

Consistency in evaluation and structure of purchasing decision advice by LLMs

Smaranda Giușcă¹, Irina Maiorescu², Gabriel-Cristian Sabou³ and Dragomirescu (Radu) Alexandra Teodora⁴

¹⁾²⁾³⁾⁴⁾ *The Bucharest University of Economic Studies, Bucharest, Romania.*

E-mail: smaranda.giusca@com.ase.ro; E-mail: irina.maiorescu@com.ase.ro;

E-mail: gabriel.sabou@csie.ase.ro; E-mail: alexandra.teodora.at@gmail.com

Please cite this paper as:

Giușcă, S., Maiorescu, I., Sabou, G.-C., & Dragomirescu (Radu), A.T. (2026). Consistency in evaluation and structure of purchasing decision advice by LLMs. In C. Vasiliu, D.-C. Dabița, A. Tziner, D. Pleșea, & V. Dinu (Eds.), *Proceedings of BASIQ International Conference on New Trends in Sustainable Business and Consumption, 12th Edition* (pp. 388-395). Editura ASE.

DOI 10.24818/BASIQ/2026/12/087

Abstract

As people turn to the use of Large Language Models (LLMs) for help in making various decisions in their life, purchasing advice is one of the areas in which LLMs should consistently turn out reliable answers. This empirical research shows how a query for help in a purchasing decision differs in the case of three LLMs. The technical nature of the query, a result of appliances being the product chosen for the experiment, was a way to ensure an equitable comparison was possible. An experiment was conducted in which ChatGPT, Mistral and Deepseek received prompts that requested a product comparison and decision regarding two sets of similar household appliances. Different structures of the answers were observed with different deviations from the user's prompt, in what concerns the order, amount and specificity of the information provided by the LLMs. In order to better judge the similarities of the product models comparisons that the three LLMs developed, a subsequent prompt requesting for evaluation through a numerical score was introduced. Although the product models were graded for three technical and quantifiable characteristics the resulted numerical score showed the three LLMs did not grade the models' characteristics similarly. Reaching a common decision for the three LLMs seems to be more of a coincidence than a predictable result. Using LLMs in making a purchasing decision is faster and might use more sources of information but it does not necessarily prove to be better than the user making a decision on his own.

Keywords:

purchasing decision, LLM, product comparison, online shopping, AI tools

DOI: 10.24818/BASIQ/2026/12/087

Introduction

The use of AI in the shape of Large Language Models (LLMs) by individual users for everyday purposes has had a considerable increase in recent years, ChatGPT users alone amounting to 10% of the globe's population in July 2025, less than three years since its release to the public (Chatterji et al., 2025).

Shopping, both online or physical, is one of those instances where AI is used. Other, more traditional, digital purchase recommendation sources like social media, online marketplace, search engines or brand website or app are already losing ground in the preferences of gen AI users, as 18% of them choose gen AI as their purchase recommendation source, only slightly surpassed by 19% who choose physical store recommendations (Wright et al., 2025).

Reliability of AI outputs in situations where it is used by non-professionals has been questioned and tested in various areas of interest so it is necessary to also check the methods employed in creating the answers to prompts of socioeconomic nature, such as requiring help in the purchasing decision. With this reasoning, for an object as technical as an appliance whose characteristics are accurately defined, quantifiable and can all be compared, the results of using technology in the purchase decision should be precise and unique. In the case of an entirely human process of purchasing decision, the results could vary

by error or due to subjectivity, either during the comparison or the decision phase. For example, there are product characteristics whose assessment are prone to subjectivity (such is the case of design features that rely significantly on aesthetics). If we take into consideration the diversity of products and the amount of information or details available nowadays it is no wonder that 74% of consumers give up shopping for the desired product and report feeling overwhelmed (Crowley et al., 2024).

The fact that consumers turn to reviews in making their purchase decision also complicates things, paying more attention to negative comments than to positive ones, with women more likely to be influenced by negative reviews than men (Chen et al., 2022). While reviews and ratings from consumers are useful, the unstructured nature of textual user reviews makes them difficult and time-consuming to use (Lackermair et al., 2013).

This paper is structured in four parts. The first section covers the relevant scientific literature regarding the process of making a purchasing decision nowadays and the uses of AI in this process. The second section presents the research methodology and the scenario built for the experiment. The third section covers the results and discussions, while the final section presents the conclusions and recommendations for future research.

1. Review of the scientific literature

The development of digital commerce and artificial intelligence has clearly changed how consumers approach online purchasing. Instead of relying only on traditional information sources, users now interact with intelligent systems that help them search, compare and evaluate products. In this context, generative AI and conversational agents have become increasingly integrated into e-commerce platforms, where they play an active role in shaping consumer decisions and simplifying the overall buying process (Anggraini et al., 2026). Other ways AI has modified digital commerce was by use of incorporated product evaluation algorithms, individualized purchasing recommendations, and improved customer care (Dixit et al., 2025).

What is particularly interesting in recent research is the idea that generative AI does not only provide information but also influences how much control consumers feel they have during decision-making. When users interact with AI-based systems, they often experience a reduction in cognitive effort, which makes the decision process feel easier and more manageable. This perceived control seems to be an important mechanism behind why these systems can increase engagement and purchase intention, especially in situations where consumers are faced with too many alternatives or unclear product differences (Anggraini et al., 2026).

At the same time, the effectiveness of generative AI is not the same across all contexts. Studies suggest that its impact depends significantly on the type of product being evaluated. For example, AI support tends to be more useful when consumers evaluate experience goods, where personal judgment and subjective evaluation are important. On the other hand, when products can be evaluated based on objective characteristics such as technical specifications or price, consumers may rely less on conversational assistance and more on structured information. This makes the role of AI in purchasing decisions quite context-dependent rather than universal (Anggraini et al., 2026).

Usage patterns in the case of ChatGPT showed that conversations on topics of Practical Guidance amounted to 29% of overall usage and those for topics related to Seeking Information has grown to 24% of all usage. Practical Guidance includes conversations on How-To Advice, Tutoring or Teaching, Creative Ideation, and Health, Fitness, Beauty, or Self-Care while the “Seeking Information” category includes conversations on Specific Info, Purchasable Products and Cooking and Recipes (Chatterji et al., 2025).

If we move from generative AI to chatbot-based systems, the literature shows a similar pattern: communication style matters a lot. Whang et al. (2022) highlight that the way chatbots deliver messages directly affects how consumers perceive control during the shopping process. When chatbot responses are more structured and personalized, users tend to feel more confident and less uncertain about their choices, which ultimately increases their intention to purchase. This suggests that it is not only what the system says that matters, but also how it is communicated. Persuasive personalization in recommendations balanced with transparency and trustworthiness proved to be the combination that is more likely to be accepted by shoppers (Sharma et al., 2025).

Another important aspect that keeps appearing in recent studies is anthropomorphism. In simple terms, this refers to how “human-like” a chatbot feels during interaction. Research shows that when chatbots behave in a more human-like way, users tend to engage more easily and respond more positively. Igarashi et al. (2026) found that anthropomorphic characteristics such as warmth and competence can significantly

increase consumer engagement and positively influence purchasing behavior, especially in environments where customers interact repeatedly with the system, such as subscription-based services.

What is also worth noting is that these human-like characteristics do not only affect behavior, but also emotional perception. Consumers often perceive anthropomorphic chatbots as more socially present and more trustworthy, which strengthens their willingness to continue interacting with them. This emotional dimension seems to be quite important in explaining why AI systems can influence not just decisions, but also long-term consumer relationships (Igarashi et al., 2026).

Beyond design and appearance, responsiveness and communication quality also play a major role in shaping consumer reactions. Pang et al. (2025) show that when AI systems provide fast, conversational and relevant responses, users tend to feel more satisfied with the overall shopping experience. Interestingly, such interactions can also encourage more impulsive purchasing behavior, mainly because they reduce hesitation and make decision-making feel more effortless.

In addition to this, responsive AI systems also contribute to a more personalized experience, which strengthens the perception that the platform understands individual needs. This kind of interaction increases engagement and makes consumers more likely to rely on AI support when making decisions in online environments (Pang et al., 2025).

Overall, the literature consistently shows that AI-driven systems influence consumer behavior through several mechanisms, such as perceived control, personalization, anthropomorphism and communication quality. However, what is still not fully explored is whether the advice provided by LLM-based systems remains consistent when consumers face similar decision-making situations. This is important because inconsistent recommendations could weaken trust and reduce the reliability of AI-assisted decision-making processes. At the moment, this aspect remains underexplored in the literature, which creates a clear research gap for further investigation (Anggraini et al., 2026; Whang et al., 2022; Igarashi et al., 2026; Pang et al., 2025).

2. Research methodology

We chose LLMs as the subjects of our experiment because when consumers are faced with purchasing decision LLMs as standalone websites are perceived as impartial as opposed to chatbots and other types of AI shopping assistant that are embedded on shopping websites.

The process of soliciting a LLM to help make a decision regarding the best option in a purchasing process falls at the intersection of “Practical Guidance” and “Seeking Information” as described by Chatterji et al. (2025), as making a purchase decision can be either based on personal details and individual preferences - practical guidance category - or highly technical when it lacks the above and it's based on characteristics, data, statistics and numerical reviews - seeking information category.

We will try to avoid any potential misuse of LLMs that can occur when asking them for advice from situations such as: an initial prompt that offers insufficient details, confusing instructions regarding the action required or the output requested, or not enough details available for aiding the LLM further in the process. Therefore we will establish a scenario, decide on the prompts to be used and pre-research the product models that are going to act as subjects for the purchasing decision requested from the LLMs.

The research is an experiment based on a set scenario in which prompts are given to LLMs in order to observe and analyze the way LLMs answer in the case of making a purchasing decision.

We created a scenario in which the user who is going to be making the purchasing decision with the help of the LLM is an adult member of a household consisting of two adults and a child. The user's gender was established as male in accordance with the fact that, while all household members are traditionally involved in making decisions regarding the household products, the father's opinion is more important when buying technological products (Križan, 2022). Like 73% of the urban EU population (Eurostat, 2025a) they live in a flat. They also own a cat, the most popular household pet in Europe (FEDIAF EuropeanPetFood, 2024). The user will request help in choosing a vacuum cleaner (the first test) and a vertical iron steamer (the second test) for this household. These appliances fit in the category of „Furnishings, household equipment and routine household maintenance” which account for 4.8% of annual consumption expenditure in a European household (Eurostat, 2025a).

Guidelines for implementing the scenario are as follows:

- The above information will not be offered initially but is established in order to be available in case the LLM could not answer and requested details.

- In order to ensure that no previous interaction with the LLMs would influence the results obtained, new accounts were created for use in the queries.
- Three LLMs were used: ChatGPT (from OpenAI, USA), Mistral (France) and DeepSeek (People's Republic of China). The selection of the three is based on popularity in each region.

To this point, based on a preliminary human-made comparison of the main technical characteristics, we have selected for the first test two models of cordless stick vacuum cleaners, Dyson V15 Detect and Dreame Z30, and for the second test two models of vertical iron steamers, Philips STH3010/70 and Xiaomi BHR9034EU. Their complete name is going to be used in the prompts but, for ease of analysis, these will be referred to only by their brand name (Dyson, Dreame, Philips and respectively Xiaomi) in the paper.

The initial prompts request the LLM for a comparison and a decision:

"I want to choose which vacuum cleaner/ vertical iron steamer to buy. Compare these two models and tell me which one is better: Dreame Z30 or Dyson V15/ Philips STH3010/70 or Xiaomi BHR9034EU".

We stated the context in which we needed the answer – choosing the type of appliance to buy – and requested the form of the answer – comparison followed by decision.

The first answer from the LLMs will give out our initial data and we could then move on to the second stage of the experiment in which we will compare the answers. In order to have a clearer comparison we acknowledge numerical comparison to be the more accurate than a text-based one so we will request the LLMs to grade the appliances on a scale. Thus the secondary prompt for the LLMs is:

"Rate the following characteristics for each of the two models: suction power, dustbin volume, battery / steam flow rate, water tank volume, cord length, on a scale of 1 to 5 (where 1 is the lowest score and 5 is the highest)."

At this point, depending on whether the LLM provides on its own a final decision or score as a result of the second prompt, a third prompt requesting it may be used:

"Give an overall score for each model."

3. Results and discussion

The answers from the two tests, the vacuum cleaners and the vertical iron steamers, are analyzed in parallel.

The efficiency of the answers to the first prompt can be analyzed in two ways. The first is to look at the information in the answers in order to see if the LLM respected the user's request in terms of answer structure, meaning comparison followed by decision. The second way of analyzing the answers' efficiency is to look at the number of sections of content that make up a complete answer and at the language used.

The initial prompt leads to answers with different structures (Table no. 1). All LLMs did the requested comparison: two of them, ChatGPT and DeepSeek, chose a table as the best option to show the comparison in both tests, while Mistral made two subsequent lists for the comparison. Looking at how the information is organized, DeepSeek is the only one that starts off with the decision then shows the comparison process, while the other two respect the order in the prompt (comparison then decision).

Out of the three, once the comparison is displayed, Mistral is the one that, in both tests, first presented recommendations for both product models matched with what the user might be interested in and then stated the chosen product model. The other two would follow the user request more closely: only after choosing one of the two product models as the recommended one would they offer additional arguments for specific uses.

In terms of sections that make up a complete answer, Mistral is the more efficient of the three, putting out in both tests a concise answer with a structure of three sections: comparison, recommendation for both products depending on user interests and final choice. Depending on the products compared, ChatGPT and DeepSeek have answers of variable length and structure, from 5 to 7 sections.

In what concerns additional information that where not required by the user, we observe that for both tests ChatGPT starts its answers with a section of product identification: photo, complete model name, star rating, price and main characteristics. It's placement at the beginning of its answers is both helpful in eliminating any user errors and a way to make the interaction friendlier by having images included in the answers.

The two tests show there is structure to the answers that each LLM provides. The only variations come from the two products that are the focus of the analysis, how obvious the differences between the two objects are having an impact on the length of the analysis and the final conclusion.

Table no. 1. Answer structure to the initial prompt in the two tests

ChatGPT answer		Mistral answer		DeepSeek answer	
vacuums	vertical iron steamers	vacuums	vertical iron steamers	vacuums	vertical iron steamers
Product identification	Product identification	Comparison in bulleted list	Comparison table	Decision	Decision
Comparison table	Comparison table	Recommendations for each model based on what the user might value more	Recommendations for each model based on what the user might value more	Comparison table	Comparison table
Recommendations for each model based on what the user might value more	Clear winner followed by presentation of the other product's strong points	Final decision with caveat for choosing the other product if the user is interested in other characteristics	Verdict section: one product is declared "clear winner" while the other is recommended only for particular uses	Opinions on different characteristics ("What the Experts Say")	Why Choose the Philips
Argumentation for the two decisions	More recommendations by situation (less detailed than above recommendations but more specific points to consider)	Offer for a more in-depth analysis should the user want it (tailored request)	No request for additional information from the user	Weak points for the selected winner ("Two Minor Drawbacks to Consider")	Why Choose the Xiaomi
More recommendations by situation (less detailed than above recommendations but more specific points to consider)	No request for additional information from the user			Final ideas if the user wants to choose the Dreame Z30 ("Go with the Dreame Z30 if...")	Final Verdict
Final ideas for both choices				Final ideas if the user wants to choose the Dyson ("Stick with the Dyson V15 if...")	Question for more information (generic request)
No request for additional information from the user				Acknowledgment that more information could improve the answer. (tailored request)	

Source: Author's own elaboration based on the answers to initial prompt

The requests for more information encountered in the two tests show two types of requests: generic requests (“Are there any other features you would like me to compare?”) and tailored requests, that use appropriate information gathered for the answer (“Would you like a deeper dive into any specific aspect, like battery life or accessory comparisons?”, “Do you have mostly carpets or hard floors? Knowing that could help narrow it down further.”). The latter feel more professional just as a knowledgeable shopping assistant. It is interesting to notice that while Mistral and DeepSeek employ both types of requests (an indication that no particular type of request has been proven as best, both types are efficient depending on other factors), ChatGPT doesn’t request further information in any of the two tests.

As for the moment LLMs are not monetized and have free for individual use accounts, the requests for more information could at one point in time also serve as a method of engaging and keeping the user on the website for longer periods of time and directing him towards affiliated links at the end of the interaction.

The second prompt was designed in order to allow the observation of results similarity by turning the information gathered for the comparisons into numerical scores (Table no. 2). The aim was to see if the three comparisons for each set of products were now similar.

By design, the three chosen characteristics for each type of product where technical quantifiable characteristics, their numerical value easily comparable even by the user. Thus, the process of grading and comparing them should have been very straightforward but the LLMs could provide a faster way to do that comparison and reach a conclusion.

Table no. 2. Scores (from 1 to 5) awarded to each product model characteristic following the second prompt

	ChatGPT	Mistral	DeepSeek	ChatGPT	Mistral	DeepSeek
	Dyson			Dreame		
Suction power	4	5	4	5	5	5
Dustbin volume	3	4	4	4	3	4
Battery	4	4	3	5	4	5
	Philips			Xiaomi		
Steam flow rate	3	4	3	4	5	3.5
Water tank volume	2	2	2	4	4	4
Cord length	4	4	3	3	3	2.5

Source: Author’s collection of data from the answers to second prompt

The ideal result for this prompt would have been getting the same numerical score for each characteristic from all three LLMs. However, that did not happen and we can explain the different scores the three LLMs gave to the same characteristics of each model as a result of the models having different inputs for AI training (such as product reviews websites, Reddit posts, YouTube videos etc.). Only in case of one characteristic (the water tank volume for the vertical iron steamers) did all LLMs agree on the exact same score in both tests.

Although the three LLMs are from different areas of the world the sources of information are international and no inference in the sense of patriotism can be deduced based solely on the given scores. This means that while for the vacuums the Chinese LLM gave the Chinese vacuum model the highest score the American LLM also did so (although each evaluated the three characteristics differently) while in the case of the vertical iron steamers the Chinese model actually received lower scores from the Chinese LLM than from the others.

The third and final prompt requesting an overall score was necessary almost every time, DeepSeek being the only one who provided an overall score on its own but only in the case of the vacuums.

As a result of this prompt, one of the tests now concludes with the same final answer from the three LLMs (Table no. 3). However, considering the nature of the products in the tests and the content of the answers received, we acknowledge this as a coincidence and cannot attribute it to the LLMs' similar reasoning.

Table no. 3 Final prompt answers for deciding on a product model

	Vacuums			Vertical iron steamers		
	ChatGPT	Mistral	DeepSeek	ChatGPT	Mistral	DeepSeek
Winner product model	Dyson	Dyson	Dreame	Xiaomi	Xiaomi	Xiaomi

Source: Author's collection of data from the answers to third prompt

A noticeable distinction between the three LLMs appears in this part of the test. Mistral fulfills the prompt to the letter adding no more information. ChatGPT and DeepSeek can feel overwhelming for the user because they both add more arguments reinforcing their scores and final decisions and they acknowledge that price should be an important factor in the final decision.

Limitations of this research primarily deal with the number of tests conducted. Consistency in the structure of the answers for each LLM on their own cannot be commented on based on just the two tests. A significantly higher number of tests (product comparisons) is required in order to reach valid observations.

Another limitation comes from the subject of the research: the LLMs are decoding language and reassembling it in appropriate answers and then, in this experiment, we ask for another codification through score. In interpreting the language of reviews, there are nuances an LLM has to deduce in order to evaluate their meaning for consumers. We can easily consider the case of a product characteristic, like the length of the electrical cordless: a long electrical cord could be an advantage or a hindrance depending on the product.

Conclusions

As users of various ages are nowadays becoming more adept and comfortable with the use of AI in everyday life in the shape of LLMs, chatbots, shopping assistants, and many others it is crucial for the consumers to be able to rely on the information they receive from them. In the case of LLMs this situation is also under the pressure of the fact that, although for the moment free accounts for individual use exist, the AI companies will soon have to be able to profit from them as well.

This paper showed that in the case of asking a LLM for help in a purchasing decision the process and results vary. Even if the input was the same and its nature was technical, LLMs gathered their information from different sources, processed and presented it in different ways. Variations arose from either specific LLM patterns of action or were dependent on the product category or the similarity of the product models in the comparison.

While using more sources and taking less time, the process of using LLMs in making purchasing decisions doesn't necessarily prove to be better than making the decision based solely on human effort and possibilities.

The experiment in this paper can be easily replicated for any category of products and with a larger number of tests in order to draw the necessary conclusions regarding the similarity between results of different LLMs and each LLM's capability to put out consistent analyses.

Acknowledgement:

Gabriel Cristian Sabou's research is partially supported by the project "Societal and Economic Resilience within multi-hazards environment in Romania" funded by European Union – NextgenerationEU and Romanian Government, under National Recovery and Resilience Plan for Romania, contract no.760050/23.05.2023, cod PNRR-C9-I8-CF 267/ 29.11.2022, through the Romanian Ministry of Research, Innovation and Digitalization, within Component 9, Investment I8.

Generative Artificial Intelligence (GenAI) usage statement

ChatGPT, Mistral and DeepSeek were used solely as subjects of the experiment in the research. No GenAI tool was used in the development of this paper, either in the planning or the writing of the paper.

References

- Angraini, L., Demoulin, N. T. M., & De Kerviler, G. (2026). Generative AI and customer engagement with the retailer: Does product type matter? *Journal of Business Research*, 211, 116182. <https://doi.org/10.1016/j.jbusres.2026.116182>
- Chatterji, A., Cunningham, T., Deming, D. J., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025, September). *How people use ChatGPT* (NBER Working Paper No. 34255). National Bureau of Economic Research. <https://doi.org/10.3386/w34255>
- Chen, T., Samaranayake, P., Cen, X., Qi, M., & Lan, Y.-C. (2022). The impact of online reviews on consumers' purchasing decisions: Evidence from an eye-tracking study. *Frontiers in Psychology*, 13, 865702. <https://doi.org/10.3389/fpsyg.2022.865702>
- Crowley, J., Standish, J., Weiss, E., Wright, O., & Blackburn, E. (2024). *The empowered consumer: The key to deeper consumer relationships: Reducing the noise around decision-making*. Accenture. Retrieved December 2, 2025, <https://www.accenture.com/content/dam/accenture/final/accenture-com/document-2/Accenture-The-Empowered-Consumer.pdf>
- Dixit, R. S., Choudhary, S. L., & Govil, N. (2025). Analyzing the impact of artificial intelligence on the online purchase decision-making process through the lens of the UTAUT 2 model. *Discovery Computing*, 28, 88. <https://doi.org/10.1007/s10791-025-09575-5>
- Eurostat. (2025a). *Housing in Europe – 2025 edition*. European Commission. Retrieved December 2, 2025, <https://ec.europa.eu/eurostat/web/interactive-publications/housing-2025#how-we-live>
- Eurostat. (2025b). *Household budget survey – statistics on consumption expenditure*. European Commission. Retrieved December 2, 2025, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Household_budget_survey_-_statistics_on_consumption_expenditure
- FEDIAF EuropeanPetFood. (2024). *FEDIAF EuropeanPetFood releases its annual Facts & Figures confirming 340 million pets in Europe*. Retrieved December 2, 2025, https://europeanpetfood.org/_/news/fediaf-europeanpetfood-releases-its-annual-factsfigures-confirming-340-million-pets-in-europe/
- Igarashi, M., Onishi, H., Sakai, O., & Hasegawa, S. (2026). How chatbot anthropomorphism shapes customer purchase decisions: Evidence from a field experiment on a cosmetic e-commerce retailer. *Journal of Retailing and Consumer Services*, 91, 104721. <https://doi.org/10.1016/j.jretconser.2025.104721>
- Križan, F., Pleša Puljić, N., & Bogнар, Z. (2022). *The role of women as purchase decision makers in the family*. [Conference presentation] 90th International Scientific Conference on Economic and Social Development – Building resilient society: National and corporate security. Zagreb, Croatia.
- Lackermair, G., Kailer, D., & Kanmaz, K., (2013). Importance of Online Product Reviews from a Consumer's Perspective. *Advances in Economics and Business*, 1(1), 1 - 5. <https://doi.org/10.13189/aeb.2013.010101>
- Pang, H., Wang, Y., Sun, M., & Wang, R. (2026). How artificial intelligence functionalities boost consumer benefit and impulse shopping? The significant role of responsive communication and conversational communication. *Computers in Human Behavior*, 181, 108999. <https://doi.org/10.1016/j.chb.2026.108999>
- Sharma, S. K., Mehta, S., Manu, T. U., Puranik, A. M., Singh, R. K., & Sharma, V. (2025). The impact of AI-powered recommendations on online purchase decisions: A consumer perspective. *Advances in Consumer Research*, 2(4), 1542–1552.
- Whang, J.-B., Song, J. H., Lee, J.-H., & Choi, B. (2022). Interacting with chatbots: Message type and consumers' control. *Journal of Business Research*, 153, 309–318. <https://doi.org/10.1016/j.jbusres.2022.08.012>
- Wright, O., Weiss, E., Leary, B., Blackburn, E., & Ducarre, J. (2025). *Me, my brand and AI: The new world of consumer engagement*. Accenture. Retrieved December 2, 2025, <https://www.accenture.com/content/dam/accenture/final/accenture-com/document-3/Accenture-Me-My-Brand-and-AI.pdf>