Gurugram University Business Review (GUBR)
Volume 6 • Issue 1 • January–June 2026
ISSN: 2582-9718 (Online)

Unveiling the Intellectual Structure of Negative Word of Mouth Research

A Bibliometric Mapping of Themes, Influence, and Evolution

Pooja Sharma[1], Samridhi Tanwar[2] & Jyoti Yadav[3]

ABSTRACT

This study explores the scope of NWOM research using bibliometric analysis to guide future research direction. The researcher followed a systematic data analysis procedure, which resulted in 553 articles after applying the pre-defined research string. Bibliometric coupling and keyword analysis were used to get a comprehensive overview of the current state of the NWOM domain. The researcher used Vos-viewer and R Studio for data analysis and visualisation. The results highlight the most influential authors, journals, nations, and organisations in the field of NWOM. The chronological publication trend shows a significant increment after 2020 in the domain of NWOM. Findings show that the USA ranks first among countries publishing the highest number of articles in this domain. Keyword analysis provides a roadmap for future research. This study will provide helpful insight to academics and marketers and will also help marketers make important decisions regarding brand management.

Keywords: Negative word of mouth, Literature Review, Biblioshiny, Social Media
Paper Type: Literature Review

INTRODUCTION

Word of mouth has long been known as a powerful driver that shapes brand perception from the consumer perspective. Over a few decades, word of mouth has been widely studied online and offline (Mirbabaie et al., 2023). The commercialisation of social media has created a feedback loop among customers and entrepreneurs (Mirbabaie et al., 2023). When customers' expectations are unmet, dissatisfied customers resort to negative word-of-mouth communication (Barber et al., 2011; Romani et al., 2012). Research shows that unsatisfied customers are more likely to respond to negative word of mouth than those who complain about a company (Ferguson & Johnston, 2011). Two types of negative communication have been identified in the market literature. One is "private complaining", and the other one is "public complaining" (Christodoulides et al., 2012; Presi et al., 2014; Nguyen, 2021). Private complaints involve negative discussions about a brand with family and friends (Zeithaml et al., 1996; Fetscherin, 2019). Public complaints occur mainly through social media platforms such as social networks or websites (Fetscherin, 2019). Negative word of mouth often significantly impacts the consumer more than online complaints (Curina et al., 2020).

The existing literature has focused on the word of mouth and electronic word of mouth without differentiating between positive and negative aspects. The overlap between word of mouth and negative word of mouth has created ambiguities, raising questions such as: Is NWOM distinctive enough to be considered a separate theme? What are the significant themes of negative word-of-mouth studies, intellectual structure, and research groups?

The study addresses this gap by conducting a comprehensive bibliometric analysis of NWOM

[1] 1Research Scholar, Department of Management, Indira Gandhi University, Meerpur, Rewari, India.
E-mail: pooja.management.rs@gmail.com
[2] Associate Professor, Department of Management, Indira Gandhi University, Meerpur, Rewari, India.
E-mail: Samridhi.tanwar@gmail.com
[3] Research Scholar, Department of Management, Indira Gandhi University, Meerpur, Rewari, India.
E-mail: jyoti.management.rs2@gmail.com

research, examining these trends, patterns, and intellectual structures. This study identifies the most influential authors using the bibliometric mapping technique. Institutions, journals and countries contributing to negative word-of-mouth research. The study contributes in three ways: Firstly, it provides theoretical insights into how NWOM has been conceptualised and operationalised and gives direction regarding less explored areas and opportunities for future research. Thirdly, it provides managerial relevance by demonstrating how understanding the NWOM can help companies design more effective crisis management strategies.

The following research questions have been developed to examine the intellectual structure and emerging trends in this area:

Research Questions

- RQ 1: What are the publication trends of research on negative word of mouth over the specified period?
- RQ2: Who are the most influential authors in the field of NWOM?
- RQ3: What are the most contributing journals in the field of NWOM?
- RQ3: What are the countries contributing the most to the domain of NWOM?
- RQ4: What are the most cited articles in the domain of NWOM?
- RQ5: Which institutions are leading in NWOM research?
- RQ6: What are the key relationships between the most contributing country, research theme and leading journal?
- RQ7: What are the most frequently occurring keywords in NWOM research?

THEORETICAL BACKGROUND

The concept of word-of-mouth marketing has evolved dramatically over the years. With the mass adoption of the Internet and social networking sites, word of mouth's impact has intensified. Marketing word of mouth dates back to the 1970s when psychologist George Silverman led teleconferenced peer-influenced groups to shape opinions about pharmaceutical products (Scott, 2023). Harrison-Walker (2001) defines word of mouth as informal person-to-person communication concerning organisations, products, and services.

Richin (1983) described how dynamic interpersonal communication among consumers contributes to generating "negative word of mouth". In the era of Web 2.0, "negative word of mouth" can severely damage a company's "brand reputation", primarily due to its rapid dissemination on "social media" (Cassidy et al., 2021).

METHODOLOGY

The data used in this study were obtained from the Scopus database on 15 January 2025. The Scopus database provides comprehensive information and broader coverage than other databases, making it more convenient for practical use. The performance indicators provided by the Scopus database are better than those of the Web of Science (Pranckutė, 2021; Ejaz et al., 2022). This systematic literature review follows a straightforward approach to collecting data from multiple databases (Jain et al., 2021).

Recently, bibliometric analysis gained popularity for evaluating scientific output (Donthu et al., 2021; Jain et al., 2021; Batra et al., 2022). This analysis quantitatively examines the content and identifies leading authors, articles, and journals in the domain of NWOM. It consists of two major components: (i) performance analysis, and (ii) scientific mapping, which work together to form the bibliometric analysis.

Defining Search Terms

A precise and clear methodology should be employed for "bibliometric analysis" (Khatib et al., 2021; Batra et al., 2022). The initial search for the sample extraction was conducted on December 25, 2025. The researcher selected keywords after reviewing the existing literature. Based on this review, the chosen keywords were ("Negative word of mouth" OR "NWOM").

Screening and Selection Procedure

"Web of Science" and "Scopus" are widely recognised databases in academic literature. Scopus is considered a more comprehensive database compared to Web of Science (Farooq, 2022; Kumar et al., 2021; Zainuldin and Lui, 2021; Batra et al., 2022; Pranckutė, 2021;

Farooq, 2022). Therefore, the researcher selected the Scopus database to extract the data.

During the initial search, 757 documents were retrieved. The data must be clean, clear, and transparent to minimise spurious results. To refine the results, the researcher applied inclusion and exclusion criteria. The documents were narrowed down in the second step through year filtration, resulting in 740 papers. Next, subject filtration was conducted in the third step, which reduced the number of documents to 648. In the fourth step, document filtration was applied, yielding 560 papers. Finally, in the fifth step, language and source filtration were used to arrive at the final count of 553 papers (Figure 1 illustrates all filtration steps).

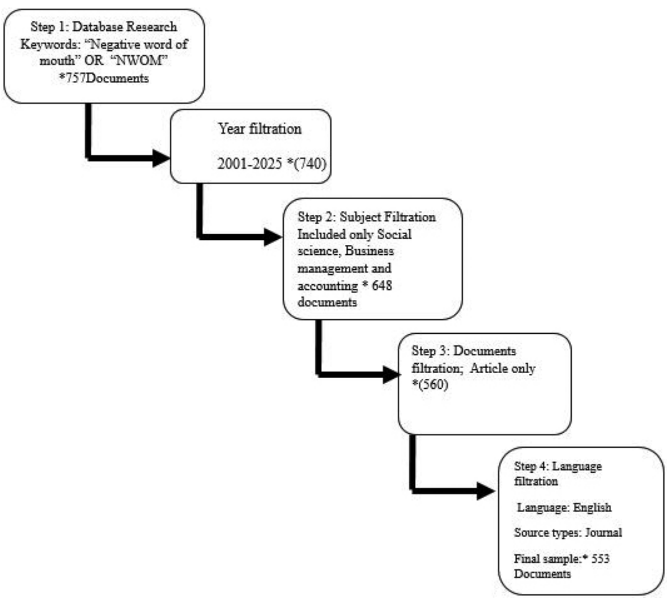


Figure 1: Process of Article Filtration
Source: Author's Elaboration

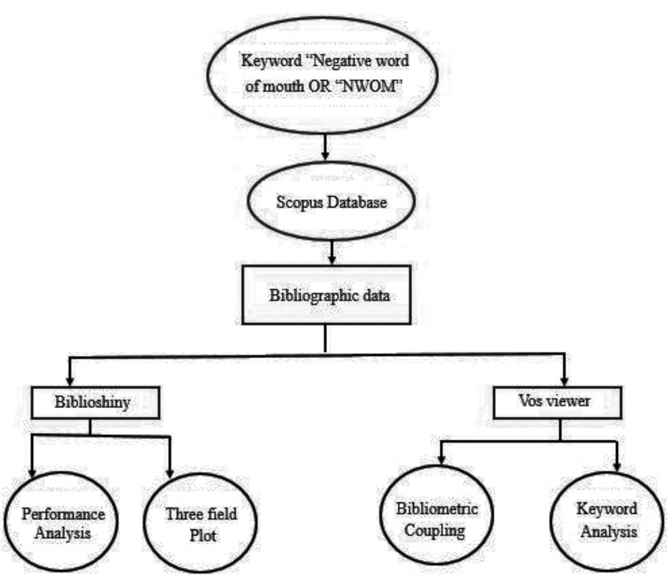


Figure 2

RESULTS AND ANALYSIS

Performance Analysis

These findings show the annual increase of documents from 5 in 2001 to 66 in 2025. In the NWOM domain, only 69 articles were published in the first decade, 2001- 2010. The number of documents published between 2011 and 2020 was 261, about four times that of the previous period. In the remaining years, 2021 to 2025, the number of papers published in this domain was 228. The graph shows the increasing trends in published documents in the NWOM domain.

Table 1: Most Profile Authors on the Topic

Authors	*Articles*	*Articles Fractionalized*
Antonetti P	8	4.00
East R	8	2.50
Mattila AS	8	2.83
Baghi I	7	3.33
Liu Y	5	1.28
Romaniuk J	5	1.50
Talwar S	5	1.05
Bagozzi RP	4	1.42
Balaji MS	4	1.37
Brady MK	4	1.33

Figure 3 shows the most influential authors in the field of NWOM. The findings revealed that Antonetti P stood at the top with eight documents, followed by East, R and Mattila As, with eight papers each.

Table 2: Most Profile Institutions

Affiliation	*Articles*
Pennsylvania State University	14
University of Modena and Reggio Emilia	12
Griffith University	10
University of Vienna	10
Florida State University	9
Ghent University	9
Kingston Business School	9
University of New South Wales	9
Auburn University	8
Kyung Hee University	8

Figure 4 shows those institutions that have made significant contributions to the field of NWOM.

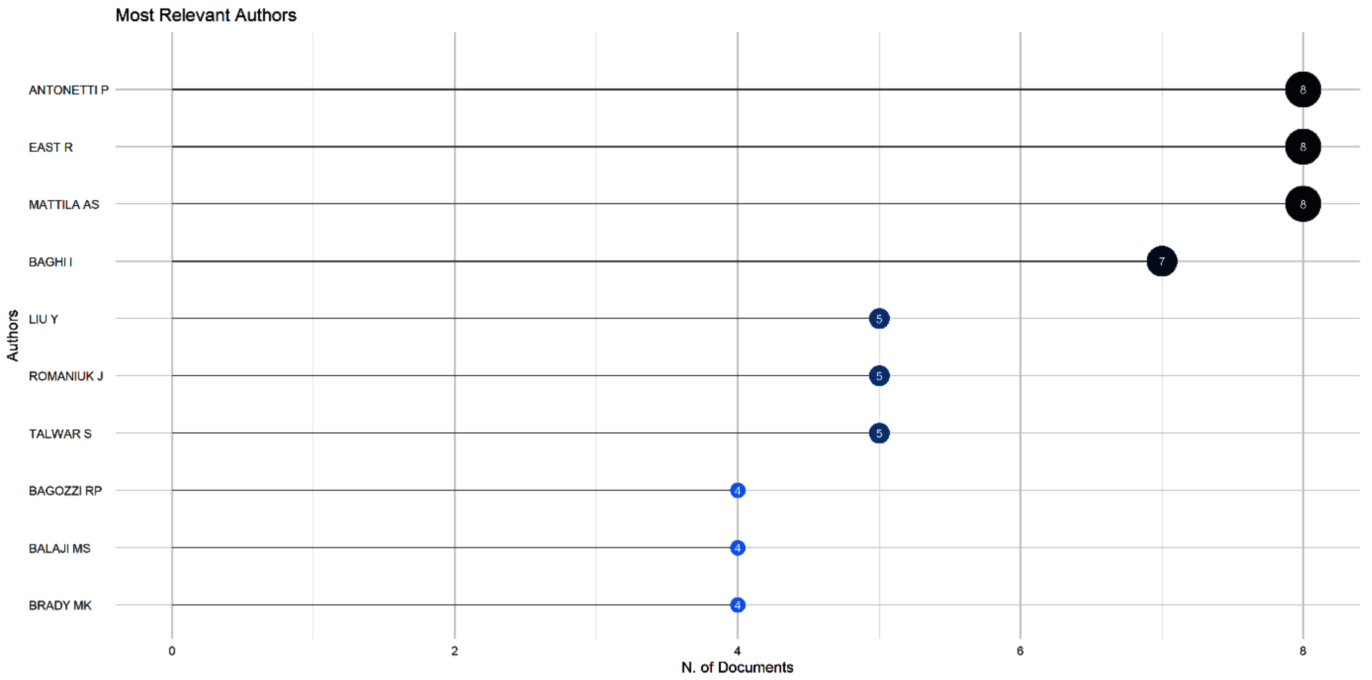


Figure 3: Most Profile Authors in the Field of NWOM

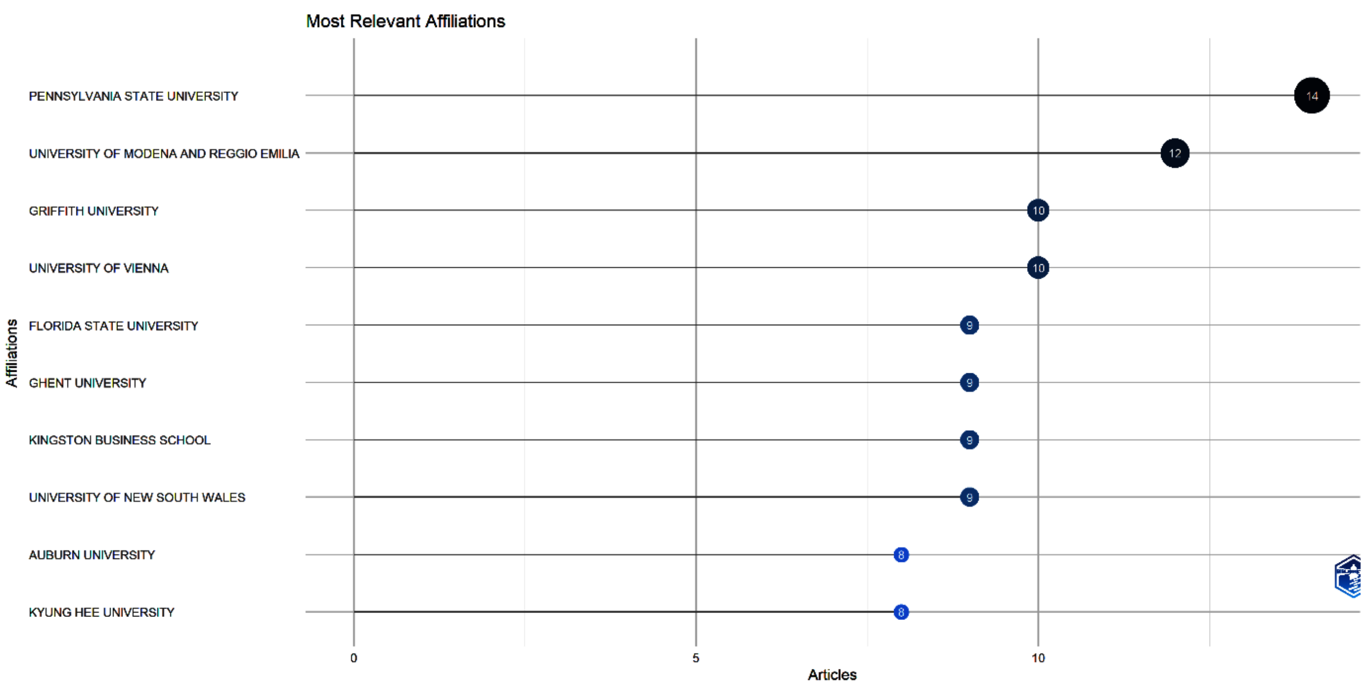


Figure 4: Most Profile Institutions in the Field of NWOM

Pennsylvania State University, the University of Modena, and Reggio Emilia are in the top position with 14 and 12 documents, respectively. Other institutions that made significant contributions are Griffith University and the University of Vienna, with 10 papers each.

Table 3: Most Profile Countries

Country	*Freq*	*Country*	*Freq*
USA	466	Germany	52
China	175	South Korea	47
Australia	108	Italy	40
UK	73	Brazil	37
India	64	Canada	28

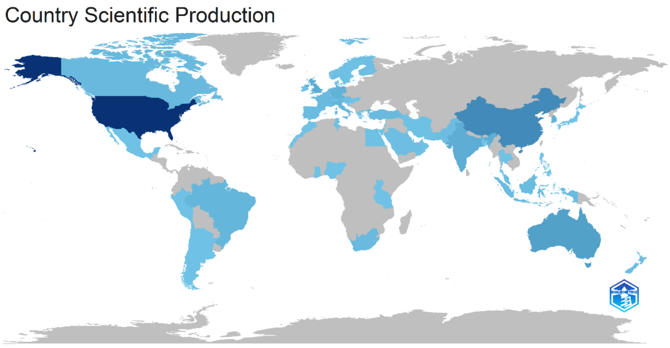


Figure 5: Most Contributing Countries on the Topic of Negative Word-of-Mouth

Figure 5 presents the countries that contribute the most to the NWOM domain. Table 3 shows that America is in the top position with 466 documents, followed by China with 175 papers and Australia with 108 documents.

Figure 6 presents the most cited document in the field of the NWOM domain. Citation analysis is an essential component of bibliometric analysis. Citation analysis is one of the approaches of bibliometric analysis used to see how one article cites another.

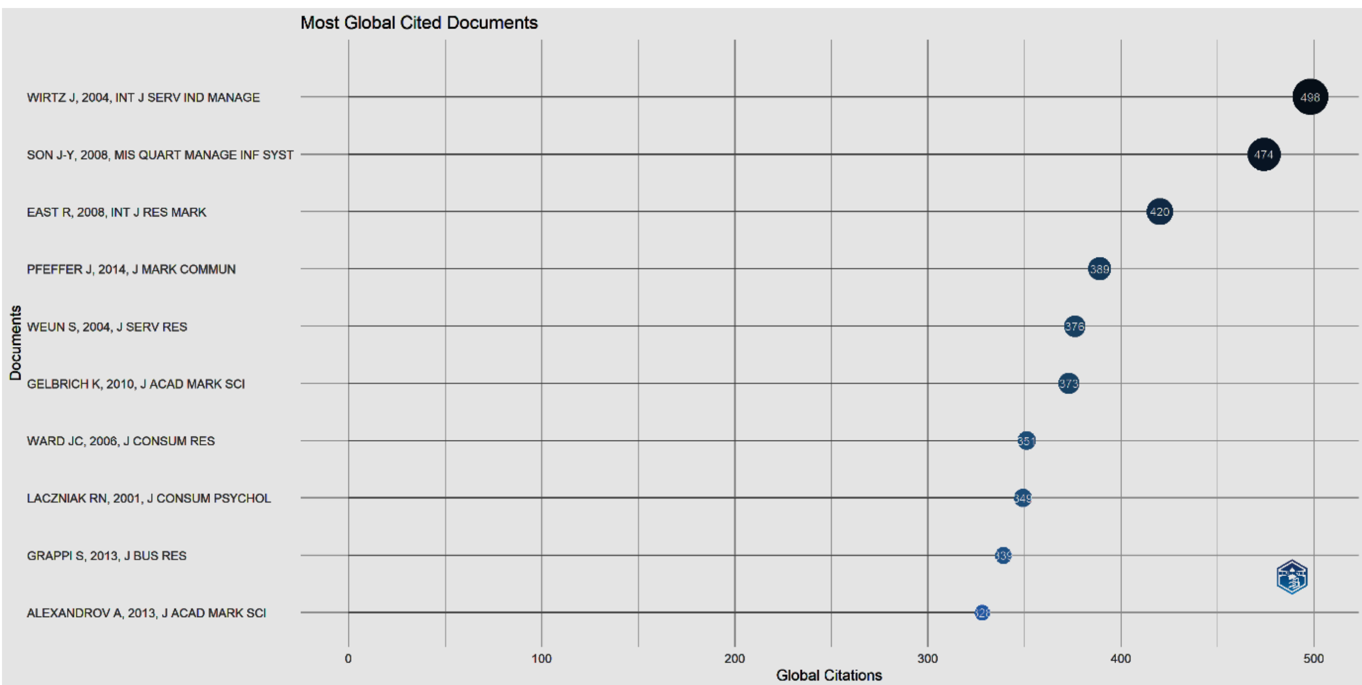


Figure 6: The Most Cited Documents on the Topic of Negative Word-of-Mouth

Table 4: The Most Cited Document in the Field of the NWOM Domain

Paper	*DOI*	*Total Citations*	*TC per Year*
WIRTZ J, 2004, INT J SERV IND MANAGE	10.1108/09564230410532484	498	23.71
SON J-Y, 2008, MIS QUART MANAGE INF SYST	10.2307/25148854	474	27.88
EAST R, 2008, INT J RES MARK	10.1016/j.ijresmar.2008.04.001	420	24.71
PFEFFER J, 2014, J MARK COMMUN	10.1080/13527266.2013.797778	389	35.36
WEUN S, 2004, J SERV RES	10.1108/08876040410528737	376	17.90
GELBRICH K, 2010, J ACAD MARK SCI	10.1007/s11747-009-0169-6	373	24.87
WARD JC, 2006, J CONSUM RES	10.1086/506303	351	18.47
LACZNIAK RN, 2001, J CONSUM PSYCHOL	10.1207/15327660152054049	349	14.54
GRAPPI S, 2013, J BUS RES	10.1016/j.jbusres.2013.02.002	339	28.25
ALEXANDROV A, 2013, J ACAD MARK SCI	10.1007/s11747-012-0323-4	328	27.33
VAN NOORT G, 2012, J INTERACT MARK	10.1016/j.intmar.2011.07.001	322	24.77
LEONIDOU CN, 2015, J BUS ETHICS	10.1007/s10551-015-2829-4	319	31.90
KAU A-K, 2006, J SERV RES	10.1108/08876040610657039	305	16.05

Paper	*DOI*	*Total Citations*	*TC per Year*
JONES MA, 2007, J SERV RES	10.1177/1094670507299382	292	16.22
WETZER IM, 2007, PSYCHOL MARK	10.1002/mar.20178	289	16.06
GEBAUER J, 2013, J BUS RES	10.1016/j.jbusres.2012.09.013	286	23.83
DONTHU N, 2021, J BUS RES	10.1016/j.jbusres.2021.07.015	263	65.75
SOSCIA I, 2007, PSYCHOL MARK	10.1002/mar.20188	256	14.22
LUO X, 2009, MARK SCI	10.1287/mksc.1080.0389	248	15.50
HEGNER SM, 2017, J PROD BRAND MANAGE	10.1108/JPBM-01-2016-1070	246	30.75

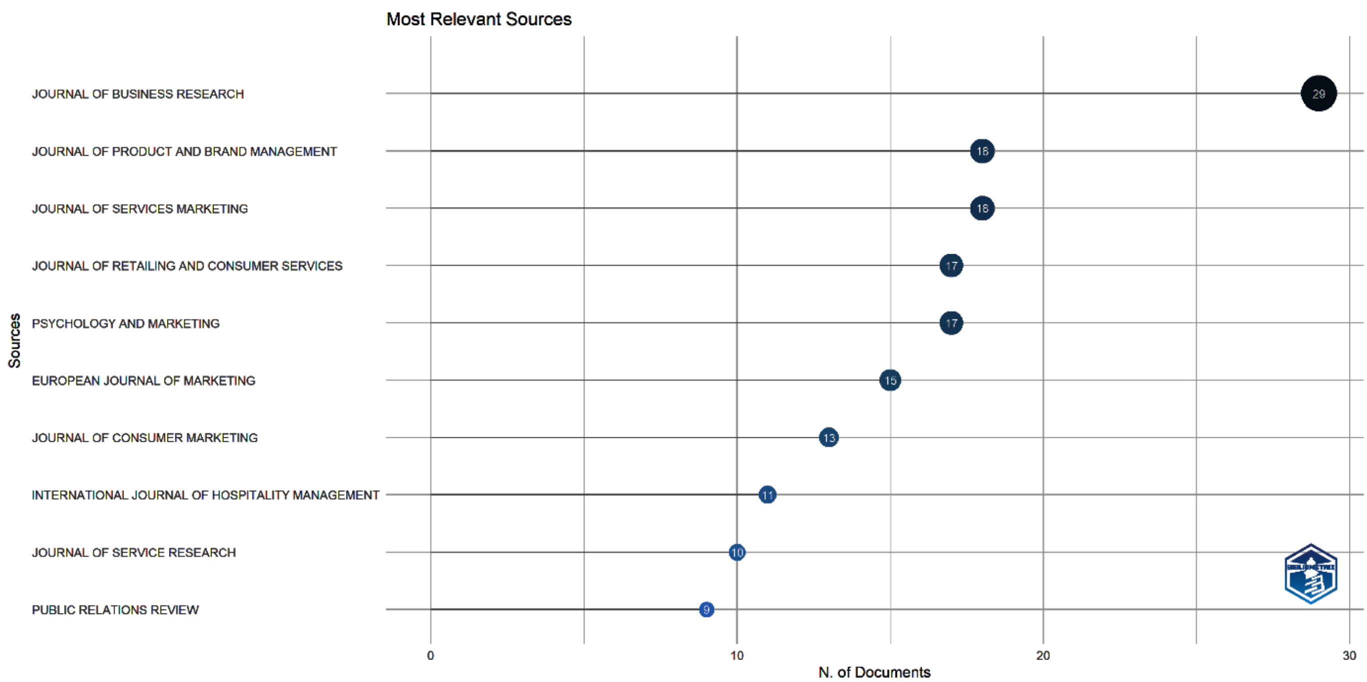


Figure 7: Presents the Most Relevant Journals in Negative Word of Mouth

Table 5: The Most Relevant Journals in the Field of Negative Word-of-Mouth

Sources	*Articles*
"Journal of Business Research"	29
"Journal of Product and Brand Management"	18
"Journal of Services Marketing"	18
"Journal of Retailing and Consumer Services"	17
"Psychology and Marketing"	17
"European Journal of Marketing"	15
"Journal of Consumer Marketing"	13
"International Journal of Hospitality Management"	11
"Journal of Service Research"	10
"Public Relations Review"	9

Figure 7 presents the most profiled journal that publishes the maximum number of articles with negative word of mouth. The table shows the journals that contributed most significantly to it. "The Journal of Business Research" is the most contributed to journal with 29 documents, followed by the "Journal of Product and Brand Management" (18), and the "Journal of Retailing Consumer Service" (17) are in the top position. Most journals are included in the Australian Business Dean Council (ABDC) and are ranked A*, A, and B.

Three-Field Plot

Figure 8 displays three field plots involving the country, keywords, and journals. This figure is based on the popular Sankey diagram (Riehmann et al., 2005). The thickness of the lines illustrates the strength or frequency of these connections. It represents the geographical distribution of research, highlighting the USA, China, Australia, the UK, and India as prominent contributors. Notable journals such as the "Journal of Business Research", "Journal of Service Marketing", and "Psychology and Marketing" are frequently linked with contributors from various countries, indicating their significant impact and relevance in the academic

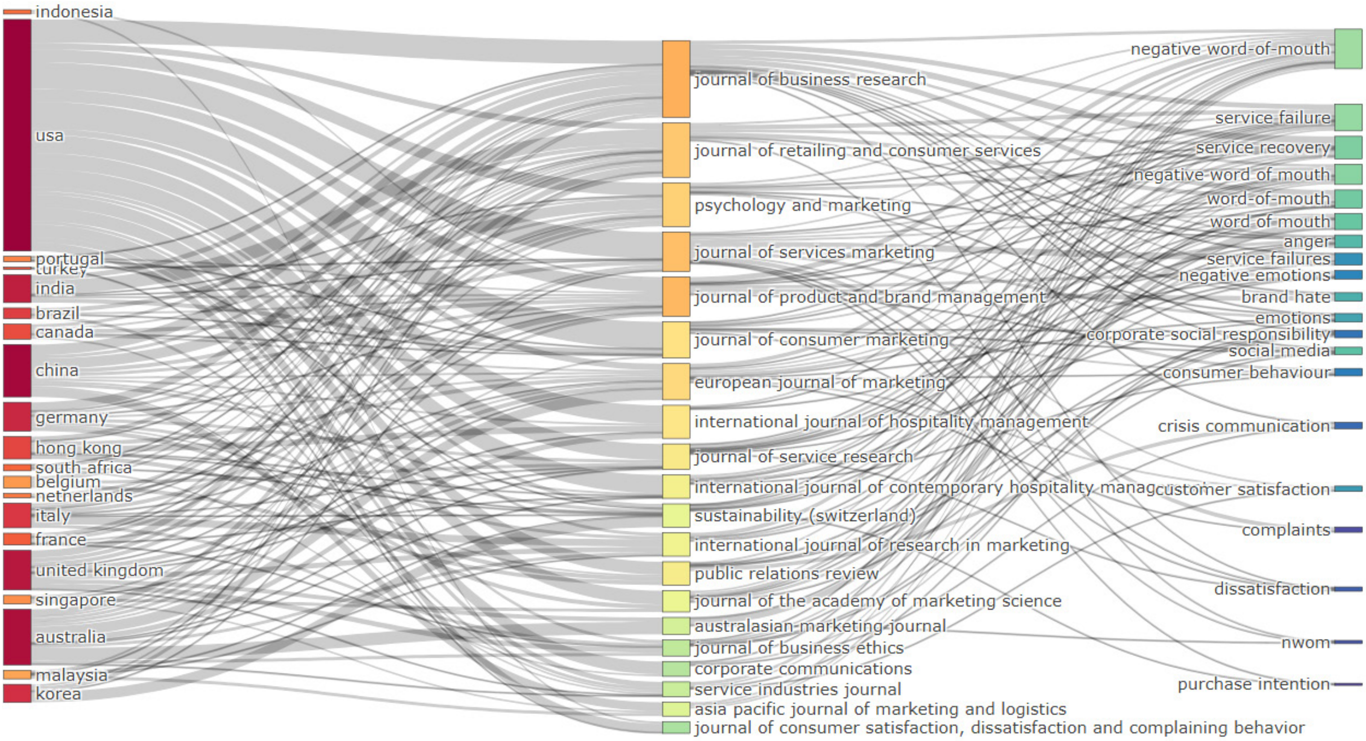


Figure 8: Three Field Plots

field. The rightmost section focuses on key research themes or keywords, including “negative word of mouth”, “service failure”, “brand hate”, and “purchase intention”.

The relationship between these three factors is shown through lines, in which the intensity of the relationship signifies the strength of the relationship. For instance, the diagram indicates that the USA has a strong link to journals such as “The Journal of Business Research” and themes such as “NWOM” and “service failure”. The UK shows a strong connection with the “Journal of Business Research” the “European Journal of Marketing” and the “Journal of Service Marketing” and themes such as “brand hate”, consumer behaviour” and “service Recovery “. China exhibits a strong connection with the journals such as “Journal of Consumer Marketing” “Journal of Service Research” and “Sustainability”, and themes like as CSR, social media, negative emotion and customer satisfaction.

Bibliographic Coupling Analysis

Bibliometric analysis reflects many dimensions of a specific topic (Nilashi et al., 2022). A strong coupling between two publications is indicated when a substantial number of citations are anticipated between them (Jain et al., 2021; Donthu et al., 2021; Bharti et al., 2023). It provides a dynamic approach to examining a subject area’s “intellectual structure” (Koseoglu, 2016; Zainuldin and Lui, 2021; Batra et al., 2022). That’s why we have used bibliometric analysis to learn about clusters in the NWOM domain. Since the documents were cited in large amounts, our threshold limit is 50 of the minimum number of citations; we got 123 qualifying items. But some of these items are not connected. In the final, 6 clusters were formed.

Cluster – 1: Consumer Word-of-Mouth Dynamics and Impact

The red colour cluster represents the largest category of consumer word-of-mouth dynamics, as shown in Figure 9. Consumers often aim to warn others about “negative experiences” with brands (Alexandrov et al., 2013). “Negative word-of-mouth” communication on social media platforms significantly affects individuals’ intentions to engage in further “negative word-of-mouth” behaviours (Balaji et al., 2016; Chakravarty et al., 2010).

Complaints and negative reviews can negatively influence consumer behaviour (Lau & Ng, 2001; Von Wangenheim, 2005). Negative word of mouth can affect the company’s stock return and increase stock

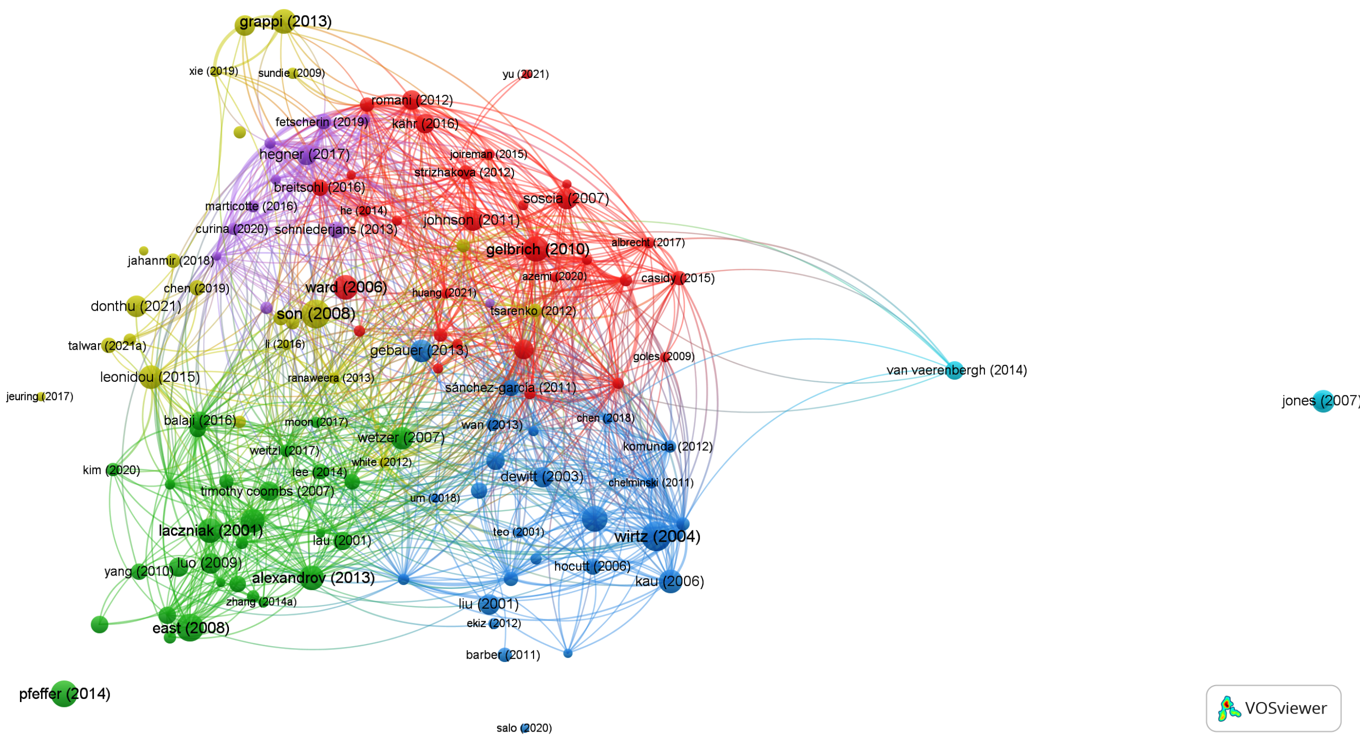


Figure 9: Bibliometric Coupling Among Documents

volatility (Luo, 2009). Research indicates that the impact of "negative word of mouth" is greater than that of positive word of mouth (Nam et al., 2010). Gender also influences how positive and "negative word of mouth" is transmitted (Zhang et al., 2014). To minimise these effects, companies can strengthen brand relationships (Weitzl & Hutzinger, 2017: Tsai et al., 2014).

Cluster – 2: Strategies for Minimising Negative Word of Mouth After Service Failure

Most articles focus on reducing negative word of mouth in this cluster. "Service recovery" influences the consumer purchase decision (Barber et al., 2011; Blodgett & Anderson, 2000). Building a strong rapport with the customer can enhance customer satisfaction and minimise negative word-of-mouth behaviour (DeWitt & Brady, 2003). By understanding customer expectations, a company can effectively manage service recovery (Holloway et al., 2005). Customer satisfaction with service recovery increases customer trust and builds brand loyalty (Ah-Keng & Loh, 2006). Service failure considerably impacts customer trust, commitment, and satisfaction (Weun et al., 2004).

Cluster 3 – Consumer Responses to Service Failure and Recovery Strategies

The blue cluster primarily discusses consumer responses to brands following a "service failure". Such failures often lead to significant consumer anger and negative word-of-mouth (Wirtz & Mattila, 2004; Gelbrich, 2009; Harmeling et al., 2015).

Negative emotion is the result of service failure, which can lead to "customer dissatisfaction" or further "negative word of mouth" (Barari et al., 2020). The negative relationship between the customer and the brand can trigger retaliatory behaviour (Johnson et al., 2011; He & Harris, 2014; Hogreve et al., 2017)

Cluster – 4: Factors Influencing Negative Word-of-Mouth

The yellow cluster highlights the factors responsible for "negative word of mouth". Previous studies show that service failure significantly influenced "negative word-of-mouth behaviour"(Cheng et al., 2005; Li et al., 2016). Brand hate within the service context also plays a significant role in the spread of negative word-of-mouth (Curina et al., 2020). Social media plays an essential role in the spread of negative word-of-mouth (Einwiller & Steilen, 2015). Customer dissatisfaction is also a primary reason for disseminating negative word-of-mouth communication (Nyer & Gopinath, 2005). A collectivist prioritising social harmony is likelier to spread negative word of mouth and complain about service failure (Wan, 2013).

Cluster – 5: Managing Negative Word of Mouth in a Service Context

The cluster in purple highlights the strategies for managing "NWOM" to mitigate the negative emotions of consumers. In the context of service failure, greenwashing also significantly affects consumer trust, affecting customer intention to revisit, which can lead to NWOM (Jones et al., 2007). The company can address the negative emotions of the consumers through effective communication, employee support, and soliciting feedback (Ranaweera & Menon, 2013; Van Vaerenbergh et al., 2014).

Cluster – 6: Consumer Responses and Behavioural Outcomes in Service Failure

The cluster in the sky-blue colour shows the behavioural outcome of service failure. When customers experience service failure, they show their negative emotions in anger, regret, Frustration, and disappointment. These emotions can lead to various behavioural outcomes such as "brand switching" and "negative word of mouth" (Wetzer et al., 2007; McColl-Kennedy et al., 2009; Tuzovic, 2010; Harrison-Walker, 2012, 2019).

Keyword Analysis

Keyword co-occurrence analysis clarifies each thematic cluster's content and offers guidance for further study (Donthu, Kumar, Mukherjee, et al., 2021; Batra et al., 2022).

Keywords within the same cluster are more likely to have similar meanings. This analysis helps researchers identify high-frequency keywords in a specific area after evaluating the existing literature (Bharti et al., 2023).

Figure 10 illustrates the literature's most commonly used "keywords" and their interconnectedness. The size of each node reflects the number of times that keyword has appeared, while the colour indicates the keyword's frequency of use over the years.

The yellow nodes represent emerging keyword trends that are less frequently searched. Recent areas of research that have been notable include "consumer forgiveness," "sharing economy," "artificial intelligence," "avoidance," "brand hypocrisy," and "brand hate." These topics present opportunities for significant research contributions.

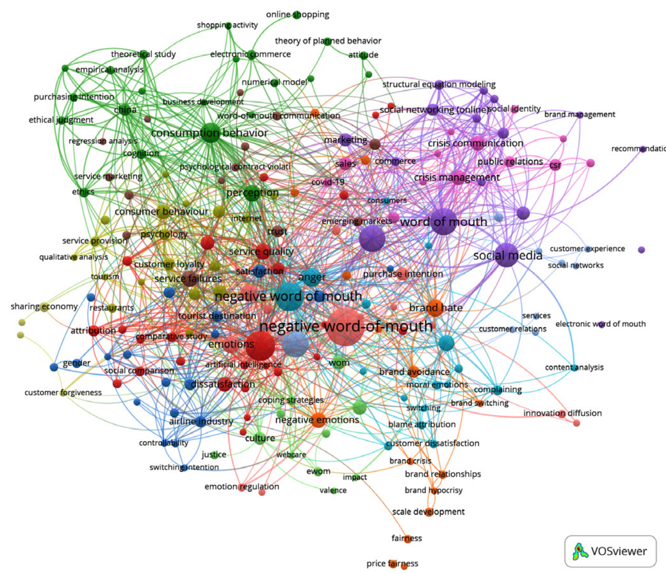


Figure 10: Threshold Limit: 3, Qualifying Amount: 203, Keyword Co-occurrence Analysis

FUTURE RESEARCH AGENDA

Consumer-Related Factors and Negative Word of Mouth

Various studies indicate that when a brand fails to meet consumer expectations, customers may spread negative word of mouth (NWOM) (Barber et al., 2011; Romani et al., 2012; Liao et al., 2023). However, several factors, including consumer characteristics such as age, gender, income level, education level, and behaviour, can influence the dissemination of "negative word of mouth. Future researchers are encouraged to explore the relationship between consumer characteristics and the spread of negative word of mouth.

Negative Word of Mouth as an Antecedent or Outcome of Brand Hate

Some studies show that negative word of mouth is an antecedent of brand hate (Bryson et al., 2021; Jabeen et al., 2022), whereas some argue that it is an outcome of brand hate (Nguyen, 2021; Pinto & Brandão, 2020). Thus, some future studies can explore this relationship.

Practical Strategies to Mitigate the Effects of Negative Word of Mouth

The literature shows that "NWOM" can damage the brand reputation, which can lead to adverse effects such as "brand dilution" and erosion of "firm value" (Bambauer-Sachse & Mangold, 2013; Verhagen et al., 2013; Balaji et al., 2016). The studies on strategies to reduce "negative word of mouth" are limited. Thus, there is a need to investigate further the practical insights for addressing NWOM online and offline.

Impact of Negative Word of Mouth on Brand Love

"Negative word of mouth" impacts the "company's brand reputation" (Cassidy et al., 2021; Bambauer-Sachse & Mangold, 2013; Verhagen et al., 2013; Balaji et al., 2016). Very few studies are available to examine how NWOM affects company brand loyalty. Future studies can explore this area.

NWOM and Brand Switching/Brand Avoidance:

"Negative word of mouth" and "brand switching" are interconnected phenomena in consumer behaviour (Rahimah et al., 2022). "Future studies" should examine the relationships between negative word of mouth, brand avoidance, and brand switching.

CONCLUSIONS/IMPLICATIONS/LIMITATIONS

Conclusions

This study highlights the growing significance of research on "negative word-of-mouth" (WOM) and its continued relevance in today's "digital age". It discusses the evolution of negative WOM research, transitioning from traditional concepts to the modern landscape where social media facilitates the rapid spread of negative feedback.

The researcher employs bibliometric analysis to examine annual publication trends, leading institutions, international collaborations, and key articles within the negative WOM domain. The findings indicate a substantial surge in research related to the "negative word of mouth" domain between 2021 and 2025.

The analysis reveals that the "Journal of Business Research" is the most influential, publishing 29 articles, followed by the "Journal of Service Marketing" with 18 articles, and the "Journal of Product and Brand Management" with 18 articles. Most of these journals are published by Emerald and Elsevier. Antonetti P is identified as the leading author, with eight published documents.

The United States is the leading country in this field, with 466 published documents, followed by China with 175 and Australia with 108. Pennsylvania State University is a prominent institution in terms of negative WOM research, having published 14 papers.

The most cited article in the domain of NWOM research is Wirtz and Mattila (2004) with 498 citations. The researcher also used keyword analysis to suggest emerging areas for further studies, including artificial intelligence, brand hypocrisy, consumer forgiveness and brand hate.

Implications

This study provides a comprehensive overview of existing literature in "negative word-of-mouth research" and identifies the top authors, countries, journals, and institutions. Keyword analysis gives the emerging theme in that area, which would help future researchers explore unexplored regions. With a complete understanding of negative word-of-mouth, marketers can develop effective strategies to mitigate the impact of negative word-of-mouth. Marketers can use this study for crisis management, improving service quality, and strategically engaging with customers. This study serves to identify current trends in the domain of "NWOM" and suggest potential avenues for future studies.

Limitations

This study has many limitations. Firstly, this study is solely based on the Scopus database. Future studies can use multiple databases for better output. Secondly, this study includes an article and a review article only during the document filtration. Future studies could include book chapters or conference papers. Lastly, the bibliometric analysis in this study was conducted using Vos Viewer and R-studio software. Future researchers might consider utilising other software tools such as CiteSpace, PoP, or BibExcel to enhance their analysis.

REFERENCES

Ah-Keng, K., & Loh, E. W. (2006). The effects of service recovery on consumer satisfaction: a comparison between complainants and non-complainants. *Journal of Services Marketing*, *20*(2), 101–111. https://doi.org/10.1108/08876040610657039

Alexandrov, A., Lilly, B., & Babakus, E. (2013). The effects of social and self-motives on the intentions to share positive and negative word of mouth. *Journal of the Academy of Marketing Science*, *41*(5), 531–546. https://doi.org/10.1007/s11747-012-0323-4

Anderson, E. W. (1998). Customer satisfaction and word of mouth. *Journal of Service Research*, *1*(1), 5–17. https://doi.org/10.1177/109467059800100102

Araslı, H., Bavık, A., & Ekiz, E. H. (2006). The effects of nepotism on human resource management. *International Journal of Sociology and Social Policy*, *26*(7/8), 295–308. https://doi.org/10.1108/01443330610680399

Balaji, Khong, K. W., & Chong, A. Y. (2016). Determinants of negative word-of-mouth communication using social networking sites. *Information & Management*, *53*(4), 528–540. https://doi.org/10.1016/j.im.2015.12.002

Bambauer-Sachse, S., & Mangold, S. (2013). Do consumers still believe what is said in online product reviews? A persuasion knowledge approach. *Journal of Retailing and Consumer Services*, *20*(4), 373–381. https://doi.org/10.1016/j.jretconser.2013.03.004

Barari, M., Ross, M., & Surachartkumtonkun, J. (2020). Negative and positive customer shopping experiences in an online context. *Journal of Retailing and Consumer Services*, *53*, 101985. https://doi.org/10.1016/j.jretconser.2019.101985

Barber, N., Goodman, R. J., & Goh, B. K. (2011). Restaurant consumers' repeat patronage: A service quality concern. *International Journal of Hospitality Management*, *30*(2), 329–336. https://doi.org/10.1016/j.ijhm.2010.08.008

Batra, S., Saini, M., Yadav, M., & Aggarwal, V. (2022). Mapping the intellectual structure and demystifying the research trend of cross-listing: a bibliometric analysis. *Managerial Finance*, *49*(6), 992–1016. https://doi.org/10.1108/mf-07-2022-0330

Bharti, Yadav, P., & Singh Yadav, A. (2023). Trends, influential authors and emerging themes in Intelligent Personal Assistants: A Bibliometric analysis. *Tuijin Jishu*, *44*(3), 4146–4163. https://doi.org/10.52783/tjjpt.v44.i3.2264

Blodgett, J. G., & Anderson, R. D. (2000). A Bayesian network model of the consumer complaint process. *Journal of Service Research*, *2*(4), 321–338. https://doi.org/10.1177/109467050024002

Blodgett, J. G., Hill, D., & Tax, S. S. (1997). The effects of distributive, procedural, and interactional justice on post-complaint behaviour. *Journal of Retailing*, *73*(2), 185–210. https://doi.org/10.1016/s0022-4359(97)90003-8

Bozkurt, E. (2023). A Bibliometric analysis of Systems Thinking research in science Education 1991–2022. *Science Education International*, *34*(3), 225–234. https://doi.org/10.33828/sei.v34.i3.6

Bryson, D., Atwal, G., Hultén, P., & Heine, K. (2021). Antecedents of luxury brand hate: A quantitative study. *Strategic Change*, *30*(1), 35–43. https://doi.org/10.1002/jsc.2387

Cassidy, R., & Shin, H. (2015). The effects of harm directions and service recovery strategies on customer forgiveness and negative word-of-mouth intentions. *Journal of Retailing and Consumer Services*, *27*, 103–112. https://doi.org/10.1016/j.jretconser.2015.07.012

Cassidy, M., Burgin, X., & Wasonga, T. A. (2021). Gender differences in perceived barriers of aspiring superintendents. *Management in Education*, 089202062098801. https://doi.org/10.1177/0892020620988010

Chakravarty, A., Liu, Y., & Mazumdar, T. (2010). The differential effects of online Word-of-Mouth and critics' reviews on pre-release movie evaluation. *Journal of Interactive Marketing*, *24*(3), 185–197. https://doi.org/10.1016/j.intmar.2010.04.001

Chen, H., Bernard, S., & Rahman, I. (2019). Greenwashing in hotels: A structural model of trust and behavioural intentions. *Journal of Cleaner Production*, *206*, 326–335. https://doi.org/10.1016/j.jclepro.2018.09.168

Chen, T., Ma, K., Bian, X., Chen, Z., & Devlin, J. F. (2018). Is high recovery more effective than expected recovery in addressing service failure? — A moral judgment perspective. *Journal of Business Research*, *82*, 1–9. https://doi.org/10.1016/j.jbusres.2017.08.025

Cheng, S., Lam, T., & Hsu, C. (2005). Testing the sufficiency of the theory of planned behaviour: a case of customer dissatisfaction responses in restaurants. *International Journal of Hospitality Management*, *24*(4), 475–492. https://doi.org/10.1016/j.ijhm.2004.10.006

Christodoulides, G., Jevons, C., & Bonhomme, J. (2012). Memo to Marketers: Quantitative Evidence for Change. *Journal of Advertising Research*, *52*(1), 53–64. https://doi.org/10.2501/jar-52-1-053-064

Curina, I., Francioni, B., Hegner, S. M., & Cioppi, M. (2020). Brand hate and non-repurchase intention: A service context perspective in a cross-channel setting. *Journal of Retailing and Consumer Services*, *54*, 102031. https://doi.org/10.1016/j.jretconser.2019.102031

DeWitt, T., & Brady, M. K. (2003). Rethinking service recovery strategies. *Journal of Service Research*, *6*(2), 193–207. https://doi.org/10.1177/1094670503257048

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis:

An overview and guidelines. *Journal of Business Research*, *133*, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Donthu, N., Kumar, S., Pattnaik, D., & Lim, W. M. (2021). A bibliometric retrospection of marketing from the lens of psychology: Insights from Psychology & Marketing. *Psychology & Marketing*, *38*(5), 834–865. https://doi.org/10.1002/mar.21472

East, R., Hammond, K., & Lomax, W. (2008). Measuring the impact of positive and negative word of mouth on brand purchase probability. *International Journal of Research in Marketing*, *25*(3), 215–224. https://doi.org/10.1016/j.ijresmar.2008.04.001

Einwiller, S., & Steilen, S. (2015). Handling complaints on social network sites – An analysis of complaints and complaint responses on Facebook and Twitter pages of large US companies. *Public Relations Review*, *41*(2), 195–204. https://doi.org/10.1016/j.pubrev.2014.11.012

Ejaz, H., Zeeshan, H. M., Ahmad, F., Bukhari, S. N. A., Anwar, N., Alanazi, A., Sadiq, A., Junaid, K., Atif, M., Abosalif, K. O. A., Iqbal, A., Hamza, M. A., & Younas, S. (2022). Bibliometric Analysis of Publications on the Omicron Variant from 2020 to 2022 in the Scopus Database Using R and VOSviewer. *International Journal of Environmental Research and Public Health*, *19*(19), 12407. https://doi.org/10.3390/ijerph191912407

Ekiz, E. H., Khoo-Lattimore, C., & Memarzadeh, F. (2012). Air the anger: investigating online complaints on luxury hotels. *Journal of Hospitality and Tourism Technology*, *3*(2), 96–106. https://doi.org/10.1108/17579881211248817

Farooq, R. (2022). A review of knowledge management research in the past three decades: a bibliometric analysis. *VINE Journal of Information and Knowledge Management Systems*. https://doi.org/10.1108/vjikms-08-2021-0169

Ferguson, J. L., & Johnston, W. J. (2011). Customer response to dissatisfaction: A synthesis of literature and conceptual framework. *Industrial Marketing Management*, *40*(1), 118–127. https://doi.org/10.1016/j.indmarman.2010.05.002

Fetscherin, M. (2019). The five types of brand hate: How they affect consumer behaviour. *Journal of Business Research*, *101*, 116–127. https://doi.org/10.1016/j.jbusres.2019.04.017

Gao, S., Chen, S., Huang, R., Guo, Y., Qiu, C., Long, S., Wu, Z., Wang, W., Qiu, H., Zhao, X., & Wang, Y. (2023). Bibliometric Analysis of Research History, Hotspots, and Emerging Trends on Flax with CiteSpace (2000-2022). *Journal of Natural Fibres*, *20*(1). https://doi.org/10.1080/15440478.2023.2194700

Gelbrich, K. (2009). Anger, frustration, and helplessness after service failure: coping strategies and effective informational support. *Journal of the Academy of Marketing Science*, *38*(5), 567–585. https://doi.org/10.1007/s11747-009-0169-6

Gyau, E. B., Sakuwuda, K., & Asimeng, E. (2023). A comprehensive bibliometric analysis and visualisation of publications on environmental innovation. *Journal of Scientometric Research*, *12*(3), 544–557. https://doi.org/10.5530/jscires.12.3.052

Harmeling, C. M., Palmatier, R. W., Houston, M. B., Arnold, M. J., & Samaha, S. A. (2015). Transformational relationship events. *Journal of Marketing*, *79*(5), 39–62. https://doi.org/10.1509/jm.15.0105

Harrison-Walker, L. J. (2001). The measurement of Word-of-Mouth communication and an investigation of service quality and customer commitment as potential antecedents. *Journal of Service Research*, *4*(1), 60–75. https://doi.org/10.1177/109467050141006

Harrison-Walker, L. J. (2012). The role of cause-and-effect in-service failure. *Journal of Services Marketing*, *26*(2), 115–123. https://doi.org/10.1108/08876041211215275

Harrison-Walker, L. J. (2019). The effect of consumer emotions on outcome behaviours following a service failure. *Journal of Services Marketing*, *33*(3), 285–302. https://doi.org/10.1108/jsm-04-2018-0124

He, H., & Harris, L. C. (2014). Moral disengagement of hotel guests' negative WOM: Moral identity centrality, moral awareness, and anger. *Annals of Tourism Research*, *45*, 132–151. https://doi.org/10.1016/j.annals.2013.10.002

Hogreve, J., Bilstein, N., & Mandl, L. (2017). Unveiling the recovery time zone of tolerance: when time matters in service recovery. *Journal of the Academy of Marketing Science*, *45*(6), 866–883. https://doi.org/10.1007/s11747-017-0544-7

Holloway, B. B., Wang, S., & Parish, J. T. (2005). The role of cumulative online purchasing experience in service recovery management. *Journal of Interactive Marketing*, *19*(3), 54–66. https://doi.org/10.1002/dir.20043

Jabeen, F., Kaur, P., Talwar, S., Malodia, S., & Dhir, A. (2022). I love you, but you let me down! How hate and retaliation damage customer-brand relationships. *Technological Forecasting and Social Change*, *174*, 121183. https://doi.org/10.1016/j.techfore.2021.121183

Jain, J., Walia, N., Singh, S., & Jain, E. (2021). Mapping the field of behavioural biases: a literature review using bibliometric analysis. *Management Review Quarterly*, *72*(3), 823–855. https://doi.org/10.1007/s11301-021-00215-y

Jalonen, H., & Jussila, J. (2016). We are developing a conceptual model for the relationship between social media behaviour, negative consumer emotions, and brand disloyalty. In *Lecture Notes in Computer Science* (pp. 134–145). https://doi.org/10.1007/978-3-319-45234-0_13

Johnson, A. R., Matear, M., & Thomson, M. (2011). A coal in the heart: Self-relevance as a Post-Exit Predictor of Consumer Anti-Brand actions. *Journal of Consumer Research*, *38*(1), 108–125. https://doi.org/10.1086/657924

Jones, M. A., Reynolds, K. E., Mothersbaugh, D. L., & Beatty, S. E. (2007). The positive and negative effects of switching costs on relational outcomes. *Journal of Service Research*, *9*(4), 335–355. https://doi.org/10.1177/1094670507299382

Khatib, S. F. A., Abdullah, D. F., Elamer, A. A., Yahaya, I. S., & Owusu, A. (2021). Global trends in board diversity research: a bibliometric view. *Meditari Accountancy Research*, *31*(2), 441–469. https://doi.org/10.1108/medar-02-2021-1194

Kim, J. K., Overton, H., Bhalla, N., & Li, J. (2020). Nike, Colin Kaepernick, and the politicisation of sports: Examining perceived organisational motives and public responses. *Public Relations Review*, *46*(2), 101856. https://doi.org/10.1016/j.pubrev.2019.101856

Köseoğlu, M. A., Rahimi, R., Okumuş, F., & Liu, J. (2016). Bibliometric studies in tourism. *Annals of Tourism Research*, *61*, 180–198. https://doi.org/10.1016/j.annals.2016.10.006

Kumar, S., Pandey, N., Lim, W. M., Chatterjee, A., & Pandey, N. (2021). What do we know about transfer pricing? Insights from bibliometric analysis. *Journal of Business Research*, *134*, 275–287. https://doi.org/10.1016/j.jbusres.2021.05.041

Laczniak, R. N., DeCarlo, T. E., & Ramaswami, S. N. (2001). Consumers' Responses to Negative Word-of-Mouth Communication: An Attribution Theory Perspective. *Journal of Consumer Psychology*, *11*(1), 57–73. https://doi.org/10.1207/s15327663jcp1101_5

Lau, G. T., & Ng, S. (2001). Individual and situational factors influencing negative Word-of-Mouth behaviour. *Canadian Journal of Administrative Sciences / Revue Canadienne Des Sciences De Administration*, *18*(3), 163–178. https://doi.org/10.1111/j.1936-4490.2001.tb00253.x

Li, M., Qiu, S., & Liu, Z. (2016). The Chinese response to hospitality service failure: The effects of face and guanxi. *International Journal of Hospitality Management*, *57*, 18–29. https://doi.org/10.1016/j.ijhm.2016.05.002

Liao, J., Li, X., Yu, G., Han, S., Rong, P., Wang, W., Li, W., & Zhou, L. (2023). Artificial intelligence assists precision medicine in cancer treatment. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/fonc.2022.998222

Luo, X. (2009). Quantifying the Long-Term impact of negative word of mouth on cash flows and stock prices. *Marketing Science*, *28*(1), 148–165. https://doi.org/10.1287/mksc.1080.0389

McColl-Kennedy, J. R., Patterson, P. G., Smith, A., & Brady, M. K. (2009). Customer Rage episodes: Emotions, expressions, and behaviours. *Journal of Retailing*, *85*(2), 222–237. https://doi.org/10.1016/j.jretai.2009.04.002

Mirbabaie, M., Stieglitz, S., & Marx, J. (2023). Negative word of mouth on social media: A case study of Deutsche Bahn's accountability management. *Schmalenbach Journal of Business Research*, *75*(1), 99–117. https://doi.org/10.1007/s41471-022-00152-w

Nam, S., Manchanda, P., & Chintagunta, P. K. (2010). The effect of signal quality and word-of-mouth on customer acquisition for a Video-on-Demand service. *Marketing Science*, *29*(4), 690–700. https://doi.org/10.1287/mksc.1090.0550

Nguyen, H. N. (2021). Antecedents and Consequences of Brand Hate Among Netizens: Empirical Evidence from Vietnam. *Journal of Asian Finance, Economics and Business*, *8*(7), 579–589. https://doi.org/10.13106/jafeb.2021.vol8.no7.0579

Nilashi, M., Abumalloh, R. A., Alrizq, M., Alghamdi, A., Samad, S., Almulihi, A., Althobaiti, M. M., Ismail, M. Y., & Mohd, S. (2022). How does eWOM impact social network sites' travel decision-making during the COVID-19 outbreak? A two-stage methodology. *Telematics and Informatics*, *69*, 101795. https://doi.org/10.1016/j.tele.2022.101795

Nyer, P. U., & Gopinath, M. (2005). Effects of complaining versus negative word of mouth on subsequent changes in satisfaction: The role of public commitment. *Psychology & Marketing*, *22*(12), 937–953. https://doi.org/10.1002/mar.20092

Pinto, Ó. D., & Brandão, A. (2020). Antecedents and consequences of brand hate: empirical evidence from the telecommunication industry. *European Journal of Management and Business Economics*, *30*(1), 18–35. https://doi.org/10.1108/ejmbe-04-2020-0084

Poffo, R. F. (2023). Exploitation and Exploration Innovation. *Revista Catarinense da Ciência Contábil*, *22*, e3373. https://doi.org/10.16930/2237-7662202333732

Pranckutė, R. (2021). Web of Science (WOS) and Scopus: the titans of bibliographic information in today's academic world. *Publications*, *9*(1), 12. https://doi.org/10.3390/publications9010012

Presi, C., Saridakis, C., & Hartmans, S. (2014). User-generated content behaviour of the dissatisfied service customer. *European Journal of Marketing*, *48*(9/10), 1600–1625. https://doi.org/10.1108/ejm-07-2012-0400

Rahimah, A., Dang, H. P., Nguyen, T. T., Cheng, J. C. P., & Kusumawati, A. (2022). The subsequent effects of negative emotions: from brand hate to anti-brand consumption behaviour under moderating mechanisms. *Journal of Product & Brand Management*, *32*(4), 618–631. https://doi.org/10.1108/jpbm-12-2021-3778

Ranaweera, C., & Menon, K. (2013). For better or for worse? Adverse effects of relationship age and continuance commitment on positive and negative word of mouth. *Social Science Research Network*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2540903

Richins, M. L. (1983). Negative Word-of-Mouth by dissatisfied consumers: a pilot study. *Journal of Marketing*, *47*(1), 68–78. https://doi.org/10.1177/002224298304700107

Richins, M. L. (1983). Word-of-mouth communication as negative information. *ACR North American Advances*. https://acrwebsite.org/volumes/6334/volumes/v11/NA-11

Riehmann, P., Hanfler, M., & Froehlich, B. (2005). Interactive Sankey diagrams. *Conference*, 233–240. https://doi.org/10.1109/infvis.2005.1532152

Romani, S., Grappi, S., & Dalli, D. (2012). Emotions that drive consumers away from brands: Measuring negative emotions toward brands and their behavioural effects. *International Journal of Research in Marketing*, *29*(1), 55–67. https://doi.org/10.1016/j.ijresmar.2011.07.001

Scott, S. (2023, September 19). Why Word of Mouth Marketing is Important. *Meltwater*. https://www.meltwater.com/en/blog/word-of-mouth-marketing

Son, J., & Kim, S. S. (2008). Internet Users' Information Privacy-Protective Responses: A Taxonomy and a Nomological Model. *Management Information Systems Quarterly*, *32*(3), 503. https://doi.org/10.2307/25148854

Sweeney, J., Soutar, G., & Mazzarol, T. (2014). Factors enhancing word-of-mouth influence: positive and negative service-related messages. *European Journal of Marketing*, *48*(1/2), 336–359. https://doi.org/10.1108/ejm-06-2012-0336

Talwar, M., Talwar, S., Kaur, P., Islam, A. K. M. N., & Dhir, A. (2021). Positive and negative word of mouth (WOM) are not necessarily opposites: A reappraisal using the dual-factor theory. *Journal of Retailing and Consumer Services*, *63*, 102396. https://doi.org/10.1016/j.jretconser.2020.102396

Tsai, C., Yang, Y., & Cheng, Y. (2014). Does relationship matter? – Customers' response to service failure. *Managing Service Quality*, *24*(2), 139–159. https://doi.org/10.1108/msq-06-2013-0113

Tuzovic, S. (2010). Frequent (flier) frustration and the dark side of word-of-web: exploring online dysfunctional behaviour in online feedback forums. *Journal of Services Marketing*, *24*(6), 446–457. https://doi.org/10.1108/08876041011072564

Van Vaerenbergh, Y., Orsingher, C., Vermeir, I., & Larivière, B. (2014). A Meta-Analysis of relationships linking service failure attributions to customer outcomes. *Journal of Service Research*, *17*(4), 381–398. https://doi.org/10.1177/1094670514538321

Verhagen, T., Nauta, A., & Feldberg, F. (2013). Negative online word-of-mouth: Behavioural indicator or emotional release? *Computers in Human Behaviour*, *29*(4), 1430–1440. https://doi.org/10.1016/j.chb.2013.01.043

Von Wangenheim, F. (2005). Post-switching negative word of mouth. *Journal of Service Research*, *8*(1), 67–78. https://doi.org/10.1177/1094670505276684

Wan, L. C. (2013). Culture's impact on consumer complaining responses to embarrassing service failure. *Journal of Business Research*, *66*(3), 298–305. https://doi.org/10.1016/j.jbusres.2011.08.009

Wang, S., & Huff, L. C. (2007). Explaining buyers' responses to sellers' violation of trust. *European Journal of Marketing*, *41*(9/10), 1033–1052. https://doi.org/10.1108/03090560710773336

Weisstein, F. L., Song, L., Andersen, P., & Zhu, Y. (2017). Examining the impacts of negative reviews and purchase goals on consumer purchase decisions. *Journal of Retailing and Consumer Services*, *39*, 201–207. https://doi.org/10.1016/j.jretconser.2017.08.015

Weitzl, W., & Hutzinger, C. (2017). The effects of marketer- and advocate-initiated online service recovery responses on silent bystanders. *Journal of Business Research*, *80*, 164–175. https://doi.org/10.1016/j.jbusres.2017.04.020

Wetzer, I. M., Zeelenberg, M., & Pieters, R. (2007). "Never eat in that restaurant, I did!": Exploring why people engage in negative word-of-mouth communication. *Psychology & Marketing*, *24*(8), 661–680. https://doi.org/10.1002/mar.20178

Weun, S., Beatty, S. E., & Jones, M. A. (2004). The impact of service failure severity on service recovery evaluations and post-recovery relationships. *Journal of Services Marketing*, *18*(2), 133–146. https://doi.org/10.1108/08876040410528737

Wirtz, J., & Mattila, A. S. (2004). Consumer responses to compensation, speed of recovery, and apology after a service failure. *International Journal of Service Industry Management*, *15*(2), 150–166. https://doi.org/10.1108/09564230410532484

Zainuldin, M. H., & Lui, T. K. (2021). A bibliometric analysis of CSR in the banking industry: a decade study based on Scopus scientific mapping. *International Journal of Bank Marketing*, *40*(1), 1–26. https://doi.org/10.1108/ijbm-04-2020-0178

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The behavioural consequences of service quality. *Journal of Marketing*, *60*(2), 31. https://doi.org/10.2307/1251929

Zhang, Y., Feick, L., & Mittal, V. (2014). How males and females differ in their likelihood of transmitting negative word of mouth. *Journal of Consumer Research*, *40*(6), 1097–1108.

Gurugram University Business Review (GUBR)
Volume 6 • Issue 1 • January–June 2026
ISSN: 2582-9718 (Online)

Assessing Users' Perception on Chatbots Effectiveness

A. Srikanth[1] & T. Mani Sri Anmitha[2]

ABSTRACT

This study looks at how K.L. Business School MBA students in Andhra Pradesh feel about voice-activated chatbots like Siri, Google Assistant, Amazon Alexa, and Bixby. It addresses the paucity of research on user assessments of chatbots performance by examining perceived efficacy, responsiveness, and overall value. Assessing general efficacy perceptions, comparing efficacy across chatbots kinds, determining user preferences, and examining the impact of demographics (gender, place of residence, and location) on perceptions and usage are the goals. A cross-sectional survey of 321 students chosen by proportionate stratified random sampling was utilized to gather data for the exploratory design. JASP, Excel, and R were utilized to analyse responses obtained via Google Forms.

The results show that consumers' opinions about chatbots efficacy are constant across chatbot types and demographic groups. Nonetheless, use patterns indicate brand domination, with Siri, Alexa, and Google Assistant becoming more popular. This implies that user choice may be influenced more by marketing, ecosystem integration, and brand exposure than by perceived performance.

Keywords: Chatbot, Alexa, Google Assistant, Siri, Bixby

INTRODUCTION

Chatbots, also known as conversational agents, are programs created to communicate with users using natural language via text, voice, or multimodal interfaces (Jalawadai, 2024). Processing user input and producing pertinent answers to voice or text enquiries is their main duty (Ghazala Bilquise, 2022). As artificial intelligence (AI) develops quickly, chatbots are using machine learning (ML) and natural language processing (NLP) more and more to mimic human interaction and improve communication accuracy. Large language models like GPT and Gemini have further changed chatbots from rule-based programs to context-aware, adaptive systems that can have dynamic, personalised discussions (Reddy, 2025). Voice-activated assistants, like as Google Assistant, Amazon Alexa, Siri, and Bixby, are now incorporated into smart phones, smart homes, and larger digital ecosystems, greatly impacting customer behaviour and service provision (Martin Adam, 2020). Depending on earlier studies, conversational AI systems with human-like characteristics increase user responsiveness and engagement (Ying Xu, 2022). Nevertheless, user opinions of chatbot efficacy continue to differ despite advancements in technology. Perceived performance and sustained engagement are still influenced by elements like trust, dependability, personalisation, and anthropomorphic design (Najaf & Ding, 2024).

NEED OF THE STUDY

Assessing how consumers see chatbots' efficacy is crucial as businesses use them more and more on digital channels. While technology developments have improved chatbot capabilities, sustained adoption and value are determined by user assessments of responsiveness, usefulness, dependability, and overall performance. In order to strategically improve chatbot design and interaction quality, it is imperative to comprehend these perceptions. In order to improve brand relationships, lower operating expenses, and

[1] Associate Professor, Business School, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.
E-mail: drakondi@kluniversity.in
[2] MBA(Final), Business School, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.
E-mail: anmithatunuguntla10@gmail.com

foster loyalty, customer pleasure is essential. While dissatisfied customers are more inclined to escalate to human agents and utilise self-service less frequently, satisfied users are more likely to trust, stick with, and refer chatbots. According to research, consistent and human-like chatbot interactions promote favourable brand perception and sustained engagement (Xu et al., 2022).On the other hand, unfavourable encounters lower returns on technology investments and erode trust in AI systems. Therefore, to improve user experience and guarantee long-term organisational results, a methodical evaluation of perceived chatbot performance is required.

LITERATURE REVIEW

According to (Adamopoulou, 2020), a chatbot is a computer software created to mimic human communication through spoken or typed exchanges, giving consumers automatic access to information and help. In general, chatbots can be divided into two categories: artificial intelligence-driven systems that employ machine learning and natural language processing (NLP) to comprehend user intent and provide adaptive responses, and rule-based systems that use rewritten scripts and decision trees (Dale, 2016). A more sophisticated type of these is voice-controlled systems, often known as voice assistants, which combine natural language processing (NLP) and automatic speech recognition (ASR) to enable smooth spoken communication between people and machines (Hoy, 2018). Prominent instances including Apple Siri, Google Assistant, Amazon Alexa, and Samsung Bixby show notable advancements in ecosystem integration, contextual awareness, and usability (Lopatovska, 2018).

Voice-activated chatbots have developed over the last ten years from simple speech-to-text tools to complex conversational agents with contextual awareness and tailored responses. More natural and dynamic interaction has been achieved by developments in ASR and NLP, which have also improved speech accuracy and cross-device interoperability. Important turning points in this history are marked by platforms like Siri, Alexa, Google Assistant, and Bixby, which are now essential components of smart speakers, smart phones, and connected home ecosystems (Darda, 2020). Their position in routine digital encounters has been reinforced by their extensive integration.

Depending on systematic reviews, voice assistants have become widely used because speech-based interaction is more convenient and intuitive than text input (Darda, 2020). These systems have revolutionised digital service access by allowing users to carry out tasks like scheduling, information retrieval, navigation, and home automation through natural voice. However, a large portion of the material now in publication places more emphasis on system performance, usability metrics, and technological architecture than on users' subjective evaluations of efficacy across platforms. As a result, nothing is known about how consumers compare the performance of the most popular commercial voice assistants in realistic settings.

According to research on user perceptions, interaction quality, personalisation, and trust have a significant effect on perceived efficacy. Perceptions of utility and credibility are influenced by conversational design characteristics like tone, responsiveness, clarity, and anthropomorphic qualities. Customer satisfaction ratings are generally greater for chatbots that react quickly, precisely, and contextually. These assessments are also influenced by demographic factors. While older users place greater emphasis on task efficiency and functional dependability, younger users frequently place more value on interactive elements and convenience. These variations underline the necessity of investigating how background, age, and experience affect opinions on chatbot efficacy (Pušnik, 2025).

Comparative research also shows that user expectations and usage context affect perceived performance. Voice assistants are exploited by many users for repetitive tasks, and opinions on their responsiveness and dependability differ depending on the user's familiarity with the device and past experiences (Sziron, 2022). People often personify their assistants, giving them human characteristics that influence their propensity to trust them and delegate difficult jobs. Despite these revelations, there is always a dearth of direct comparative studies across Siri, Alexa, Google Assistant, and Bixby across various demographic groups.

There are still significant gaps in the literature, despite the fact that chatbot evolution and usability models are covered in great detail. Instead of concentrating only on voice-activated assistants, numerous studies also examine text-based systems or overall interaction quality. There aren't many comparative studies of the main viable platforms, especially in formal academic settings. Furthermore, rather than being a direct result of perceived efficacy across platforms, demographic factors are frequently associated with adoption. In order to address these gaps, the current study integrates demographic data to provide a thorough understanding of how effectiveness is defined and assessed across platforms and user groups. It also looks at user perceptions of chatbot effectiveness across top voice assistants within a single framework.

RESEARCH GAP

The majority of studies on voice-activated chatbots, including Samsung Bixby, Apple Siri, Google Assistant, Amazon Alexa, IBM Watson, and others, have mostly concentrated on usability, user adoption, and technological advancement. Few research, nevertheless, thoroughly investigate how users perceive chatbot efficacy. It is unknown how chatbot type affects perceived performance because existing research frequently assesses specific platforms rather than comparing several assistants within a single analytical framework. While demographic factors like age, gender, and education are often associated with adoption, they are rarely examined in terms of perceived efficacy. Furthermore, neither the distinction between platform-specific and common user experiences nor the analysis of the ways in which various demographic characteristics interact to influence impressions have received ample attention. There are still insufficient comparative analyses of the leading commercial voice assistants. Therefore, a systematic assessment integrating platform comparison and demographic impact is necessary, which this study aims to provide from a consumer perception perspective.

OBJECTIVES OF THE STUDY

The following are the objectives formulated from the gaps identified by the study.

1. To assess users' perception towards chatbot effectiveness.
2. To examine the effect of type of chatbots onchatbot effectiveness from consumers' perception. (Unique Responses)
3. To analyse chatbot effectiveness among all the chatbots. (Repeated Responses)
4. To study the influence of consumers' demographics on chatbot effectiveness.
5. To examine the association between respondents' demographics and their choice of chatbots.

HYPOTHESES

The following are the alternative hypotheses formulated in-line with the above stated objectives.

$H_1$1: Users perception towards chatbot effectiveness is positive.

$H_1$2: Users' perception towards chatbot effectiveness varies by type of chatbot.(Where respondents' response to the most preferred chatbot is considered – Single Response)

$H_1$3: The preference of consumers towards choosing chatbots is not the same. (Where respondents' response for all chatbots is considered – Repeated response)

$H_1$4: There is a significant impact of demographics on chatbot effectiveness.

$H_1$5: There is an association between chosen demographic variables (Gender, Location and Residential status) and their choice of chatbots.

There are sub-hypothesis for 4th and 5th hypothesis, they are mentioned in results and discussion section.

METHODOLOGY

This section helps in meeting the objectives one by one to understand the respondents' perception on Chatbots Effectiveness. So, this includes the Research Design, Sample Design, Data Collection and Data Analysis adopted by the study.

Research Design

The study has used mixed methodological research design including Exploratory Research Design and Descriptive Research Design, wherein Literature review method of Exploratory Research Design is used for understanding the theoretical and empirical studies on the research topic and helps in identifying the research gap & Cross-Sectional Survey method of Descriptive Research Design is used in collecting the data.

Sample Design and Data Collection

In this design of study, target population are youngsters using voice-activated chatbots to know their perception on their effectiveness, particularly MBA pupils at KL Business School, KLEF. A sample of 175 from a sample frame of 321, is collected using a structured Questionnaire with likert items along with demographic variables including Gender, Location and Residential Status through Google Forms from randomly selected students using Proportionate Stratified Random Sampling (Singh, 2014).

Table 1: Sample Determination Table using Cochran Formulae

Section	*Population*	*Population%*	*Sample%*	*Sample*
Sec A	79	24.61	24.61	43
Sec B	68	21.18	21.18	37
Sec C	47	14.64	14.64	26
Sec D	59	18.38	18.38	32
Sec E	68	21.18	21.18	37
Total	321	100	100	175

To determine the sample elements from each section, sample () in R is used, which is from the base package.

Data Analysis

The study includes two phases of analysis including pilot study and final study. Initially, Questionnaire, a survey instrument is tested for its validity and reliability as a part of Pilot study. Under validity, face validity of instrument is conducted by two industry professionals and five academic members to guarantee the questionnaire's relevance and intelligibility. In addition, Construct validity is also conducted to check the factor validity of Chatbots Effectiveness. To check the reliability of Instrument, Cronbach's alpha is analysed. Response bias is examined by using a chi-square test after data cleaning was completed in R.

As a part of final study, several statistical tools are adopted by the study. For the first objective, Wilcoxon Signed Rank Test is used to measure the Chatbots Effectiveness (i.e., Users' satisfaction scores), as the data is not-normally distributed (Nahm, 2015).

To meet the second, One-way ANOVA is used to examine the effect of type of chatbots (five – AmazonAlexa, Apple Siri, Samsung's Bixby, IBM's Watson and Google's Google Assistant) on Users' Satisfaction score towards Chatbots Effectiveness, being type of chatbots is acting as Independent variable with five categories and Chatbots Effectiveness as dependent variable (María José Blanca, 2017).

For the third, the Friedman Test is used as the data is ranked across the five chatbots (Nahm, 2015).In the fourth Objective, the influence of respondent's demographics on their Satisfaction score towards Chatbots effectiveness is assessed using respective statistical tools namely Mann-Whitney U test for Gender and Residential Status, One-Way ANOVA for location and choice of chatbots on time-spent by respondents (Nahm, 2015). Finally, the fifth objective assess the association between respondents' demographics (i.e., Gender, Residential Status and Location) and their Chatbots Selection (McHugh, 2013) using Chi-square test.

RESULTS AND DISCUSSION

After conducting pilot study analysis, the study further proceed with final study with 16 statements and the conducted reliability analysis is as follows,

Table 2: Reliability Score

Measure	*Value*	*Remarks*
Raw Alpha	0.96	>0.5, reliable.

Table 3: Threshold Values

Method	*Alpha*
Feldt	0.96
Duhachek	0.96

Source: R
Package: psych
Function: alpha ()

From the above Table 3, "Cronbach's alpha of 0.96, the questionnaire exceeded the 0.90 level for excellent internal consistency, indicating outstanding reliability" (Nunnally, 2016). The questionnaire is a very trustworthy tool because of its narrow 95% confidence intervals from Feldt's and Duhachek's methodologies (0.95–0.97), which validate the stability and accuracy of the reliability estimate (Nunnally, 2016).

Overall, the results establish that the questionnaire is a highly reliable instrument. With these results, the study further proceed for final study.

For **data cleaning** the study has usedthe following packages and functions in R:

Base Package- colSums(),sum(),is.na()
VIM Package-irmi()

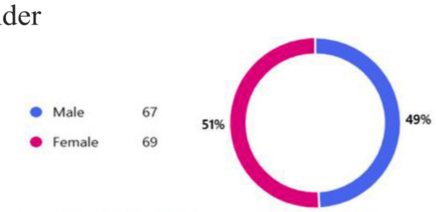
The **descriptive analysis** of data for this study is as follows:

- Categorical Data:
  Gender (2) – Male, Female
  Location (4) – Vijayawada, Guntur, Tenali, Others
  Residential status (2) – Dayscholar, Hosteler
  Chatbots (4) – Amazon_Alexa, Google_Assistant, Siri, Bixby
- Continuous Data:
  Score – Satisfaction from the Likert scale
  Time – No. of hours spent on the selected chatbot.

Bias Justification for data collected:

H_o: Proportions are equal.
H_1: Proportions are not equal.

1. Gender



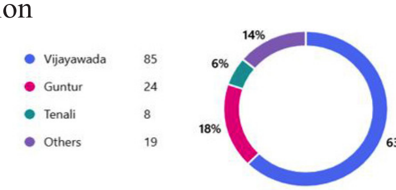

```
> chisq.test(c(51,49),p=c(1/2,1/2))

        Chi-squared test for given probabilities

data:  c(51, 49)
X-squared = 0.04, df = 1, p-value = 0.8415
```

P-value is 0.8145 greater than 0.05, acceptnull hypothesis can conclude that there is no bias (Stefanos Bonovas 1, 2023).

2. Location

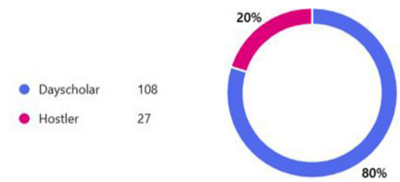



```
> chisq.test(c(63,18,6,14),p=c(1/4,1/4,1/4,1/4))
        Chi-squared test for given probabilities

data:  c(63, 18, 6, 14)
X-squared = 78.208, df = 3, p-value < 2.2e-16
```

P-value is 2.2e-16 less than 0.05, reject null hypothesis and conclude that proportions are not equal (Stefanos Bonovas 1, 2023).

3. Are you a



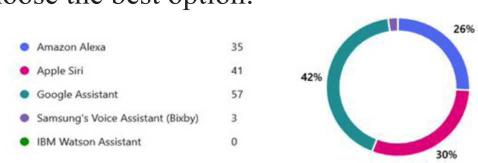

```
> chisq.test(c(80,20),p=c(1/2,1/2))

        Chi-squared test for given probabilities

data:  c(80, 20)
X-squared = 36, df = 1, p-value = 1.973e-09
```

P-value is 1.973e-09 less than 0.05, reject null hypothesis and can say that proportions are not equal (Stefanos Bonovas 1, 2023).

4. From the following Voice Activated Chatbots, choose the best option:



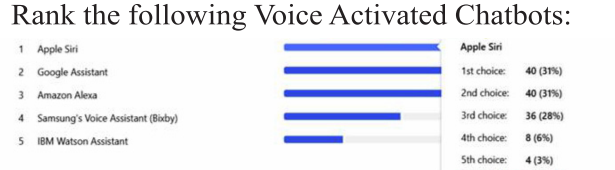

```
> chisq.test(c(30,42,26),p=c(1/3,1/3,1/3))

        Chi-squared test for given probabilities

data:  c(30, 42, 26)
X-squared = 4.2449, df = 2, p-value = 0.1197
```

For analysis the study removed Samsung's Voice Assistant (Bixby) and IBM Watson Assistant for some tests to remove bias. P-value is 0.1197 greater than 0.05,accept null hypothesisand can say that there is no biasin Voice activated chatbots which are Amazon_Alexa, Apple_Siri, Google_Assistant. (Stefanos Bonovas 1, 2023)

5. Rank the following Voice Activated Chatbots:




```
> chisq.test(c(40,40,36,8,4),p=c(1/5,1/5,1/5,1/5,1/5))

        Chi-squared test for given probabilities

data:  c(40, 40, 36, 8, 4)
X-squared = 50.75, df = 4, p-value = 2.518e-10
```

P-value is 2.518e-10 less than 0.05,reject null hypothesis and can say that proportions are not equal (Stefanos Bonovas 1, 2023).

For analysis the study removed Samsung's Voice Assistant (Bixby) and IBM Watson Assistant for some tests to remove bias.

```
> chisq.test(c(40,40,36),p=c(1/3,1/3,1/3))

        Chi-squared test for given probabilities

data:  c(40, 40, 36)
X-squared = 0.27586, df = 2, p-value = 0.8712
```

P-value is 0.8712 greater than 0.05,accept null hypothesisand can say that there is no bias in Voice activated chatbots which are Amazon_Alexa, Siri, Google_Assistant (Stefanos Bonovas 1, 2023).

The analysis shows no bias in the selection of the main chatbots or in gender distribution. Some bias appears in location and when less prominent chatbots are included, leading to their exclusion from certain assessments. Overall findings remain consistent, and future research should use larger, more geographically balanced samples.

The following results are in-line with the above-stated objectives.

From Objective 1

$H_0$1: Users perception towards chatbot effectiveness (Score) is not positive. [mu <= 3 (shows dissatisfaction)]

$H_1$1: Users perception towards chatbot effectiveness (Score) is positive. [mu>3 (shows satisfaction)]

Code Snippet (Source: R)

> shapiro.test(Score)

p-value = 0.0002847

As the p-value is < 0.05, **proceed for wilcoxon signed rank test** (Nahm, 2015)**.**

Table 4: Wilcoxon Signed Rank Test Results

Variable_name	*Test*	*p*	*Remarks*
Score	Wilcoxon	0.123	> 0.05, accept null hypothesis

Source: Jeffreys's Amazing Statistics Program (JASP)

Interpretation

From the above table 4, as the p-value is greater than 0.05, accept null hypothesis and can say that users are dissatisfied with all the functions of chatbots (Stefanos Bonovas 1, 2023).

Murray (2024) reports a sharp fall in trust, with usage dropping from 73% to 60% in just over a year. This decline confirms that current chatbots fail to meet user expectations of reliability and performance.

From Objective 2

$H_0$2: Users' perception towards chatbot effectiveness does not vary by type of chatbot. [(meanScore(Amazon Alexa)=meanScore(AppleSiri)=meanScore(Google Assistant)]

$H_1$2: Users' perception towards chatbot effectiveness varies by type of chatbot. [(meanScore(Amazon Alexa)#meanScore(AppleSiri)#meanScore(Google Assistant)]

Table 5: ANOVA Results

Variables	*Test*	*p*	*Remarks*
Dependent Variable – Score Independent Variable – Chatbots	ANOVA	0.938	> 0.05, accept null hypothesis

Source: JASP

Interpretation

From the above table 5, as the p-value is 0.938 greater than 0.05,accept null hypothesis and conclude that Users' perception is not influenced by type of chatbot (Stefanos Bonovas 1, 2023).

Customers rated Amazon Alexa, Google Assistant, and Siri similarly for user pleasure and chatbot efficacy, suggesting that platform or brand differences have little bearing on perceived performance. Customer perceptions are influenced more by performance expectations and trust than by brand identity, and competitive advantages are instead provided by elements such ecosystem integration, marketing, and value-added features (Joshi, 2021).

From Objective 3

$H_0$3: The preference of consumers towards choosing chatbots is same. [median(Amazon_Alexa)=median(Google_Assistant)=median(Siri)=median(Bixby)=median(IBM)]

$H_1$3: The preference of consumers towards choosing chatbots is not the same. [median(Amazon_Alexa)#median(Google_Assistant)#median(Siri)#median(Bixby) #median(IBM)]

Code Snippet (Source: R)

```
>friedmanTest(y=outcome, groups=treatment,blocks=participants)
```

(Nahm, 2015)

p-value = 2.2e-16

Interpretation

As the p-value is less than 0.05, reject null hypothesis which concludes that median of all the chatbots are different (Stefanos Bonovas 1, 2023).

POST-HOC Test

Code Snippet (Source :R)

```
>frdAllPairsMillerTest(y=outcome,groups=treatment,blocks=participants)
```

(Nahm, 2015)

1 2 3 4

2 1 - - -

3 1 1 - -

4 1.6e-08 1.2e-08 1.1e-09 -

5 < 2e-16 < 2e-16 < 2e-16 2.9e-07

Interpretation

As group-wise clear differences are visible,there were no discernible differences between the Google Assistant-Amazon Alexa, Siri-Amazon Alexa, Bixby-Amazon Alexa, and Siri-Google Assistant groups.

But Bixby was very different from Siri, Google Assistant, and Amazon Alexa.

When IBM was compared to all other chatbots (Amazon Alexa, Google Assistant, Siri, and Bixby), the differences were extremely noticeable.

User evaluations of efficacy are consistent across chatbot kinds, indicating that performance expectancy and trust are more important than brand labels (Joshi, 2021).

Bixby/IBM lags behind because of chatbots that have a bad communication style or a limited social presence, as demonstrated by (Na Cai, 2024).

From Objective 4

Gender on Users' perception on chatbots effectiveness (Score)

$H_0$4.1: Gender does not affect the Score. [(median Score(f)=medianScore(m)]

$H_1$4.1: Gender affects the Score. [medianScore(f) #medianScore(m)]

Code Snippet (Source :R)

```
> f<-subset(Score,Gender=='Female')
> m<-subset(Score,Gender=='Male')
> shapiro.test(f)
p-value = 0.002011
> shapiro.test(m)
p-value = 0.05369
```

As the p-value for independent variable Gender in which female is 0.002011 which is less than 0.05,conclude that data is not normal, proceed for Mann Whitney U Test (Nahm, 2015).

Table 6: Mann Witney U Test Results

Variable_name	*Test*	*p*	*Remarks*
Score	Mann-Whitney	0.347	> 0.05, accept null hypothesis

Source: JASP

Interpretation: From the above table 6, as the p-value is 0.347 greater than 0.05, accept null hypothesiswhich says that Gender does not affect the Score (Stefanos Bonovas 1, 2023). Users of both sexes have comparable opinions about chatbots. Companies can concentrate on creating and promoting gender-neutral services. Language and technology traits have a bigger influence on satisfaction than demographics (Kokošková, 2023).

Location on Users' Perception on Chatbots Effectiveness (Score)

$H_0$4.2: The Score does not vary by location. [(meanScore(Vijayawada)=meanScore(Guntur)= meanScore(Tenali)]

$H_1$4.2: The Score varies by location. [(meanScore (Vijayawada)#meanScore(Guntur)#meanScore (Tenali)]

Table 7: ANOVA Results

Variables	*Test*	*p*	*Remarks*
Dependent Variable – Score Independent Variable –Location	ANOVA	0.939	> 0.05, accept null hypothesis

Source: JASP

Interpretation: From the above Table 7, as the p-value is 0.939 greater than 0.05,acceptnull hypothesiswhich saysthat the users' perception does not vary by location (Malcolm Koo 1, 2025).Chatbots received similar ratings from users in Vijayawada, Guntur, Tenali, and other areas. Although local language integration may improve inclusivity, businesses can grow chatbots geographically without losing perception.Adoption is more influenced by trust and transparency than by geography or demographics (John, 2025).

Residential Status on Users' perception on chatbots effectiveness (Score).

$H_0$4.3: There is no impact of Residential Status on Score [medianScore(D)=medianScore(H)]

$H_1$4.3: There is no impact of Residential Status on Score [medianScore(D)#medianScore(H)]

Code Snippet (Source :R)

```
> D<-subset(Score,DH=='Dayscholar')
> H<-subset(Score,DH=='Hostler')
> shapiro.test(D)
p-value = 0.00126
> shapiro.test(H)
p-value = 0.05232
```

As one of the groups data which is Dayscholar p-values less than 0.05 reject null hypothesis and conclude that data is not normal, proceed with Mann Whitney U Test (Nahm, 2015).

Table 8: Mann Witney U Test Results

Variable_name	*Test*	*p*	*Remarks*
Score	Mann-Whitney	0.239	> 0.05, accept null hypothesis

Source: JASP

Interpretation: From the above table 8, as the p-value is greater than 0.05,acceptnull hypothesiswhich says that Residential Status has no influence on the perception of the users (Stefanos Bonovas 1, 2023).

There is no discernible difference between day students' and hostelers' opinions of chatbot effectiveness, suggesting that student demographics have adopted them consistently. This implies that rather than segment-specific modification, chatbot development can concentrate on functional perfection.Linguistic and technological characteristics, not demographic factors like living arrangement, are what influence user happiness with chatbots (Kokošková, 2023).

Choice of Chatbots onNo. of Hours Spent by the User (Time Spent -TI)

$H_0$4.4: The choice of chatbot does not affect the no. of hours spent by User. [(meanTimeSpent(Amazon Alexa)=meanTimeSpent(AppleSiri)=meanTime Spent(Google Assistant)]

$H_1$4.4: The choice of chatbot affects the no. of hours spent by User. [(meanTimeSpent(Amazon Alexa)#meanTimeSpent(AppleSiri)#meanTime Spent(Google Assistant)]

Code Snippet (Source :R)

Test Homogeneity:

H_0: Homogeneity exists

H_1: Heterogenity exists

>bartlett.test(TI~Chatbots,data=dt)

p-value = 7.152e-08

\#As the p-value is very much less than0.05,reject null hypothesis and hence conclude that heterogeneity exists (Stefanos Bonovas 1, 2023).

Independence of observations presumed.

Table 9: ANOVA Results

Variables	*Test*	*p*	*Remarks*
Dependent Variable- No. of Hours Spent by the User (Time Spent – TI) Independent Variable – Chatbots	ANOVA	0.328	> 0.05, accept null hypothesis

Source: JASP

Interpretation: From the above table 9, as the p-value is greater than 0.05,accept the null hypothesiswhich concludes that the choice of chatbot does not affect the no. of hours spent (Stefanos Bonovas 1, 2023). Individuals use Google Assistant, Siri, and Alexa for comparable periods of time, demonstrating functional parity in usage duration. Instead of investing effort in differentiating chatbot performance, businesses can concentrate on task fulfillment, engagement quality, and personalization.

Residential Status on No. of Hours Spent by the User (Time Spent – TI)

$H_0$4.5: Residential Status does not affect the no. of time spent [medianTimeSpent(D)=medianTimeSpent(H)]

$H_1$4.5: Residential Status affects the no. of time spent [medianTimeSpent(D)#medianTimeSpent(H)]

Code Snippet (Source :R)

```
>D<-subset(TI,DH=='Dayscholar')
> H<-subset(TI,DH=='Hostler')
> shapiro.test(D)
p-value = 9.678e-06
> shapiro.test(H)
p-value = 1.982e-08
> #As the p-values of groups are less than 0.05,data is not normal.
```

Proceed for Mann Whitney U Test (Nahm, 2015).

Table 10: Mann Witney U Test Results

Variables	*Test*	*p*	*Remarks*
Independent Variable – Residential Status Dependent Variable – No. of Hours Spent by the User (Time Spent – TI)	Mann-Whitney	0.923	> 0.05, accept null hypothesis

Source: JASP

Interpretation: From the above table 10, as the p-value is 0.923 which is greater than 0.05,acceptnull hypothesis which says that Residential Status does not have any effect on time spent (Stefanos Bonovas 1, 2023).

Students use chatbots in the same way regardless of their living arrangement, thus whether they are day students or hostelers has no bearing on how much time they spend using them.Human-like characteristics, trust, and interaction have a greater influence on chatbot adoption in e-commerce than demographic characteristics (Najaf & Ding, 2024).

Gender on No. of Hours Spent by the User (Time Spent – TI)

$H_0$4.6: Genderdoes not affectthe no. of hours time spent. [mediandTimeSpent(f)=medianTimeSpent(m)]

$H_1$4.6: Gender affects the no. of hours time spent. [mediandTimeSpent(f)#medianTimeSpent(m)]

Code Snippet (Source :R)

```
> f<-subset(TI,Gender=='Female')
> m<-subset(TI,Gender=='Male')
> shapiro.test(f)
p-value = 2.944e-11
> shapiro.test(m)
p-value = 0.00058
> #As the p-values of groups are less than 0.05, data is not normal, proceed for Mann Whitney U Test (Nahm, 2015).
```

Table 11: Mann Witney U Test Results

Variables	*Test*	*p*	*Remarks*
Independent Variable – Gender Dependent Variable – No. of Hours Spent by the User (Time Spent – TI)	Mann-Whitney	0.168	> 0.05, accept null hypothesis

Source: JASP

Interpretation: From the Table 11, as the p-value is greater than 0.05, accept null hypothesiswhich says that the timespent on the selected chatbot does not differ by gender (Stefanos Bonovas 1, 2023).Instead of gender-specific customisation, developers should focus on universal attributes (easiness, clarity,and trust). Trust and past experience outweigh demographic factors like gender (Kozak & Fel, 2024)

Location on No. of Hours Spent by the User (Time Spent – TI)

$H_0$4.7: Location does not make any difference on time spent with respective to chosen chatbot. [(meanTimeSpent(Vijayawada)=meanTimeSpent (Guntur)=meanTimeSpent(Tenali)]

$H_1$4.7: Location makes difference on time spent with respective to choosen chatbot. [(meanTimeSpent (Vijayawada)#meanTimeSpent(Guntur)#meanTime Spent(Tenali)]

Code Snippet (Source :R)

```
>#Ho: Homogenity exists
>#H1:Heterogenity exists
>bartlett.test(TI~Location,data=dt)
p-value = 0.004924
> #As the p-value is less than 0.05, reject null hypothesisand hence Heterogenity exists. (Stefanos Bonovas 1, 2023)
> #Independence of observations presumed.
```

Table 12: ANOVA Results

Variables	*Test*	*p*	*Remarks*
Independent Variable – Location Dependent Variable – No. of Hours Spent by the User (Time Spent – TI)	ANOVA	0.835	> 0.05, accept null hypothesis

Source: JASP

Interpretation: From the Table 12, as the p-value is 0.835which is greater than 0.05, accept the null hypothesis which concludes that Location does not make any difference on time spent with respective to selected chatbot (Stefanos Bonovas 1, 2023).

Chatbots were used for comparable amounts of time by students in various cities. This implies that chatbot adoption is stable across regions in comparable socio-cultural contexts, enabling developers to put functionality ahead of regional customisation. Adoption is more strongly influenced by trust and transparency than by geography (Colpan, 2024).

From Objective 5

Gender on Choice of Chatbots

$H_0$5.1: Users' Gender has no connection with their choice of chatbots.

$H_1$5.1: Users' Gender has no connection with their choice of chatbots.

Table 13: Contingency Table Results

Variables	*Test*	*p*	*Remarks*
Variable 1 – Gender Variable 2 – Chatbots	Chi-Squared Test (χ^2)	0.791	> 0.05, accept null hypothesis

Note: Continuity correction is available only for 2×2 tables.
Source: JASP

Interpretation: From the above table 16, as the p-value is 0.0791, which is greater than 0.05, accept null hypothesis which concludes that gender has no connection with the chatbot chosen by the users (Stefanos Bonovas 1, 2023).

Residential Status on Choice of Chatbots

$H_0$5.2: Choice of Chatbots does not depend on Users' Residential Status.

$H_1$5.2: Choice of Chatbots depends on Users' Residential Status.

Table 14: Contingency Table Results

Variables	*Test*	*p*	*Remarks*
Variable 1 – Dayscholar/ Hosteler Variable 2 – Chatbots	Chi-Squared Test (χ^2)	0.926	> 0.05, accept null hypothesis

Note: Continuity correction is available only for 2×2 tables.
Source: JASP

Interpretation: From the above table 17, as the p-value is 0.926, which is greater than 0.05, accept null hypothesis which says that chatbots choice does not depend on day scholar or hosteler (Stefanos Bonovas 1, 2023).

Chatbots were similarly selected by day scholars and hostel residents, demonstrating that brand preference is not influenced by residential setting. Chatbot adoption patterns are explained by trust and interaction rather than demographics (Najaf & Ding, 2024).

Location on Choice of Chatbots

$H_0$5.3: Users' Location and their Choice of Chatbots are not related.

$H_1$5.3: Users' Location and their Choice of Chatbots are related.

Table 15: Contingency Table Results

Variables	*Test*	*p*	*Remarks*
Variable 1 – Location Variable 2 – Chatbots	Chi-Squared Test (χ^2)	0.397	> 0.05, accept null hypothesis

Note: Continuity correction is available only for 2x2 tables.
Source: JASP

Interpretation: From the above table 18, as the p-value is 0.397 which is greater than 0.05accept null hypothesis, concludes that location and selection of chatbots are not related (Stefanos Bonovas 1, 2023).

Similar chatbot preferences were displayed by students in Vijayawada, Guntur, Tenali, and other locations, suggesting that brand adoption is consistent across geographic boundaries. Adoption is more strongly influenced by trust and transparency than by demographics based on geography (Colpan, 2024).

FINDINGS OF THE STUDY

1. Users' perception towards chatbots effective is not much positive.

2. Users' perception on chatbots effectiveness is same irrespective of the type of chatbots (For the best Chosen Chatbots from all the respondents).
3. Users' preferences towards choosing chatbots vary from those using all the chatbots in terms of efficacy, trust and communication style aspects.
4. Users' Gender, Location and their Residential Status are not affecting Chatbots Effectiveness as they have no role in the choice of those chatbots.
5. Users' Time spent on Chatbots does not depend on the choice of chatbots.
6. Even, User's demographics namely Gender, Location and their Residential Status have no role in their time spent on their chosen chatbots.

IMPLICATIONS OF THE STUDY

Companies can use these findings to improve chatbot functions and increase overall user satisfaction. Since Bixby and IBM Watson show lower user engagement compared to Google Assistant, Amazon Alexa, and Siri, these companies should focus on improving brand awareness and device integration. As user perceptions do not vary by gender or location, chatbot improvements should follow gender-neutral and location-neutral approaches.

FUTURE SCOPE OF RESEARCH

This paper provides a foundational overview of chatbots and users' perception towards voice activated chatbots. For further exploration, consider delving deeper (Yuan et al., 2024) (Yuan et al., 2024) into specific areas like the Impact of Voice-Activated Chatbots on Elders. Change the target audience and analyse the users' perception to obtain most robust results. Use other statements in Likert scale for more variables and analyse deeper.

REFERENCES

Adamopoulou, E. (2020). *Chatbots: History, technology, and applications* (Vol. 584). Springer.

Bertaria Sohnata Hutauruk, R. P. (2024). The Effectiveness of Artificial Intelligence by Chatbot in Enhancing the Students' Vocabulary. *ResearchGate*.

Colpan, E.W. (2024). Reconciling the tension between contextualization and generalization in qualitative international business research. *Journal of International Business Studies, 34*(2), 1–15.

Dale. (2016). The return of the chatbots. *Cambridge University Press, 22*(5), 811–817.

Daniel Maxwell, B. O. (2025). Generative AI in Higher Education: Demographic Differences in Student Perceived Readiness, Benefits, and Challenges. *Springer Nature*.

Darda, P. (2020). Voice Assistant: A Systematic Literature Review. *ResearchGate*, 1–8.

Dutta, S. (2024). *Chatbot Effectiveness in Enhancing Guest Communication: Insights from Secondary Data.* Tuijin Jishu.

Escobar-Grisales, D. (2023). Evaluation of effectiveness in conversations between humans and chatbots using parallel convolutional neural networks with multiple temporal resolutions. *Springer Nature*.

Ghazala Bilquise, S. I. (2022). Emotionally Intelligent Chatbots: A Systematic Literature Review. *Human Behavior and Emerging Technologies*, 1-23.

Gomes, S. L. (2025). Anthropomorphism in artificial intelligence: a game-changer for brand marketing. *Volume 11, article number 2*.

Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Taylor & Francis, 37*(1), 81–88.

Hussain, K. (2025). Humanizing generative AI Brands: How brand anthropomorphism converts customers into brand heroes. *ScienceDirect, Volume 19*.

Hutauruk. (2024). *The effectiveness of artificial intelligence by chatbot in enhancing the students' vocabulary* (Vol. 6). Journal of English Teaching and Applied Linguistics.

Jalawadai, M. (2024). Chatbots: Conversational AI Shaping Our World. *International Journal of Creative Research Thoughts(IJCRT), 12*(4), 142–146.

John, B. (2025). The Role of Trust in AI-Powered Chatbots: Impact on User Adoption and Retention. *Computers in Human Behavior* .

Joshi, H. (2021). Perception and Adoption of Customer Service Chatbots among. *Proceedings of the 17th International Conference on Web Information Systems and Technologies - Volume 1: WEBIST, 10*(3).

Jyosthna, M., Venkata Subbaiah, P., & Natalia, K. (2024). Exploring the Chatbot usage intention-A mediating role of Chatbot initial trust. *Research Gate*.

Kikani, D., & Ramchandani, S. M. (2025). *The Impact of AI Chatbots on Customer Service: Efficiency VS Human Touch* (Vols. volume10, Issue3). IJRTI.

Kokošková, J. B.-M. (2023). Integrating chatbots in education: insights from the Chatbot-Human Interaction Satisfaction Model (CHISM). *SpringerOpen, 20*(62), 1-14.

Kozak, J., & Fel, S. (2024). How sociodemographic factors relate to trust in artificial intelligence among students in Poland and the United Kingdom. *Scientific Reports, 14*(28776), 28776.

Lopatovska, I. (2018). Personification of the Amazon Alexa: BFF or a Mindless Companion. *ACM Association for Computing Machinery corporate, 55*(1), 1–10.

Magno, F. (2023). *The effects of chatbots' attributes on customer relationships with brands: PLS-SEM and importance–performance map analysis* (Vols. Volume 35, Issue 5). Emerald Publishing.

Mahima Anna Varghese1, O. I., Poonam Sharma1, O. I., & Maitreyee Patwardhan2, O. I. (2024). *Public Perception on Artificial Intelligence–Driven Mental Health Interventions: Survey Research* (Vol. 8). JMIR Publications.

Malcolm Koo 1, 2. a.-W. (2025). Likert-Type Scale. *SN Comprehensive Clinical Medicine*.

María José Blanca, R.A. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29*(4), 552–557.

Martin Adam, M. W. (2020). AI-based chatbots in customer service and their effects on user compliance, (pp. 1–6). Association for Information Systems.

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica, 23*(2), 143–149.

Murray, S. (2024). *GenAI and Voice Assistants: Adoption and Trust Across Generations.* Deloitte.

Na Cai, S. G. (2024). How the communication style of chatbots influences consumers' satisfaction, trust, and engagement in the context of service failure. *Computers in Human Behavior, 11*(687), 1–15.

Nahm, F. S. (2015). Nonparametric statistical tests for the continuous data: the basic concept and the practical use. *Korean Journal of Anesthesiology, 68*(1), 8–14.

Najaf, M., & Ding, Y. (2024). *Interactivity, humanness, and trust: a psychological approach to AI chatbot adoption in e-commerce* (Vol. 146). Elsevier.

Nunnally, J. C. (2016). *Understanding Coefficient Alpha: Assumptions and Interpretations.* McGraw-Hill.

Pušnik, M. (2025). Differences in User Perception of Artificial Intelligence-Driven Chatbots and Traditional Tools in Qualitative Data Analysis. *Computers in Human Behavior Reports*.

Reddy, N. T. (2025). A Modular Retrieval-Augmented Conversational AI Chatbot System with Integrated Recommender Engine Using Local LLMs. *2*.

Singh. (2014). Sampling techniques and determination of sample size in applied statistics research: An overview. *International Journal of Economics, Commerce and Management, 2*(11), 1–7.

Singh, A.S. (2014). *Sampling Techniques & Determination of Sample* (Vols. vol. II, Issue 11). International Journal of Economics, Commerce and Management.

Stefanos Bonovas 1, 2. a. (2023). On p-Values and Statistical Significance. *Hippokratia, 27*(1), 1–4.

Stöhr, C. (2024). *Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study.* Computers and Education: Artificial Intelligence.

Sziron, M. (2022). Investigating user perceptions of commercial virtual assistants: A qualitative study. *SN Computer Science, 3*(6), 1–12.

Trupti Rathi1, D.B. (2022). *Questionnaire as a Tool of Data Collection in Empirical Research* (Vol. 6). Journal of Positive School Psychology.

Xu, Y., Zhang, J., Chi, R., & Deng, G. (2022). Enhancing customer satisfaction with chatbots: The influence of anthropomorphic communication styles and anthropomorphised roles. Volume 14 (Issue 2).

Ying Xu, J. Z. (2022). Enhancing customer satisfaction with chatbots: The influence of communication styles and consumer attachment anxiety (pp. 1-10). Association for Information Systems.

Yuan, L., Sun, T., Dennis, A., & Zho, M. (2024). *Perception is reality? Understanding user perceptions of chatbot-inferred versus self-reported personality traits.* Computers in Human Behavior: Artificial Humans.

ANNEXURE – QUESTIONNAIRE

Questionnaire on Voice-Activated Chatbots

1. Gender
   (a) Male
   (b) Female
2. Location
   (a) Vijayawada
   (b) Guntur
   (c) Tenali
   (d) Others
3. Provide your Residential Status
   (a) Dayscholar
   (b) Hosteller
4. Number of hours spent on these voice-activated chatbots. (e.g.: 1 hour spent = 1) ...
5. From the following voice-activated chatbots, choose the best option.
   (a) Amazon Alexa
   (b) Apple Siri
   (c) Google Assistant
   (d) Samsung's Voice Assistant (Bixby)
   (e) IBM Watson Assistant
6. Provide your agreement with the following statements that focus on usage of the chosen product that measures Chatbots Effectiveness.

Likert Scale used in Questionnaire.

S. No.	*Statements*	*1*	*2*	*3*	*4*	*5*
1	More voice clarity					
2	Takes less instructions					
3	Takes less response time					
4	Provides more accurate results					
5	Validates or accepts maximum results					
6	Can easily operate other devices					
7	Can perform multitask routines					
8	Compatible with every device					
9	It remembers instructions given					
10	Hands-free calling and messaging					
11	Sets reminders effortlessly					
12	Schedules meetings effortlessly					
13	Gets directions and traffic live reports					
14	Online shopping and orders					
15	Health and fitness assistance.					
16	Helps in language translation.					

1 – Strongly Disagree; 2 – Disagree; 3 – Neutral; 4 – Agree; 5 – Strongly Agree.

7. Rank the following voice activated chatbots.
   (a) Amazon Alexa
   (b) Apple Siri
   (c) Google Assistant
   (d) Samsung's Voice Assistant (Bixby)
   (e) IBM Watson Assistant

Thank you for sparing your valuable time in filling this survey form.