The Role of Voice-Activated Technologies in Shaping Podcast Engagement and Service Satisfaction: A Theoretical and Empirical Synthesis Through HMC Framework

Dr. Praveen Kumar Pandey (*)
Postdoctoral Researcher, Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia
Orcid Id - 0000-0002-9256-0093
praveen.pandey2022@gmail.com
pdf.praveenpandey@lincoln.edu.my

Prof. (Dr.) Sailesh Suryanarayan Iyer
Adjunct Research Faculty and Supervisor, Lincoln Global Post Doctoral Programme, Lincoln University College Malaysia, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia

Professor & Principal, Narnarayan Shastri Institute of Technology-Institute of Forensic Sciences and Cyber Security (Affiliated to NFSU, Gandhinagar) Ahmedabad
Orcid Id - 0000-0002-0619-9829
drsaileshiyer@gmail.com

Corresponding Author (*) - Dr. Praveen Kumar Pandey, Postdoctoral Researcher, Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia & can be reached at praveen.pandey2022@gmail.com

Sustainable Global Societies Initiative · Aligned with SDG 9 (Industry, Innovation and Infrastructure)

1. Refined Conceptual Foundation and Theoretical Positioning

The listening experience is now mediated by a voice-activated device placed between listener and catalogue, with the percentage of podcasts recognized, speed of recognition and embedded intelligence impacting the listening experience. The project questions are the following: To what extent do two voice podcast qualities, Voice Interface Quality (VIQ) and Smart Features Integration (SFI) influence the users' service judgments, how do users' service judgments influence engagement and satisfaction with the service, and which groups of users are most affected by the results of the two service qualities, depending on their Technological Readiness (TR)? These connections are justified by Conference 1 and tested by Conference 2.

The study maintains the working of human-machine communication (HMC) theory on the point of measurement. HMC believes that the user remains focused on the experience until he or she is defocussed from it and the machinery nature of the source becomes clear; then, beliefs about machines rule how the

user reads the experience (Guzman & Lewis, 2020). Thus, VIQ and SFI are specified in relation to actual signals of performance (misrecognition, lag, and relevance of the recommendation), while TR is modelled as a moderator of the technology-to-evaluation relationships, and not as another aspect. Multiple gratifications are used to explain the continued use while functional cues influence attitudes (Pitardi & Marriott, 2021) and social cues lead to trust development, we thus thought it was necessary to route each cue through three evaluative process methods: perceived usefulness, service quality, and interactive experience quality. Each construct in Table 1 is defined and labeled as measured or not measured.

Table 1. Construct definitions, measurement-model specification, and operationalised measures.

Construct (items)	Conceptual definition	Specification; scale	Anchor source (adapted)
Voice Interface Quality, VIQ (5)	Judgement of the spoken interaction layer: recognition accuracy, robustness to accents and noise, latency, naturalness of the voice.	Reflective, first-order; 7-pt Likert	Pitardi & Marriott (2021)
Smart Features Integration, SFI (5)	Perceived depth of embedded AI capability: personalised, context-aware recommendations, cross-device continuity, ecosystem links.	Reflective, first-order; 7-pt Likert	McLean & Osei-Frimpong (2019)
Perceived Usefulness, PU (4)	Degree to which the platform is believed to improve the listening task.	Reflective, first-order; 7-pt Likert	Davis (1989)
Service Quality, SQ (5)	Overall judgement of the reliability, responsiveness, and assurance of the technology-mediated service.	Reflective, first-order; 7-pt Likert	Parasuraman et al. (1988)
Interactive Experience Quality, IEQ (4)	Judgement of the interaction as smooth, intuitive, and enjoyable.	Reflective, first-order; 7-pt Likert	Pitardi & Marriott (2021)
Technological Readiness, TR (16)	Tendency to take up new technology: optimism and innovativeness set against discomfort and insecurity.	Higher-order, reflective-formative (TRI 2.0, 4 dimensions); 7-pt	Parasuraman & Colby (2015)
User Engagement, UE (10)	Strength of cognitive, affective, and behavioural investment in the platform.	Higher-order (3 dimensions); 7-pt Likert	Established multidimensional consumer-engagement scales
Service Satisfaction, SAT (4)	Overall post-use evaluation of the service.	Reflective, first-order; 7-pt Likert	Standard post-use satisfaction scales

2. Conceptual Model and Hypotheses (Restated for Estimation)

Figure 1 reproduces the model validated at Conference 1. For Conference 2, each composite claim is split into its own structural path, so every path can be tested and reported on its own.

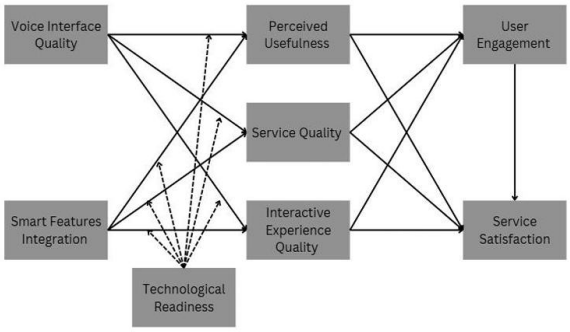

Fig. 1: Conceptual Model (Self-sourced)

H1a–H1c / H2a–H2c. VIQ (H1) and SFI (H2) are positively related to (a) perceived usefulness, (b) service quality, and (c) interactive experience quality.

H3a–H3c / H4a–H4c. Perceived usefulness, service quality, and interactive experience quality are positively related to user engagement (H3) and to service satisfaction (H4).

H5. User engagement is positively related to service satisfaction.

H6a–H6c / H7a–H7c. Technological readiness positively moderates the links from VIQ (H6) and from SFI (H7) to perceived usefulness, service quality, and interactive experience quality.

The implied indirect and moderated-mediation paths are tested and reported in Section 4.

3. Research Methodology

It is a post-positivist, deductive work, based on a quantitative, cross-sectional survey (question is explanatory and variance oriented, and survey data enable the whole system of effects to be estimated). Only one time point can't demonstrate causality. This is achieved by providing paths based on causal arguments, distinguishing between predictor and criterion measurement within the instrument, and by statistically testing for/common method variance; the satisfaction measure will be lagged and will be included in Conference 3.

In the domain, that of smart-speaker podcast listening in India, we are looking at a burgeoning market with both both fast uptake, and even more variable technological readiness and hard multilingual recognition conditions – just the sort of variation for which a moderation test is there. The target audience are adult listeners (18+) who have opened and/or interacted with podcast content using any voice-activated device within the last 3 months (Amazon Alexa, Google Assistant or Apple Siri). There is no sampling frame and therefore a non-probability sample is used, based on the online panel and screened by quotas (age band and gender as well as metro vs non-metro location). It follows that according to the inverse-square-root method (Kock & Hadaya, 2018) that it is adequate to test a path coefficient close to

0.13 with 80% power with approximately 384 cases, and the desired number of usable responses is established at 400.

All of the latent variables are assessed by multi-item reflective scales taken from well-designed instruments (Table 1) measured by seven-point Likert scales. Technological readiness is a reflective-formative higher order construct formed by the four TRI 2.0 dimensions (Parasuraman & Colby, 2015) and user engagement as a higher-order construct of cognitive, affective, and behavioural dimensions, both of which are estimated using the disjoint two-stage approach.

Predictor and outcome blocks in the questionnaire are separated, and the order of the items within the predictor and outcome blocks are randomized (Podsakoff et al., 2003). Common method variance is handled procedure (anonymity, no leading demands, attention checks) and statistically: Harman's single factor test as a baseline, the final judgement is based on the full-collinearity VIF-check (cut-off value is 3.3).

4. Analytical Strategy

There are two tests of the model. The main estimator used in SmartPLS 4 is PLS-SEM, which can be used to analyze a large model including higher order constructs, parallel mediators, interaction terms and, apart from its explainational function, also prediction (Hair et al., 2019). The measurement part is first assessed for reliability (composite reliability ≥0.70, Cronbach's α ≥0.70), convergent validity (loadings ≥0.70 and AVE ≥0.50), and discriminant validity (HTMT < 0.85); the formative part is tested for collinearity and the statistical significance of outer-weights. Then, inner VIF below 3.3 and bootstrapping (10 000 subsamples, bias-corrected confidence intervals) are performed in the structural model, and R-squared and f-squared are provided to avoid the pitfalls of inferring significance as practical weight in the structural model.

The indirect effects are tested by bootstrapping the mediation which classifies each as complementary, competitive or indirect-only. Moderation (H6, H7) applies the two-stage method with product terms and illustrates simple-slope plots and the conditional indirect model reduces the use of equation parameters to values at representative levels of the moderator. PLSpredict performs out-of-sample prediction and importance–performance map analysis converts the prediction results into design priorities.

Given the assumption of symmetry and additivity in PLS-SEM and the often real nature of users' judgment, fuzzy-set qualitative comparative analysis (fsQCA) is included with the question: "What combinations of VIQ, SFI, PU, SQ, IEQ, and TR are sufficient for a high level of engagement and high satisfaction (Pappas & Woodside, 2021)? Responses are scaled into fuzzy-set membership scores, a truth-table is created and

reduced, and configurations are analyzed for consistency (usually greater than 0.80) and coverage, including a run of necessary-condition analysis first. The differences between the two methods, if any, are openly reported.

5. Ethical Considerations

The study will commence as soon as a clearance is given from the institutional Ethics committee. The participation is voluntary and therefore based on informed consent, the right to withdraw without penalty. The collection does not include any personally identifiable information, responses are anonymized, and questions related to attitude towards AI and data protection are formulated in a neutral manner.

6. Anticipated Contributions

The study, in theory, contributes to transforming human-machine communication from the largely conceptual and laboratory settings to a field survey account of voice mediated media communication. This is methodologically a repeatable instrument for doing voice podcast research and combining PLS-SEM with fsQCA to report net effects and sufficient configurations. It translates into practice and policy the opportunities for platform designers to invest in ways that foster greater engagement and satisfaction, and for SDG 9, how a human-centred infrastructure approach to voice can expand participation in the digital audio economy in the emerging market.

7. Limitations and Bridge to Conference 3

There are three known limitations, which are openly discussed: the cross sectional design does not allow to make conclusions about causal relationships, the non-probability sample does not allow for generalisation and self-reported instruments can have method variance. Two relationships will be re-estimated, findings will be validated against a holdout sample, and if indicated, a time-lagged satisfaction measure will be added to the empirical study conducted in Conference 3 so that it is not redesigned.

References

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A Human–Machine Communication research agenda. New Media & Society, 22(1), 70–86.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31(1), 2–24.

Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. Information Systems Journal, 28(1), 227–261.

McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa… Examine the variables influencing the use of artificial intelligent in-home voice assistants. Computers in Human Behavior, 99, 28–37.

Pappas, I. O., & Woodside, A. G. (2021). Fuzzy-set qualitative comparative analysis (fsQCA): Guidelines for research practice in information systems and marketing. International Journal of Information Management, 58, 102310.

Parasuraman, A., & Colby, C. L. (2015). An updated and streamlined technology readiness index: TRI 2.0. Journal of Service Research, 18(1), 59–74.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. Journal of Retailing, 64(1), 12–40.

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. Psychology & Marketing, 38(4), 626–642.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879–903.