Decision Fatigue in Digital Catalogues: The Effect of Inventory Size on Decision-Making Efficiency

Iman Syed, Samarth Kulkarni, David Yusuf
Dept. of Electrical Engineering and Computer Science
York University
Toronto, Ontario, Canada M3J 1P3
imansyed@my.yorku.ca, samrk@my.yorku.ca, davidy03@my.yorku.ca

ABSTRACT

This study examined how product selection size impacts consumer decision-making efficiency and task difficulty in online shopping environments. Participants completed six tasks involving clothing inventories of 10, 25, and 50 items. For each task, they were given a list of requirements and asked to find the matching item. Performance was measured by time taken, number of mouse clicks, incorrect submissions, and pages visited. User experience was assessed through post-task Likert-scale questionnaires. Results showed that larger inventory sizes led to longer task times, more pages visited, and increased task-completion difficulty. In contrast, mouse clicks and incorrect submissions remained relatively consistent across inventory sizes. These findings suggest that decision fatigue increases with inventory size, negatively affecting user efficiency and experience. The study highlights the need to balance variety with usability in digital shopping interfaces, as an overwhelming number of choices may affect user ability to efficiently shop.

Keywords

Decision fatigue, choice overload, online shopping, e-commerce interfaces, cognitive load, user experience, consumer behavior, product variety, Decision-making efficiency.

INTRODUCTION

In the digital age, online shopping has become an integral part of modern consumer behavior, offering unparalleled convenience and access to a vast variety of products. While this abundance of options is often seen as beneficial, studies by Iyengar & Lepper [3] and Scheibehenne et al. [5] show that too many choices can hinder decision-making, which is referred to as *choice overload* or *decision fatigue*. When users navigate interfaces that present hundreds to thousands of similar items, the increased cognitive effort can lead to delayed decisions, reduced satisfaction, or even decision avoidance.

This study explores how the number of available product options in an online shopping interface affects users' decision-making experience. Specifically, it investigates user performance and difficulty when interacting with three different inventory sizes: 10, 25, and 50. These range of inventories reflect search results of real-world websites, looking at how easily users find their desired item with the given range.

Past research by Iyengar and Lepper [3] and Scheibehenne et al. [5] has examined the psychological consequences of extensive choice, while Gwizdka [2] explored the effects of cognitive load in web-based environments. More recently, Wang et al. [6] investigated interface design elements that contribute to decision fatigue. However, few studies have directly examined how inventory size influences user decision-making in structured, task-based selection interfaces.

This user study examines both quantitative measures (e.g., time to decision) and qualitative feedback (e.g., user difficulty) as participants complete guided decision tasks on three inventory sizes. The goal is to understand how increase in choice complexity impacts user experience and difficulty, and to provide insights for future e-commerce interface design that can help reduce decision fatigue.

RELATED WORK

Consumer research and cognitive psychology have made decision fatigue and choice overload a significant concern, especially when it comes to online shopping. Below is a review of key studies summarizing relevant research.

A study by Wang et al. [6] investigated choice fatigue in online shopping, and found that as the amount of options increased on the websites, the consumers displayed greater cognitive load and also a decrease in their decision making efficiency. Their results indicated how choice fatigue can lead customers to make quick decisions that are usually less than ideal. This offers clear evidence of how overwhelming the variety in online retail environments may reduce the quality of choices.

In their 2007 study, Fasolo et al. [1] explored the effects of reducing the number of attributes displayed during decision tasks. The results showed that, particularly in cases where attributes are redundant or when the customer lacks significant prior preferences, reducing the complexity may help decision making without sacrificing satisfaction. According to this study, simplified layouts could decrease the amount of options available in both online and physical shopping environments.

Pignatiello et al. [4] provided a conceptual analysis of decision fatigue, where they reviewed how making the same decisions over and over again might reduce one's ability for self control and reduce the quality of their decisions. Although the review mainly focuses on healthcare and daily life decisions, it offers a valuable theoretical foundation for understanding how decision fatigue can show up in some online shopping contexts, where consumers must make a lot of small judgements in quick succession.

In order to evaluate the reliability of the choice overload phenomenon, Scheibehenne et al. [5] conducted a meta-analysis of 63 experimental conditions from 50 studies. Though the overall effect size was virtually zero, the authors found a number of moderators (such as the decision-makers' expertise, the nature of the task, and prior preferences) that can influence whether choice overload will have a negative impact on motivation or satisfaction. Their analysis emphasizes the necessity for deliberate online shopping interface design by highlighting the fact that choice overload is context-dependent rather than universal.

METHOD

This study followed an experimental methodology, observing participant behaviours and responses under carefully controlled and manipulated tasks and conditions.

Participants

72 participants were gathered for this experiment. As the study involves online shopping, the target population consisted of individuals who were comfortable using computers. Participants who were comfortable navigating websites and had experience online shopping previously became the most ideal match. As a result, university students matched this demographic extremely well and were selected as the participant pool. Outside of a criteria match, university students were easily accessible, and with a research team consisting of university students, strong connections and access to university facilities to advertise were already granted.

Neither age nor gender was treated as a key variable in the study. The selected inventory sizes showed images suited for all gender identities, and tasks were designed to avoid any gender-based complications. Instead, the primary variable of participant eligibility is their technological experience. Participants chosen were aware of how to navigate their device controls and were confident in effectively using a variety of different websites. This ensured that the observed behaviour reflects cognitive decision-making involving choice and preference and not basic computer operations.

Participants were recruited by posting ads online as well as using connections.

Interested participants enrolled in the experiment by clicking a provided link shown on the advertising material. This took them to a pre-experiment survey where information about their email, age, gender (while stated as not a key variable, balancing is still important), computer model, access to a laptop or desktop, and digital experience was asked. It also included a consent form that the participant could download, sign, and upload. This form outlined the nature of the study and addressed all ethical requirements.

Due to the existing connections that made up the participant amount, this study was voluntary.

Apparatus

The software and hardware for this experiment were minimal. Each participant needed access to a desktop or laptop and the Zoom application.

As shown in Figure 1, during the Zoom session, a group member was on the call with the participant. The participants shared their incognito screen, and the group member recorded the participant's screen activity from their own device. Group members made sure to avoid recording the participant's menu and task bar for privacy reasons.

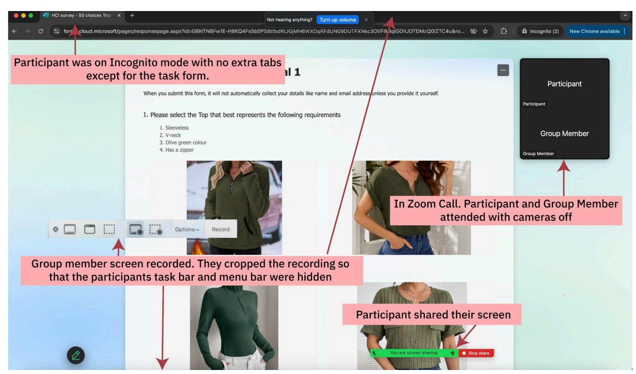

Figure 1. Screenshot of example Zoom meeting and screen recording layout and format.

After the experiment was completed, the research team reviewed each screen recording. The number of mouse clicks, the number of pages visited, the task completion times, and incorrect submissions were all compiled and calculated manually. These metrics and insights were then inserted into an Excel spreadsheet alongside the participant's ID number and their corresponding questionnaire responses.

Procedure

Participants were asked to join a Zoom meeting. They were given instructions on what they were required to do, were told where the required information was found, were told that if the blue button at the bottom of the page said

'Next,' then there were more options available on the next page, and were informed that if they did submit the wrong item, they would be told so by the researcher and would have to do the task again. Once informed, they were given one practice trial with 25 items (median inventory size), allowing them to get an idea of how to navigate between pages as well as understand where the item's requirements were located.

Once done with the practice trial, participants were given the first link to the tested form (the link was determined by their counterbalancing group) and were asked to read the requirements and submit the correct item.

After each inventory size was complete (ex., Task 1 and 2's forms for size 10 were submitted correctly), the participant was asked to complete a six-question Likert-scale questionnaire that relayed their experience while doing the tasks. Once done, the next set of tasks for the next inventory size was completed as well as another questionnaire with the same questions. Finally, the last inventory size's tasks and questionnaire were completed in the same format.

Tasks for each inventory size varied slightly to reduce learning effects. Specifically, task 1 were tops and task 2 were bottoms. The tasks and their requirements were:

- Size 10, Task 1: Half Sleeve, Has imagery of cats and plants on it, Sleeves are folded
- Size 10, Task 2: Light blue sweatpants, Star graphic on right thigh, White stripes on the side, Flared bottom pants
- Size 25, Task 1: Half sleeve, Number '33' on it, Two stripes on sleeves, Green colour
- Size 25, Task 2: Dark blue shorts, Two squiggly stripes on side, Cotton
- Size 50, Task 1: Sleeveless, V-neck, Olive green colour, Has a zipper
- Size 50: Task 2: Medium-wash skinny jeans with a snug fit, Women's jeans, Large cargo pocket on thigh, No visible rips or distressing.

Design

The study was a 3 x 2 within-subject design, with 3 levels for the independent variable of inventory size (10, 25, 50) and 2 tasks (task 1 = tops, task 2 = bottoms). Counterbalancing was also included. 12 groups were created by gathering n! of the 3 inventory sizes (6 groups), and each group was then divided into 2. One group would start with task 1 and then task 2, while the other would start with task 2, then task 1.

The dependent variables of the study included Time(s), Pages Visited, Incorrect Submissions, Mouse Clicks, and Difficulty.

Time(s) was the time taken between the user clicking 'Start Now' and when they clicked 'Submit.' If a participant submitted an incorrect image, they were required to repeat the task until a correct submission was made. In such cases, the time for each attempt was recorded and summed to calculate the total time taken for that task.

Pages Visited was the number of times the participant loaded another page. By default, the page visited counter was 1, and the total number of visits depended on how many times 'Back' or 'Next' was clicked.

Incorrect Submissions was the number of times the participant submitted the wrong image.

Mouse Clicks was the number of times the participant selected an image up until they clicked 'Submit.' Mouse clicks did not include the participant clicking 'Start Now,' 'Next,' 'Back,' or 'Submit.'

Difficulty was the participant's feedback and experience completing the task. They rated their experiences on a scale of 'Strongly Disagree,' 'Disagree,' 'Neither Agree nor Disagree,' 'Agree,' and 'Strongly Agree' (1-5 scale). Participants answered 6 questions:

1. I had trouble deciding which item to choose.
2. I felt mentally tired after completing these tasks.
3. I felt frustrated during the tasks.
4. I felt overwhelmed by the number of items shown.
5. I had difficulty staying focused during the task of finding the right item.
6. I felt impatient to finish the task.

The experiment took a maximum of 15 minutes to complete between the introduction and the final questionnaire. In total, 72 practice trials, 432 measured trials, and 216 questionnaire results were gathered within a week, resulting in about 1 ½ hours of work being done per researcher every day (varying depending on the number of participants available).

RESULTS AND DISCUSSION

Statistically significant differences were found across inventory sizes for time taken, number of pages visited, and questionnaire responses (difficulty). In contrast, the number of mouse clicks and incorrect submissions were not statistically significant and may have been caused by random effects.

To assess potential order effects, each dependent variable was analyzed using the Kruskal-Wallis test to evaluate the effectiveness of the counterbalancing design. The results indicated that counterbalancing successfully offset order effects for Time(s) ($\chi^2 = 0.658$, $p > .05$, df = 11), Pages Visited ($\chi^2 = 0.681$, $p > .05$, df = 11), Mouse Clicks ($\chi^2 = 3.396$, $p > .05$, df = 11), and Incorrect Submissions ($\chi^2 = 5.239$, $p > .05$, df = 11). However, the Difficulty variable, found through questionnaire responses, showed a significant effect of task order, suggesting that counterbalancing did not fully eliminate order effects in this case ($\chi^2 = 26.50$, $p < .01$, df = 11).

Time(s)

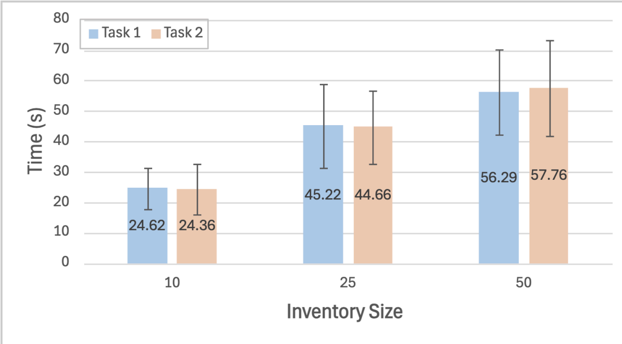

Figure 2. Time(s) by Inventory Size. Error bars show ±1 *SD*.

As shown in Figure 2, the mean selection time increased consistently with inventory size. Task 1 results showed a mean of 24.62 seconds for the 10-item catalogue, 45.22 seconds for the 25-item catalogue, and 56.29 seconds for the 50-item catalogue. Task 2 results showed a mean of 24.36 seconds for the 10-item catalogue, 45.66 seconds for the 25-item catalogue, and 57.76 seconds for the 50-item catalogue. These results suggested that larger inventories led to longer decision times, even when the target item appeared early in the list. Observational data support this pattern, as participants often chose to browse through all available options to ensure no better alternatives were found on later pages.

The combined-task average selection time for the 25-item inventory (24.46 seconds) was 84.61% higher than that of the 10-item inventory (45.21 seconds). The 50-item inventory (56.27 seconds) showed a 129.77% increase compared to the 10-item inventory and a 24.46% increase compared to the 25-item inventory. These findings indicate that even moderate increases in inventory size significantly affected decision-making time.

Normality tests were conducted for all six size/task categories (size 10: task 1, size 10: task, size 25: task 1, size 25: task 2, size 50: task 1, size 50: task 2), with only three meeting the criteria for normal distribution. As a result, Friedman's test was used to compare inventory sizes. Two separate tests were conducted: one for task 1 and one for task 2. For task 1, the results revealed a statistically significant difference between inventory sizes ($\chi^2 = 85.36$, $p < .0001$, df = 2). Post hoc analysis confirmed significant differences across all three pairwise comparisons. Similarly, task 2 showed a significant difference between inventory sizes ($\chi^2 = 95.36$, $p < .0001$, df = 2), with post hoc tests again confirming significance in all pairings.

Pages Visited

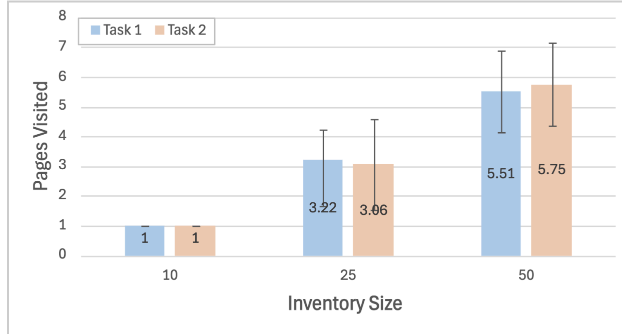

Figure 3. Pages Visited by Inventory Size. Error bars show ±1 *SD*.

As seen in Figure 3, the average number of pages visited for both tasks of inventory size 10 was one page. This is due to there only being a single page of options for the participant to browse through. For the larger inventory sizes, multiple pages were accessible, resulting in more pages being visited during the participants' search. Specifically, for an inventory size of 25, task 1 had an average of 3.22 pages visited, and task 2 had an average of 3.06 pages visited. The combined mean of both tasks for size 25 (3.14 pages visited) was found to have a 214% increase from the 10 inventory size. For size 50, task 1 had a mean of 5.51, and task 2 had a mean of 5.75 pages visited. The total mean of size 50 (5.63 pages) showed a 78.3% increase over size 25.

Normality of data was assessed using the Lilliefors test across all size combinations of inventory sizes and tasks. Since three out of six combinations were not normally distributed, the non-parametric test Friedman was used. Two tests were conducted for each task to evaluate the differences in the number of pages visited across inventory sizes. Results showed that the difference was statistically significant for task 1 ($\chi^2 = 130.979$, $p < .0001$, df = 2) and task 2 ($\chi^2 = 128.140$, $p < .0001$, df = 2). Post hoc analysis for both tasks further confirmed that the difference was significant for all pairwise comparisons.

Mouse Clicks

The mean number of clicks remained relatively consistent across inventory sizes. For task 1, participants averaged 1.22 clicks for the 10-item catalogue, 1.17 clicks for the 25-item catalogue, and 1.24 clicks for the 50-item catalogue. For task 2, participants averaged 1.18 clicks for the 10-item catalogue, 1.14 clicks for the 25-item catalogue, and 1.17 clicks for the 50-item catalogue. The total average number of clicks across both tasks was 1.20 for the 10-item catalogue, 1.15 for the 25-item catalogue, and 1.20 for the 50-item catalogue. These results suggested that inventory size had minimal impact on the number of clicks required to complete the tasks.

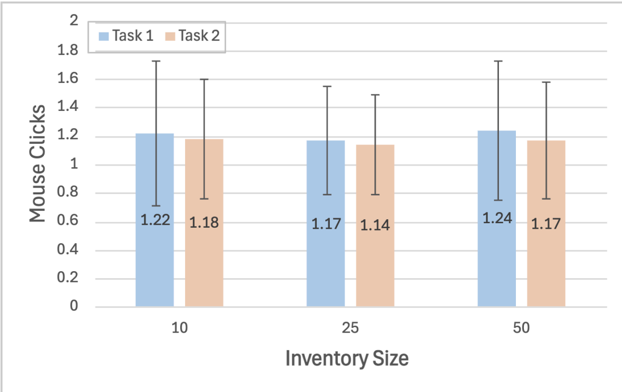

Figure 4. Mouse Clicks by Inventory Size. Error bars show ±1 *SD*.

Figure 4 shows that the average number of clicks for both tasks remained relatively stable across inventory sizes (10, 25, and 50 items). While task 1 consistently required slightly more clicks than task 2 for all inventory sizes, the overall differences were minimal, reflecting the non-significant statistical results from the Friedman test.

Normality tests for all six click categories indicated that the data was not normally distributed. Therefore, Friedman's test was used to compare inventory sizes for each task. For task 1, no statistically significant difference was found between inventory sizes ($\chi^2 = 0.35$, $p > .05$, df = 2), with post hoc pairwise comparisons confirming no significant differences. Task 2 also showed no statistically significant differences ($\chi^2 = 0.15$, $p > .05$, df = 2), again confirmed by post hoc results.

Incorrect Submissions

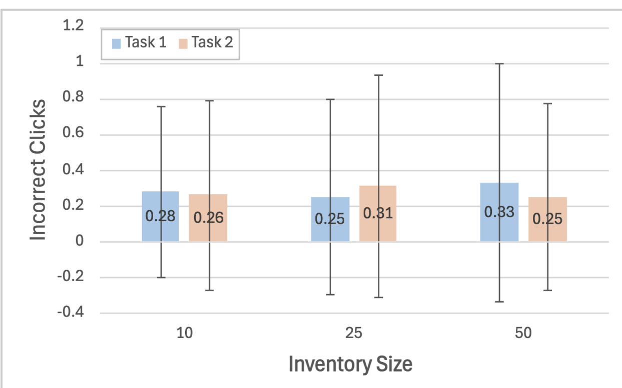

Figure 5. Incorrect Submissions by Inventory Size. Error bars show ±1 *SD*.

The average number of incorrect submissions was low across all inventory sizes, with minimal variation between conditions. For Task 1, the 10-item inventory had an average of 0.28 incorrect submissions, the 25-item inventory averaged 0.25, and the 50-item inventory averaged 0.33. Task 2 showed a similar pattern, with averages of 0.26, 0.31, and 0.25, respectively. As shown in Figure 5, no clear trend emerged between inventory size and incorrect submissions, and error rates remained consistently low.

Normality tests indicated that the data were not normally distributed; therefore, Friedman's test was used to compare conditions. For Task 1, the analysis showed no statistically significant difference between inventory sizes ($\chi^2 = 0.33$, $p > .05$, df = 2), with post hoc pairwise comparisons confirming no significant differences. Similarly, Task 2 also showed no significant differences ($\chi^2 = 0.02$, $p > .05$, df = 2), with post hoc tests again confirming no meaningful variation across inventory sizes.

These findings suggest that inventory size did not have a measurable impact on the number of incorrect submissions, in contrast to the clear effect observed for selection time.

Questionnaire Results

A Likert-scale questionnaire was conducted after each inventory size was complete. Participants were asked a series of six questions varying from 'Strongly Disagree,' 'Disagree,' 'Neither Agree nor Disagree,' 'Agree,' and 'Strongly Agree.' The answers to the six questions were then converted into a 1-5 rating scale. Responses indicated the difficulty in the task. Responses closer to 'Strongly Agree' meant that the user struggled to complete the task with ease.

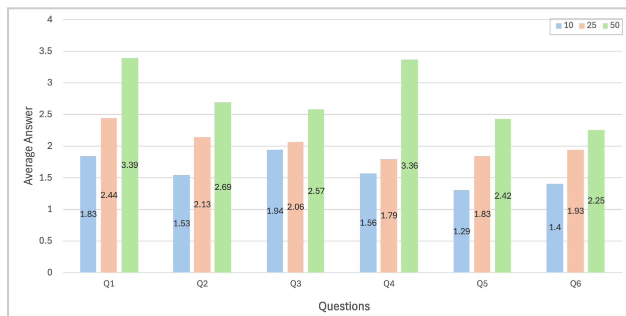

Figure 6. Average Answer by Questions for three inventory sizes (10, 25, 50).

Figure 6 showed the results of all inventory sizes for the six questions. The mean of all the questions for size 10 was calculated to be 1.59. In other words, Most participants found that they strongly disagreed on the difficulty of the task. From this, it was concluded that participants found that the smaller catalogue size was easily navigable and were able to find their required item with ease.

The mean for size 25 was 2.03, which showed that participants disagreed with the majority of the questions. Specifically, participants were found to not be overwhelmed by the increase in options (mean of 1.79), but still found difficulty in selecting the correct item (mean of 2.44).

Inventory size 50's questionnaire results showed the total mean of all answers was 2.78, showing that participants began leaning more towards the uncertainty of whether the task was difficult or not. For the individual question means, it is found that participants began to feel overwhelmed by the increase in catalogue size (mean of 3.36) and also struggled slightly with the task of finding the correct item (mean of 3.39).

Although the mean responses across inventory sizes suggested a relationship between inventory size and questionnaire answers, a Wilcoxon Signed-Ranks test was conducted to statistically confirm this relationship. The goal was to determine whether changes in inventory size led to significant differences in participants' responses.

Three pairwise comparisons were evaluated: 10 vs. 25, 10 vs. 50, and 25 vs. 50. 10 vs. 25 revealed a significant difference (z = -6.75, df = $p < .0001$), suggesting that even a moderate increase in inventory size influenced participant responses. Similarly, the 25 vs. 50 comparison also demonstrated a strong significant effect (z = -9.85, df = $p < .0001$), following the same conclusion that even small changes in inventory size show changes in answers. Finally, the 10 vs. 50 comparison also showed a strong significant effect (z = -13.33, df = $p < .0001$), confirming that even between larger inventory sizes, differences are present. These results collectively indicated that as inventory size increases, it has a consistent and significant impact on user responses.

CONCLUSION

This study investigated the impact of product selection size on consumer decision-making efficiency and its application in online shopping environments. The results demonstrated that as inventory sizes increased (10, 25, 50 items), participants took significantly longer to make decisions, visited more pages, and reported greater difficulty and mental fatigue in selecting their desired item. In contrast, other measures such as mouse clicks and incorrect submissions remained fairly consistent across inventory sizes, suggesting that these differences may be due to random variation.

These findings confirmed that decision fatigue is evident both in participant experiences and in quantitative performance data as inventory size increases. While this study was limited to a maximum of 50 items, the consistent upward trend in task time and difficulty suggests that larger inventories can significantly affect the online shopping experience. This implied that online platforms must find a careful balance between offering variety and avoiding user decision fatigue. Presenting too many options may discourage purchases by overwhelming users, even if it increases browsing time.

Future work could investigate the issue with counterbalancing with the questionnaire results as well as explore more in-depth methods to increase external validity. Specifically, future works may investigate the role of filtering tools while making decisions or will look more into the cognitive process of participants through think-aloud techniques.

REFERENCES

1. Fasolo, B., McClelland, G. H., & Todd, P. M., Escaping the tyranny of choice: When fewer attributes make choice easier. *Marketing Theory*, 7(1) (2007), 13–26.
2. Gwizdka, J., Distribution of cognitive load in web search. *Journal of the American Society for Information Science and Technology*, 61(11) (2010), 2167-2187.
3. Iyengar, S. S., & Lepper, M. R., When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6) (2000), 995–1006.
4. Pignatiello, G. A., Martin, R. J., & Hickman, R. L., Decision fatigue: A conceptual analysis. *Journal of Health Psychology*, 25(1) (2018), 123–135.
5. Scheibehenne, B., Greifeneder, R., & Todd, P. M., Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, 37(3) (2010), 409–425.
6. Wang, Y., Mo, D. Y., & Ho, G. T .S., How choice fatigue affects consumer decision-making in online shopping. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, (IEEE 2023), 1–6.