Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 19 No. 14 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijim.v19i14.56997

SPECIAL FOCUS PAPER

From Touch to Talk: Transforming Mobile Interactions with Voice-Based Artificial Intelligence

Mehak Mittal (✉),
Sanjay Manocha

Bharati Vidyapeeth (Deemed to be University), Pune Institute of Management and Research, New Delhi, India

mehak15597@gmail.com

ABSTRACT

The rapid evolution of mobile technology has commenced in a new era of interaction paradigms, with "voice-based artificial intelligence (VAI)" emerging as a major disruptive influential force in user engagement. This study explores the factors that influence the adoption of voice-based AI on smartphones among Indian consumers. Using a quantitative research design, data were gathered from 500 Indian mobile users who have prior experience with voice assistants. The study is based on constructs from the "technology acceptance model (TAM)" and related theories to examine the influence of "Perceived Usefulness", "Perceived Ease of Use", "Perceived Enjoyment", "Social Influence", "Trust", and "Perceived Privacy Risk" on "Attitude to Use" and "Actual Usage of voice-based AI". To determine the measurement and structural aspects of the model, the dataset was tested using "Partial Least Squares Structural Equation Modelling (PLS-SEM)". Based on the findings, "Perceived Usefulness", "Perceived Ease of Use", "Perceived Enjoyment", and "Trust", and "Social Influence" positively influence users' "Attitude to Use", which has a positive effect on "Actual Usage". "Perceived Privacy Risk" negatively affected Attitude to Use. This study enhances the existing literature by focusing on the primary reasons for adoption in India. The results provide marketers and developers with insightful information. It implies that stakeholders may promote wider adoption and use of voice-based AI technology by building confidence, improving user experience, and resolving privacy concerns.

KEYWORDS

voice-based AI, mobile technology, technology acceptance model (TAM), partial least squares structural equation modelling (PLS-SEM), India, technology adoption, voice assistants

1 INTRODUCTION

Mobile technology has transformed from traditional input methods to modern methods which are more natural and seamless. These new methods have brought a change in people's interactions with digital systems. "voice-based artificial

Mittal, M., Manocha, S. (2025). From Touch to Talk: Transforming Mobile Interactions with Voice-Based Artificial Intelligence. *International Journal of Interactive Mobile Technologies (iJIM)*, 19(14), pp. 42–56. https://doi.org/10.3991/ijim.v19i14.56997

This article is a revised version of a paper presented at the Annual Conference on Innovating for Impact, held on February 19–20, 2025, at the Manipal Dubai Campus, Dubai, UAE. Article submitted 2025-05-03. Revision uploaded 2025-05-28. Final acceptance 2025-06-03.

intelligence (AI)" is a very significant revolution in human-computer interactions. Voice assistants such as "Apple Siri", "Google Assistant", "Amazon Alexa", and "Samsung Bixby" enable users to manage their devices via natural spoken language. Voice assistants have revolutionised mobile interactions and changed how consumers use their smartphones. Sending messages, setting alarms and reminders, planning daily activities, finding information, and even controlling smart home devices with voice commands are some of the primary purposes for which users use voice assistants [1], [2].

Voice assistants are intelligent AI-powered software devices, which use "speech recognition", "machine learning", and "natural language processing". Voice assistants facilitate natural and intuitive communication between humans and computer devices. They are built in such ways that they can comprehend what users are saying and give accurate, reliable responses by proper processing. Voice assistants are a blessing in circumstances where typing or tapping is not practical and unsafe [3] and in countries like India where users consider their smartphones as the primary means of accessing digital services, be it big business man or students.

Voice assistants are primarily helpful in situations of multi-tasking like cooking and driving. People with disabilities can also greatly benefit from such technologies. Voice assistants are becoming more accessible due to the growing support for local regional languages and dialects by the product developers. Such developments play a crucial role in adoption for India's diverse multilingual population [4].

Cloud computing, big data analytics, and intelligent algorithms are some of the technologies that have contributed to the rise of voice-based AI. This has led to more precise, customised, personalised, and organic human-computer interactions. Even with such radical transformation and high rates of adoption, some people still hesitate to adopt voice AI in India. Such users are sceptical about various things like privacy, data security, and cultural nuances [5]. This difference in adoption across different regions demonstrates the importance of studying the factors influencing the adoption of voice assistants across a multicultural country like India.

Researchers generally rely on the "technology acceptance model (TAM)" for studying and understanding how individuals accept new technologies. According to this model, if users find technology to be useful, easy to use and beneficial, people will use it. This thought process creates a positive attitude in consumer minds, leading to actual usage [6]. Experts and past researchers have advised including more factors in the traditional TAM, though, due to the complexity and interactivity of AI technologies [6]–[8].

When studying the Indian population, it becomes more important to add additional elements in the TAM model due to the nation's sociocultural dynamics, usage of approximately over 22 languages across the nation, expanding digital reach, and rising awareness about data privacy and trust [9].

One of the most important drivers is "Perceived Usefulness," which is "how much users believe voice assistants help them save time and make tasks easier by enhancing their productivity and convenience". In India, where mobile phones are the main gateway to the internet, voice assistants can be especially helpful in situations like cooking or traveling, where using hands is difficult [6], [10]. Another major factor is "Perceived Ease of Use" i.e., "people are more likely to try and keep using voice assistants if they find them simple and user-friendly". The ability to interact using natural spoken language makes voice assistants easier to use, especially for those who may struggle with traditional touchscreen interfaces and people with certain disabilities [10].

One key factor that encourages people to adopt voice assistants is "Perceived Enjoyment" they get from using them [10]. Many users find it fun and satisfying to interact with AI through voice, which helps build a positive user attitude toward continued use. "Social Influence" also plays an important role, especially in a collectivist country like India, where people often value the opinions and behaviours of those around them. When friends, family, or peers use and recommend voice assistants, it creates a favourable atmosphere for others to follow and leads to technology adoption [9]. "Trust" is another major factor which recommends people are more likely to use voice assistants if they believe the technology is dependable, respects their privacy, and gives accurate, reliable results. As awareness around digital safety grows, voice assistants that are seen and perceived as secure and non-intrusive are more likely to be accepted [10].

On the other hand, certain concerns hold users back from adopting voice assistants widely in India. One of the biggest concerns is privacy. Many users worry about how their voice data is being recorded, stored, or possibly misused [11]. With data protection laws and guidelines still in the development phase in India, apprehensions about surveillance, data leakage, and unauthorised access remain high. Another challenge is the technology's difficulty in handling different Indian languages and accents, which can limit its usability for many users.

By examining both enablers and barriers to adoption, this study provides a comprehensive understanding of the psychological, social, and local influences shaping how Indian users engage, adopt and accept voice-based AI on their smartphones. It also proposes a framework that connects these factors to better understand their influence on users' attitudes and actual use of voice assistants.

1.1 Need and significance of the study

Given the accelerating penetration of smartphones and the increasing integration of AI-driven services into mobile platforms, it is vital to understand what drives or inhibits the use of VA in India. While prior research tried to cater an understanding about voice AI adoption in Western settings, relatively few empirical studies focus on emerging economies with complex user profiles such as India. This study helps bridge the gap by examining how people adopt voice-based AI on their smartphones using an expanded version of the "Technology Acceptance Model (TAM)", which includes factors like "Trust", "Perceived Privacy Risk", "Perceived Enjoyment", and "Social Influence" in addition to the original elements of the model.

Recent literature has also expanded the scope by incorporating constructs like Trust, Perceived Enjoyment, and Privacy Concerns to better capture user motivations and apprehensions in engaging with voice-enabled technologies. However, much of this study has been carried out in Western contexts, where technology usage patterns, digital literacy, and privacy norms differ significantly from those in developing countries like India. Moreover, while there is growing interest in voice assistants, few empirical studies have specifically examined their adoption among Indian smartphone users, particularly within a mobile-first ecosystem marked by linguistic diversity and variable internet access [12], [13]. Existing studies often overlook the interplay between psychological drivers and contextual factors unique to India.

This creates a need for a research that includes both functional and socio-cultural dimensions influencing technology adoption in India.

RQ: "What are the key psychological and contextual factors influencing the adoption and actual usage of VAI on mobile devices among Indian consumers?"

2 MATERIALS AND METHODS

The rapid development of voice-based AIs has received attention from scholars in relation to its acceptance, effect, and usability. The TAM, created by Davis [6], has been integral in form of a framework attempting to explain user behaviour towards new technologies. It advocates that "Perceived Usefulness (PU)" and "Perceived Ease of Use (PEOU)" are the leading determinants of intention. Other models, like UTAUT by Venkatesh et al. [7], added "Social Influence" and "Facilitating Conditions", further extending these frameworks in previously unexplored contexts.

In regard to voice assistants, Hoy [2] examined the historical development of voice assistants and underscored the importance of recognition and natural language processing (NLP) technologies for voice interfaces.

Gupta and Arora [14] noted that the adoption of voice assistants within urban Indian settings tends to be influenced by social circles, supporting the notion of sociological diffusion. Adoption continues to be influenced by PU. Pradhan et al. [15] stated in a study conducted among Indian digital users that users found voice assistants to be adopted only for basic and primary functions.

Lee and Choi [16] studies the association between usability of voice commands and the benefits users perceive that they get from using voice commands. Researchers highlighted the direct proportion between the accuracy of the assistant's understanding and execution user queries and user satisfaction. These findings support the fact that users consider the effectiveness and productivity of voice interfaces in their adoption decisions.

Factors like "Perceived Ease of Use (PEOU)" also play an important role. McLean and Osei-Frimpong [10] noted that users are likely to adopt more advanced voice assistants with simple yet intuitive interfaces which require minimal energy to use. This is critical in mobile infrastructures where users are looking for quick on-the-go solutions. Apart from functional properties, emotional context is becoming increasingly relevant. There are several studies, including that by Zhou et al. [17], which define "Perceived Enjoyment" as "the intrinsic delight discovered in using the technology". They emphasise it as one of the distinct reasons for technology adoption and continuous usage. This explains how enjoyment increases user experience, especially in young generations and children.

The trust relationship between users and the AI systems stands as one more major thing to be considered. Lau et al. [18] found that when companies are transparent about how they handle user data and have a strong brand reputation, users are more likely to trust their voice assistants. Gefen et al. [19] also emphasised that trust helps reduce users' uncertainty and concerns about risk, which allows users to become more open to using voice-based technologies. Lutz and Newlands [20] added that people often worry about how their voice data is collected and stored, which directly influences their level of trust in the voice assistants. Past researchers have underlined the importance of building trust in the eyes of consumers and users. "Trust" helps to attract new customers and retain the current ones.

One of the main obstacles to the adoption of technology is “Perceived Privacy Risk.” People worry about data security and privacy. Such worries and apprehensions are particularly prevalent in societies with low levels of technological infrastructure and low awareness about remedies for keeping data safe, according to Lau et al. [18]. This is supported by other researchers also, like by Dhamija and Yadav [21], who underlined that many Indians refrain from adopting voice assistants for financial or personal tasks due to privacy concerns and the possibility of money theft. They believe that voice assistants will steal their credit/debit cards or UPI information and allow scammers and thieves to steal money from their bank accounts.

“Social influence” is another important factor in technology adoption, particularly in cultures that emphasise group behaviour and collectivist societies like India. Venkatesh and Bala [22] found that people’s decisions about using technology are often shaped by opinions of their friends and family and social media advertisements. Kapoor et al. [23] showed that recommendations and seeing others use voice assistants in public increase the chances of adoption.

For collectivist societies like India, “Social Influence” proves to be another significant driver in the adoption of technology. Venkatesh and Bala [22] discovered that social media advertisements and the opinions of friends and family are really important to users and influence their decision-making process when technology adoption is in question. It is found by Kapoor et al. [23] that word-of-mouth recommendations and testimonials can greatly motivate people in the positive direction. Thus, it can be concluded that word-of-mouth recommendations and public demonstrations of voice assistant usage contribute to greater acceptance. Similarly, Gupta and Arora [24] observed that in Indian metropolitan cities, people often follow trends set by tech-savvy peers, showing how group behaviour can drive technology use.

Thus, it is concluded that while voice assistants have strong potential for mainstream adoption, there remains a gap in understanding how diverse user segments, particularly in developing economies like India, perceive and interact with these technologies. They advocated for more localised and user-centred research to bridge this gap.

2.1 Hypotheses Development

Based on the literature review done and theoretical framework proposed as depicted in Figure 1, the following hypotheses are developed and used for further testing:

- H1: “Perceived Usefulness has a positive effect on Attitude to Use voice assistants”.
- H2: “Perceived Ease of Use has a positive effect on Attitude to Use voice assistants”.
- H3: “Perceived Enjoyment has a positive effect on Attitude to Use voice assistants”.
- H4: “Trust has a positive effect on Attitude to Use voice assistants”.
- H5: “Social influence has a positive effect on Attitude to Use voice assistants”.
- H6: “Perceived Privacy Risk has a negative effect on Attitude to Use voice assistants”.
- H7: “Attitude to Use has a positive effect on Actual Usage of voice assistants”

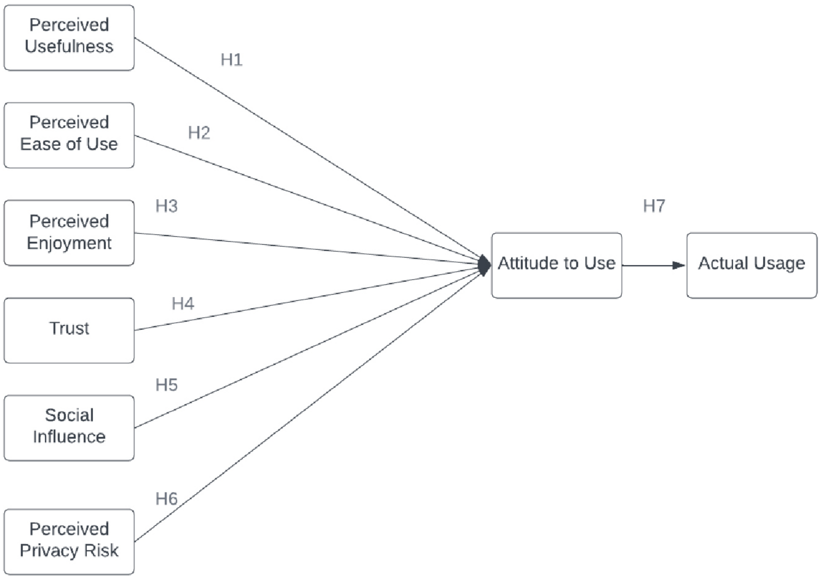

Fig. 1. Proposed framework

2.2 Research Methodology

The study is quantitative and cross-sectional in nature. Major independent factors of this study are "Perceived Usefulness", "Perceived Ease of Use", "Perceived Enjoyment", "Trust", "Social Influence", and "Perceived Privacy Risk".

The sampling technique was non-probability convenience sampling. Data was collected from smartphone users (had prior experience of using voice assistants) across five metropolitan cities in India. There was no well-defined sampling frame for this kind of study. Digital channels like social media, emails, and messaging applications were used to collect responses from the 500 target respondents.

The questionnaire used for data collection was framed on adapted statements from past studies to ensure reliability and validity. Each construct had five statements, making a total of 40 items. 5 point Likert scale was used, ranging from "Strongly Disagree" to "Strongly Agree". Pre and pilot testing of the questionnaire helped remove ambiguous items and added to reliability and validity. Final Data collection took place over four months, from January to April 2025. The formula used in sample size selection has been showed in Table 1.

Table 1. Criteria for determining the sample size

Name of the Researcher	Formula Given by the Researcher/Technique	Sample Size Recommended as per the Criteria
Hair et al., 1999	25 multiplied by the number of factors	8 * 25 = Minimum sample size of 200

The study utilised "Partial Least Squares Structural Equation Modelling (PLS-SEM)" for data analysis. This advanced multivariate technique is well-suited for examining complex models involving several latent variables and interrelated paths. PLS-SEM was used to evaluate both the measurement model and the structural model. The measurement model analysis was focused on confirming the reliability and validity of the research instrument, while the structural model tested the study's hypotheses. The model assessment followed established guidelines, including checks for convergent validity, discriminant validity, composite reliability, and the statistical significance of the path relationships. For data analysis including both measurement and structural model, "Partial Least Squares Structural Equation Modelling (PLS-SEM)" was used.

3 RESULTS

Tables 2 and 3 list demographic and consumer behaviour factors.

Table 2. Demographic characteristics

Demographics		Total (N = 500)	%
Gender	Male	365	73
	Female	135	27
Age	Below 25	93	18.6
	25–34	147	29.4
	35–44	70	14
	45–54	118	23.6
	55 and above	72	14.4
Highest Educational Qualification	High School	132	26.4
	Graduate	140	28
	Post-graduate	223	44.6
	Doctorate	5	1
Experience of VA usage on mobile devices	Less than 6 months	23	4.6
	6–12 months	104	20.8
	1–2 years	267	53.4
	More than 2 years	106	21.2
Location	National Capital Region (NCR)	146	29.2
	Mumbai	80	16
	Kolkata	70	14
	Chennai	87	17.4
	Bangalore	117	23.4

Table 3. Consumer behaviour characteristics

Characteristics		Total (N = 500)	%
Frequency of using VA	Rarely (Less than once a week)	17	3.4
	Occasionally (1–2 times a week)	102	20.4
	Frequently (3–5 times a week)	180	36
	Daily	201	40.2
VA Preference	Google Assistant	150	30
	Amazon Alexa	174	34.8
	Apple Siri	130	26
	Samsung Bixby	46	9.2

(Continued)

Table 3. Consumer behaviour characteristics *(Continued)*

Characteristics		Total (N = 500)	%
Reasons for using VA	Sending texts or making calls	470	94
	Searching the internet/news	345	69
	Setting alarms or reminders	365	73
	Playing music or videos	409	81.8
	Navigating or checking maps	230	46
	Smart home control	120	24

3.1 Assessment of Multivariate Normality

Prior to choosing a "structural equation modelling (SEM) method", it was crucial to confirm that the dataset is multivariate normal or not. "Covariance-based SEM (CB-SEM)" relies on the assumption that the variables have a multivariate normal distribution. To assess this, Mardia's coefficients for skewness and kurtosis were computed using the Web-Power tool [25] to evaluate the multivariate normality of the data. The analysis indicated that the data did not follow a normal distribution. As a result, PLS-SEM was chosen, as it can handle non-normal data and is effective in predictive modelling and theory development.

3.2 Common Method Bias (CMB) Assessment

Since the study used a self-administered questionnaire based on Likert scales to collect responses on various latent constructs, there was a possibility of CMB. CMB refers to a type of measurement error that stems from the data collection method itself rather than genuine associations and relationships between variables/constructs [26]. To mitigate this issue, a full collinearity test was conducted following the procedure suggested by Kock [27]. This method involves examining the inner variance inflation factor (VIF) values for all constructs. If these values are below the benchmark value of 3.3, the model is considered to be free from significant common method bias. For this each construct was tested against a randomly generated variable, and all inner VIF values were found to be below the 3.3 threshold, indicating that the results were not influenced and deteriorated by CMB [28].

3.3 Reliability and Validity of the Questionnaire Used

For checking the "Reliability and Validity" of the questionnaire, "internal consistency", "convergent validity", and "discriminant validity" were checked. Cronbach's alpha and composite reliability (CR) were computed to assess internal consistency. Cronbach's alpha and composite reliability values for each construct were more than 0.70 [29], which is above the threshold value. Convergent validity was measured using the Average Variance Extracted (AVE). For all constructs, AVE values were greater than 0.50. This supported the presence of convergent validity [29]. A summary of the reliability and convergent validity results is provided in Table 4.

To check whether the different constructs in the study were clearly separate from one another, discriminant validity was tested using two widely accepted methods: the "Fornell-Larcker criterion" and the "Heterotrait-Monotrait (HTMT) ratio."

First, the HTMT values for all pairs of constructs were found to be below 0.85. This means that the constructs are not too closely related and are measuring different things, as shown in Table 5. Second, the "Fornell-Larcker method" was applied. In every case, the square root of the AVE was higher than the corresponding correlations. These results, displayed in Table 6, confirm that the measurement model has strong discriminant validity [30], [31].

Table 4. Reliability and convergent validity

	Cronbach's Alpha	Composite Reliability (rho_a)	Composite Reliability (rho_c)	Average Variance Extracted (AVE)
ATT	0.902	0.903	0.928	0.72
PE	0.893	0.893	0.921	0.7
PEOU	0.8	0.803	0.882	0.714
PPR	0.717	0.729	0.823	0.538
PU	0.89	0.891	0.919	0.694
SI	0.911	0.915	0.934	0.739
TR	0.864	0.869	0.902	0.647
USE	0.836	0.839	0.891	0.671

Table 5. Heterotrait-Monotrait (HTMT) ratio

	ATT	PE	PEOU	PPR	PU	SI	TR	USE
ATT								
PE	0.728							
PEOU	0.596	0.618						
PPR	0.438	0.191	0.246					
PU	0.61	0.613	0.664	0.218				
SI	0.499	0.435	0.321	0.188	0.468			
TR	0.548	0.449	0.411	0.138	0.445	0.44		
USE	0.563	0.565	0.553	0.169	0.787	0.57	0.456	

Table 6. Fornell-Larcker criterion

	ATT	PE	PEOU	PPR	PU	SI	TR	USE
ATT	0.848							
PE	0.654	0.837						
PEOU	0.508	0.526	0.845					
PPR	–0.36	–0.161	–0.196	0.734				
PU	0.548	0.548	0.562	–0.181	0.833			
SI	0.453	0.392	0.275	–0.159	0.421	0.86		
TR	0.486	0.396	0.342	–0.113	0.392	0.39	0.804	
USE	0.49	0.49	0.453	–0.139	0.68	0.5	0.392	0.819

3.4 Structural Model Assessment

The structural model was assessed by analysing the size and statistical relevance of the path coefficients. To evaluate the significance of these relationships, a bootstrapping technique with 5,000 resamples was employed. Results are summarised in Table 7.

Table 7. Path coefficients

Direct Effects	Coefficient	T Statistics	p Values
Path			
Attitude –> Actual Usage	0.49	12.786	0.000
Perceived Enjoyment –> Attitude	0.38	7.649	0.000
Perceived Ease of Use –> Attitude	0.102	2.608	0.005
Perceived Privacy Risk –> Attitude	–0.218	7.591	0.000
Perceived Usefulness –> Attitude	0.122	2.913	0.002
Social Influence –> Attitude	0.119	3.255	0.001
Trust –> Attitude	0.182	4.781	0.000

Table 8. Explanatory power, predictive power and model fit

Explanatory Power: R Square	R Square	R Square Adjusted
Attitude	0.74	0.72
Actual Usage	0.61	0.60
Predictive Power: Q^2		
Attitude	0.571	
Actual Usage	0.323	
Effect Size: f Square		
Attitude –> Actual Usage	0.316	
Perceived Enjoyment –> Attitude	0.206	
Perceived Ease of Use –> Attitude	0.015	
Perceived Privacy Risk –> Attitude	0.109	
Perceived Usefulness –> Attitude	0.02	
Social Influence –> Attitude	0.025	
Trust –> Attitude	0.06	
Model Fit		
SRMR	0.068	

The analysis showed that all the factors in the model had a significant influence on users' attitudes toward using voice-based AI on mobile phones because all the path coefficients were found to be significant. This means that elements like how useful, easy, enjoyable, and trustworthy people find the technology, along with social pressure, all play a positive role in shaping their attitude. Privacy concerns, on the other hand, had a negative effect. Additionally, the study results highlighted that a more positive attitude toward the technology leads to greater actual use of technology.

Predictive relevance was examined through Q^2 values calculated using the PLSPredict technique. The Q^2 values for “Actual Usage” and “Attitude to Use” exceeded 0.25 and 0.50, respectively, indicating moderate to high predictive relevance. According to guidelines, Q^2 values greater than 0, 0.25, and 0.50 reflect low, moderate, and high predictive accuracy, respectively [32]. These findings affirm the robustness of the model in terms of explanatory and predictive performance.

Table 8 highlights key results showing how well the proposed model explains and predicts consumer adoption behaviour. The R^2 values show how much of the change in user attitude and actual use of voice-based AI can be explained by the factors in the model. The adjusted R^2 values, which give a more accurate picture as they consider the number of predictors used, are also mentioned. To understand how much each predictor factor individually contributes to the outcome factor, effect sizes (f^2) were also calculated.

The model’s overall fit was checked using the standardised root mean square residual (SRMR), which came out to 0.068. Since values under 0.08 are seen as acceptable, the proposed model is a good fit [32]. The predictive ability of the model was measured using Q^2 values through the PLSPredict method. The Q^2 values for both “Actual Usage” and “Attitude to Use” were above 0.25 and 0.50, showing that the model has moderate to strong prediction power. According to standards, Q^2 values above 0 show low prediction accuracy, values above 0.25 show moderate accuracy, and values above 0.50 indicate high accuracy.

4 DISCUSSION

This study offers important valuable insights about the factors which drive/hinder Indian smartphone users to adopt voice-based AI on their mobile devices. Using an extended version of the TAM, the research examined how various factors “Perceived Usefulness”, “Perceived Ease of Use”, “Trust”, “Perceived Enjoyment”, “Social Influence”, “Perceived Privacy Risk”, and their influence on “Attitude to Use” further influencing “Actual Usage” of voice assistants.

“Perceived Usefulness” was found to be a key factor influencing users’ attitudes. This supports earlier findings that users are more likely to adopt any technology if they believe that adoption of a given technology will help them work more efficiently or save their time [6]. Voice assistants, by enabling hands-free interactions and quick access to information, align well with the modern user’s demand for convenience and functionality in this busy and tech-savvy world. Also, “Perceived Ease of Use” had a strong positive influence on users’ attitudes. It depicts that people prefer using technologies that are easy to use [7], [10].

“Trust and Perceived Enjoyment” also had a positive influence in driving adoption. Trust is especially important in AI applications, where users often share personal information and rely on the system for information retrieval and to respond correctly. When people feel confident about their data safety and feel that the system works reliably, they are more likely to use it. “Perceived Enjoyment” reflects intrinsic motivation; the more pleasurable the experience, the higher the likelihood of repeated usage, which aligns with the findings from prior studies [10]. “Social Influence” had a positive influence as well, showing that recommendations from friends or popular trends can affect people’s willingness to adopt new technology

[17], [23]. In a low-context, culturally collectivist country like India, where community opinions often guide personal decisions, social influence proves to be an important factor when studying consumer behaviours.

On the other hand, "Perceived Privacy Risk" has a negative effect on user attitudes, aligning with previous research suggesting that privacy concerns about personal data collection, storage and usage remain a major barrier in the adoption of AI technologies [10], [12]. Users are concerned about data breaches, surveillance, and misuse of their personal information, especially private and financial data, which inhibit their willingness to adopt voice assistants despite VA's functional benefits. To summarise, this study contributes to the limited body of empirical work focused on voice AI adoption in India. The findings of this study have both theoretical and practical implications. From a theoretical point, the results support the extended TAM model while highlighting the need for the inclusion of contextual and social factors. Practically, product developers, marketers and business companies of voice-based AI systems should focus on making their systems easy to use, enjoyable, and trustworthy while also clearly addressing users' privacy concerns through measures like transparent data policies.

5 CONCLUSION

This paper study had an objective to examine the key determinants influencing the adoption of VAI on mobile devices among Indian users, using an extended version of TAM framework. The results highlight that people are more likely to use voice assistants when they find them useful, easy to use, enjoyable, and trustworthy. Users should have less or minimal privacy concerns, as privacy concerns still act as a major barrier, highlighting the need for secure and transparent AI systems.

The study is one of the major contributions to the existing research on voice AI, especially in developing and mobile-first countries like India, where rapid digital transformation is shaping new user behaviours. By extending TAM with the inclusion of emotional and contextual factors, the research offers a more holistic and deeper understanding of adoption behaviour toward VA systems. The moderately high explanatory and predictive power of the structural model also validates the robustness of the proposed framework for similar future technology adoption studies.

However, there are several limitations that must be acknowledged. First, the study used a cross-sectional design, which doesn't allow for inferring causality and change in behaviour over time. Longitudinal studies may help to provide a more holistic understanding and better insights of user behaviour and changing perceptions in rapidly changing environments. Second, data was collected through self-reported questionnaires, which may have introduced certain biases despite robust statistical checks. Third, the study focused only on urban smartphone users in metropolitan cities, which may not fully capture the adoption patterns in semi-urban or rural regions of India, as the access, attitudes and experiences toward technology might differ significantly across different regions.

In order to overcome these limitations, future studies should use longitudinal designs. Researchers may also perform a comparative study between urban and rural smartphone users across India or different countries. Also, qualitative methods like focus groups and interviews could help in better and deeper understanding. For better insights, researchers may also think about including several additional elements like user habits and preferred languages.

In conclusion, the study provides timely and relevant insights into the behavioural drivers of mobile voice AI adoption in India. As voice technologies continue to evolve and integrate into our daily life, understanding user attitudes and concerns will be of major help to developers, policymakers, and marketers aiming to foster broader, wider, and responsible adoption of AI-driven interfaces.

To sum up, the study offers some great insights about factors influencing India's adoption of mobile voice AI. Product developers, marketers and policy makers can benefit greatly by enhancing trust and reducing consumers' privacy concerns.

6 REFERENCES

[1] S. K. Sharma and Y. J. Choi, "Adoption of AI-based voice assistants: TAM model perspectives," *Telematics and Informatics*, vol. 59, p. 101561, 2021.

[2] M. Hoy, "Alexa, Siri, Cortana, and More: An introduction to voice assistants," *Medical Reference Services Quarterly*, vol. 37, no. 1, pp. 81–88, 2018. https://doi.org/10.1080/02763869.2018.1404391

[3] J. Lee, K. Kim, and H. Lee, "Mobile voice assistants and user acceptance: A review and future directions," *International Journal of Human–Computer Interaction*, vol. 37, no. 11, pp. 1082–1097, 2021.

[4] A. Bhatia and M. Goyal, "Accents, dialects, and digital divide: Performance of voice AI in India," *Journal of AI and Society*, vol. 38, no. 2, pp. 401–414, 2023.

[5] S. Purington, J. Taft, S. Sannon, N. Bazarova, and M. Taylor, "Alexa is my new BFF: Social roles, user satisfaction, and personification of voice-enabled AI," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017, pp. 2853–2859. https://doi.org/10.1145/3027063.3053246

[6] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989. https://doi.org/10.2307/249008

[7] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Quarterly*, vol. 27, no. 3, pp. 425–478, 2003. https://doi.org/10.2307/30036540

[8] R. Gefen, E. Karahanna, and D. W. Straub, "Trust and TAM in online shopping: An integrated model," *MIS Quarterly*, vol. 27, no. 1, pp. 51–90, 2003. https://doi.org/10.2307/30036519

[9] A. Kapoor, S. Dwivedi, and S. Piercy, "Consumer acceptance of artificial intelligence-enabled voice assistants: An empirical study," *Technol. Forecast. Soc. Change*, vol. 163, p. 120431, 2021.

[10] G. McLean and K. Osei-Frimpong, "Hey Alexa … examine the variables influencing the use of artificial intelligent in-home voice assistants," *Comput. Human Behav.*, vol. 99, pp. 28–37, 2019. https://doi.org/10.1016/j.chb.2019.05.009

[11] Y. C. Wang and C. Y. Lin, "Predicting user acceptance of voice assistants: An integration of trust, perceived risk, and TAM," *J. Retail. Consum. Serv.*, vol. 68, p. 103103, 2022.

[12] B. R. Cowan *et al.*, "'What can I help you with?' Infrequent users' experiences of intelligent personal assistants," *Interacting with Computers*, vol. 29, no. 3, pp. 344–361, 2017.

[13] A. Moorthy and K.-P. L. Vu, "Examining the effect of age and distraction on touchscreen smartphone use while driving," *Human Factors*, vol. 57, no. 4, pp. 671–682, 2015.

[14] R. Gupta and S. Arora, "Influence of social conformity on voice assistant adoption in Urban India," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 3, pp. 45–58, 2021.

[15] A. Pradhan, M. Mehta, and L. Findlater, "Accessibility came by accident: Use of voice-controlled intelligent personal assistants by people with disabilities," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2018, pp. 1–13. https://doi.org/10.1145/3173574.3174033

[16] E. Lee and J. Choi, "Perceived usefulness and user acceptance of voice AI in smartphones," *Telematics and Informatics*, vol. 59, p. 101561, 2021. https://doi.org/10.1016/j.tele.2020.101561

[17] R. Zhou, X. Jiang, and M. Knijnenburg, "Investigating the role of enjoyment in user engagement," *User Modeling and User-Adapted Interaction*, vol. 29, pp. 329–377, 2019.

[18] J. H. Lau, B. C. Zimmerman, and A. Schaub, "Alexa, are you listening?: Privacy perceptions, trust, and voice assistants," *Proc. ACM Human-Computer Interaction*, vol. 3, no. CSCW, pp. 1–24, 2019. https://doi.org/10.1145/3274371

[19] D. Gefen, E. Karahanna, and D. W. Straub, "Trust and TAM in online shopping: An integrated model," *MIS Quarterly*, vol. 27, no. 1, pp. 51–90, 2003. https://doi.org/10.2307/30036519

[20] C. Lutz and G. Newlands, "Privacy and smart speakers: A multicountry survey on adoption and usage," *ACM Trans. Internet Technol.*, vol. 21, no. 1, pp. 1–29, 2021.

[21] P. Dhamija and R. Yadav, "Privacy concerns and adoption of voice assistants in India: A study of millennials," *Technological Forecasting and Social Change*, vol. 183, p. 121939, 2022.

[22] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions," *Decision Sciences*, vol. 39, no. 2, pp. 273–315, 2008. https://doi.org/10.1111/j.1540-5915.2008.00192.x

[23] K. Kapoor, S. Dwivedi, N. Piercy, and Y. K. Dwivedi, "Influence of word-of-mouth and social media on technology adoption," *Information Systems Frontiers*, vol. 22, no. 2, pp. 289–309, 2020.

[24] S. Gupta and R. Arora, "Voice assistants and urban tech adoption in India," *Asian Journal of Management Science*, vol. 12, no. 1, pp. 112–124, 2023.

[25] Z. Zhang and K.-H. Yuan, "WebPower: An R package for statistical power analysis," *Behavior Research Methods*, vol. 50, no. 2, pp. 750–758, 2018.

[26] N. Kock, "Common method bias in PLS-SEM: A full collinearity assessment approach," *International Journal of e-Collaboration (IJeC)*, vol. 11, no. 4, pp. 1–10, 2015. https://doi.org/10.4018/ijec.2015100101

[27] N. Kock and G. S. Lynn, "Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations," *Journal of the Association for Information Systems*, vol. 13, no. 7, pp. 546–580, 2012. https://doi.org/10.17705/1jais.00302

[28] N. Kock, "WarpPLS User Manual: Version 6.0," ScriptWarp Systems, Laredo, TX, 2017.

[29] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, 3rd ed. New York, NY, USA: McGraw-Hill, 1994.

[30] C. Fornell and D. F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error," *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50, 1981. https://doi.org/10.1177/002224378101800104

[31] J. Henseler, C. M. Ringle, and M. Sarstedt, "A new criterion for assessing discriminant validity in variance-based structural equation modeling," *Journal of the Academy of Marketing Science*, vol. 43, pp. 115–135, 2015. https://doi.org/10.1007/s11747-014-0403-8

[32] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 2nd ed. Thousand Oaks, CA: SAGE Publications, 2017.

7 AUTHORS

Ms. Mehak Mittal is a Research Scholar at the Bharati Vidyapeeth (Deemed to be University), Pune Institute of Management and Research, New Delhi, India (E-mail: mehak15597@gmail.com).

Dr. Sanjay Manocha is an Assistant Professor at the Bharati Vidyapeeth (Deemed to be University), Pune Institute of Management and Research, New Delhi, India (E-mail: sanjay.manocha@bharatividyapeeth.edu).