an AR feature, then performing some scenarios, and finally filling out a questionnaire. All questionnaires used Likert scale (5 and 7). There are 23 respondents that filled up the usability testing.

The scenario that was performed in this study was an instruction that could be done by the respondents on the Maybelline Brand page on Tokopedia that had an AR feature. The respondents were asked to:

1. Look at the “Virtual Try On” feature on your smartphone screen, then choose one lipstick colour that you like the most, and then add it to the cart.
2. Choose four lipstick colours that you like at once, then compare them simultaneously on one side-by-side screen, and then save the photo to the gallery.
3. Reject Choose your selfie photo from the photo gallery on your smartphone, then try this “Virtual Try On” feature on your photo by choosing one of the colours that you like.

This study uses two instruments for testing, namely SUS and UEQ. The System Usability Scale (SUS) was analyzed by three categories: acceptability ranges, grade scale and adjective ratings. Acceptability was used to see the user’s acceptance of the application or in other words the category that shows whether the SUS score of a system can be accepted or not by the user, grade scale to see the (grade) of the application or in other words the category that shows the quality of usability of a system and adjective rating to see the rating of the software produced or in other words the category that shows the user’s perception of the usability of a system. There were three levels of acceptability, not acceptable, marginal (low and high) and acceptable. In the meantime, the grades were A, B, C, D, and F on the grade scale. There were numerous levels for the adjective rating, including worst imaginable, poor, ok, good, excellent, and best imaginable. (Paolis, Gatto, Corchia and Luca, 2023). The questionnaire follows the standard format of SUS has
shown in Table 2 (Brooke, 1995).

Table 2. Instrument SUS
(Source: Personal documentation)

<table>
<thead>
<tr>
<th>No.</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I think that I would like to use this system frequently</td>
</tr>
<tr>
<td>2</td>
<td>I found the system unnecessarily complex</td>
</tr>
<tr>
<td>3</td>
<td>I thought the system was easy to use</td>
</tr>
<tr>
<td>4</td>
<td>I think that I would need the support of a technical person to be able to use this system</td>
</tr>
<tr>
<td>5</td>
<td>I found the various functions in this system were well integrated</td>
</tr>
<tr>
<td>6</td>
<td>I thought there was too much inconsistency in this system</td>
</tr>
<tr>
<td>7</td>
<td>I would imagine that most people would learn to use this system very quickly</td>
</tr>
<tr>
<td>8</td>
<td>I found the system very cumbersome to use</td>
</tr>
<tr>
<td>9</td>
<td>I felt very confident using the system</td>
</tr>
<tr>
<td>10</td>
<td>I needed to learn a lot of things before I could get going with this system</td>
</tr>
</tbody>
</table>

According to Schepp (2023) the UEQ contains 6 scales with 26 items:

1. Attractiveness: whether or not the user finds it to be interesting to use.
2. Perspicuity: which refers to how easy a system is for users to use on their first try.
3. Efficiency: when the user completed tasks without additional effort.
4. Dependability: which refers to the user’s sense of control over the system.
5. Stimulation: Is using the product exciting and motivating?
6. Novelty: Do customers find the product’s inventiveness and creativity appealing?

Every question in the SUS analysis has a scale value based on a 5-point Likert scale, with 1 indicating “strongly disagree” and 5 indicating “strongly agree.” SUS produces a single number that represents the overall usability of the system being studied, which in this case is the AR Try-On feature from Tokopedia. Add up all the values of each item before calculating the SUS score. The value range for each item is 0 to 4. The value of odd-numbered items (i.e., 1, 3, 5, 7, and 9) = scale position - 1, and the value of even-numbered items (i.e., 2, 4, 6, 8, and 10) = 5 - scale position. After obtaining the total value, then the total = (even item value + odd item value) * 2.5. The final step in calculating the SUS score is to find the average of the total item values of all respondents’ answers.

With UEQ analysis, each question has a question value, based on a 7-point Likert scale, from 1 which indicates ‘negative’ to 7 which indicates ‘positive’. These seven scales are used to reduce bias from well-known products. UEQ has item values ranging from -3 to 3, with -3 indicating the most negative answer, 0 indicating a neutral answer, and +3 indicating a very positive answer. The UEQ calculation begins by finding the mean, variance and standard deviation of each respondent’s answer, which is then analyzed based on the categories of attractiveness, efficiency, perspicuity, dependability, stimulation and novelty, and finally comparing the impression values (mean and variance) in groups with the evaluation values.

Result

In the qualitative method, according to the question in Table 1, on general background, the author wants to find out about the habits of the respondents when they are going to buy a makeup product offline, because this will affect their interaction habits in AR. When AR try-on pro-
vides an experience that is similar to the offline shopping experience, then the respondents will get more engagement with this technology. From the respondents’ answers: Before the COVID-19 pandemic, respondents usually bought makeup by coming to an offline store and trying on the makeup they would choose on the back of their hand, lips or face to make sure the product they would buy was not allergic on their skin, had a texture that matched their preferences, and did not make a mistake in choosing a colour that matched their skin tone. In trying out these products, respondents often tried several products at once, at least 2-3 products.

From the question of General experience with offline and online shopping, the author wants to know the concerns of the respondents when buying makeup products online, in order to find out whether this AR try-on feature can solve the respondents’ problems. From the respondents’ answers: There are some negative experiences that have been experienced by the respondents when shopping online, 3 of them have experienced buying makeup with a colour that turned out to be different from their expectations, as a result some of them ended up giving it to their friends, some still wore it because of the expensive price factor, some left it at home and did not use it, and bought back a color that they thought ‘hopefully’ matched their skin. Therefore, to avoid regrettable decisions, often participants prefer to buy expensive cosmetics in offline store, whereas they claim “nothing to lose”, when buying inexpensive cosmetics as they can easily replace them with more suitable ones without feeling any loss.

In Figure 2, the stages of the scenario that users need to go through for usability testing are depicted. The purpose of usability testing is to assess the users’ understanding of the user flow designed by the Tokopedia application, specifically regarding the guidance provided to enter the AR virtual try-on feature. The author aims to identify any difficulties encountered by users during the user flow. Additionally, the author seeks to examine the user interactions within the AR virtual try-on, including the ease of understanding the user interface, the placement of buttons and interface elements, and the AR interactions such as swiping and clicking, in relation to the user experience. Based on the results of the usability testing, the author found that one participant encountered difficulties initially while attempting to enter the AR virtual try-on and had to uninstall the Tokopedia application before reinstalling it. Regarding the user flow, the participants did not encounter any difficulties from the initial entry into the Tokopedia application until accessing the AR virtual try-on, as they were already familiar with the user experience of Tokopedia, including searching for the Maybelline brand, finding the requested product, and entering the AR virtual try-on feature. As for the user interface in the AR virtual try-on, participants were slightly confused by the new icons they encountered for the first time, as shown in Figure 3.
For the questions Experience with AR as virtual Try-on tool, AR virtual technology try-on trial and AR virtual try-on experience, the author wants to know whether the categories that are indicators in this study, such as: acceptability ranges, grade scale and adjective ratings as well as attractiveness, efficiency, perspicuity, dependability, stimulation and novelty are in accordance with the quantitative data results. From the respondents’ answers:

1. For this AR technology, all respondents have tried it before, but for ‘fun’ like what is on social media in the form of AR filters, but they have never used it for shopping for makeup. (novelty)

2. AR should have been utilized not only for fun, but also for its function. And maybe later it can reach other products not just makeup. (acceptability range, grade scale)

3. When trying it out, respondents felt enjoyed, could choose from the colours available, and could also experiment with new colours that they had never tried before. In a short time, respondents were able to do the scenario prepared by the author. Respondents rated it as very easy to use, the icons provided were also informative, and did not make them confused. There was a high sense of enthusiasm after trying this AR try-on feature by asking “Where else is this AR?” so that respondents can try other products again. (attractiveness, perspicuity, efficiency, dependability)

4. But because all five respondents are not consumers of the Maybelline Brand, respondents do not know whether the colours seen on the screen match their original colours. If the quality of the colours seen on the screen is the same as their original ones, respondents would be happy to repeat orders by buying new colours and even recommending them to their close people. (stimulation, adjective rating).

Discussion

In the quantitative method, 23 data were obtained from respondents. The test results of the System Usability Scale from the interpretation of the respondents are shown in Table 3.

Table 3. Test result of SUS
(Source: Personal documentation)
The SUS score of 23 respondents was 77.61. According to Figure 4, these results have a level of acceptability ranges in the “acceptable” category. Acceptable with a SUS value above 70, it means that the system has a good level of usability and can be accepted by users. At the grade scale level it is at “B+” with a SUS value between 77.2 to 78.8, meaning that the system has good usability quality. Whereas at the adjective rating level it is in the “Good” category, where the SUS score is between 71 to less than or equal to 85, meaning that users feel that the try-on AR feature is easy to use and satisfying. In this context, the user interface (UI) elements, such as buttons, sliders, or menus, play a significant role in facilitating intuitive navigation and interaction within the AR Virtual Try-On feature. The seamless integration with the Tokopedia e-commerce platform enables users to effortlessly transition from trying out products to making purchases.

Table 4. Test result of UEQ
(Source: Personal documentation)

<table>
<thead>
<tr>
<th>UEQ Scales (Mean and Variance)</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attractiveness</td>
<td>1.957</td>
<td>0.57</td>
</tr>
<tr>
<td>Perspicuity</td>
<td>2.043</td>
<td>1.19</td>
</tr>
<tr>
<td>Efficiency</td>
<td>2.043</td>
<td>0.73</td>
</tr>
<tr>
<td>Dependability</td>
<td>1.957</td>
<td>0.94</td>
</tr>
<tr>
<td>Stimulation</td>
<td>2.250</td>
<td>0.60</td>
</tr>
<tr>
<td>Novelty</td>
<td>2.207</td>
<td>0.59</td>
</tr>
</tbody>
</table>

The results of the UEQ test obtained from the results of 23 respondents were analyzed using mean, variance, and std deviation is shown in Table 4. From the results of the calculation of the mean and variance, it shows that all categories are positive, with the green arrow indicator pointing up. These observations indicate the effectiveness and reliability of the AR Virtual Try-On feature. Real-time rendering and precise tracking ensure that virtual product trials accurately align with the user’s facial features. Nevertheless, there are occasional technical challenges, such as minor lag or disruptions, that may impact usability and the overall user experience.

Obtained an impression value greater than 0.8 (value > 0.8 represents a positive evaluation and value <0.8 represents a negative evaluation) with the lowest score 1.8 in the Dependability category with the understanding that there are still some respondents who are confused in their interactions and the highest is 2.4 in the category Novelty stated that this AR feature apart from being just entertainment, this technology has a new use, namely for shopping. In order to enhance user-friendliness, Tokopedia can implement refinements and enhancements to the arrangement of buttons, interface components, and instructional materials within the feature. By making appropriate changes, users will have a better understanding and seamless experience when using the feature.

Table 5. Test result of UEQ by categories
(Source: Personal documentation)
According to Table 5, it can be seen that attractiveness and dependability have the lowest score of 1,957 with the understanding that, there is still confusion in using the interaction felt by respondents (dependability) has an impact on the Overall impression of the product (attractiveness) in user experience. While the highest score is at stimulation of 2,250 with the understanding that respondents are excited and motivated to use the product, there is a possibility because of the novelty of the AR function in shopping.

Cronbach alpha values range from 0 to 1. The higher the Cronbach alpha value, the more reliable the questionnaire is made. Interpret the Cronbach alpha results according to the following criteria: low reliability is defined as alpha < 0.50, moderate reliability as 0.50 < alpha < 0.70, sufficient reliability as alpha > 0.70, strong reliability as alpha > 0.80, and perfect reliability as alpha > 0.90. It can be seen in Table 6, Cronbach alpha obtained > 0.8 and > 0.9. Strong and perfect reliability at Cronbach alpha > 0.9 means that the research instrument used has a very high level of consistency or stability. That is, the respondent’s answers to the statements in the questionnaire did not change from time to time. Cronbach alpha value > 0.9 indicates that all items in the questionnaire are positively correlated and there are no unreliable items.

Conclusion

From the results of SUS and UEQ, as well as the results of the interviews, it was found that all categories of SUS, namely acceptability range, grade scale, and adjectives and categories from UEQ, namely attraction, efficiency, perspicuity, dependability, stimulation and novelty of AR features on Tokopedia, had positive values and have good reception. Due to the novelty element of the AR function, the positive attractiveness and novelty values encourage and stimulate users to use AR Try-on for purchasing purposes other than entertainment. We can therefore conclude that consumers find the AR Try-on function an enjoyable shopping experience. Although in reality, e-commerce consumers come from diverse user bases with varying brand preferences and habits. This research will have a deeper grasp of how various user can interact with this AR Try-on functionality by incorporating people from diverse back-

Table 6. Test result of UEQ by categories
(Source: Personal documentation)
grounds.

However, Tokopedia can optimize the button arrangement, interface elements, and instructions in the feature to promote clarity and ease of use, as some respondents reported misunderstanding or trouble using the AR try-on option. This will improve dependability and user experience. It’s also critical to give consumers who might run into problems clear support or direction.

The factors that influence the assessment are that the respondents who conducted the questionnaire were not brand loyalists from Maybelline, so there was a slight bias in the results.

As the conclusion, the participants demonstrated keen interest in utilizing the AR virtual try-on feature and recognized its practicality in choosing and virtually testing makeup products. However, as non-users of Brand Maybelline, they had limited awareness of how accurately the color representation in the AR virtual try-on matched the original colours. Hence, it is essential to ensure the quality and precision of the colour representation aligns with the real products. This will foster trust and comfort among users when making purchase decisions.

Based on the existing limitations, the authors hope that in the future it will be even better in finding suitable respondents. And there is a further analysis of the comparative interaction of AR in various media.

References


