LUISS

Degree Program in Marketing - Analytics and Metrics

Course of Marketing Metrics

How voice interaction shapes consumer decision making: a comparative study of intelligent assistant and digital channels

Michele Costabile

SUPERVISOR

Matteo De Angelis

CO-SUPERVISOR

786291

CANDIDATE

Academic Year 2024/2025

To my supervisor, co-supervisor, and assistants,

for their guidance and support throughout this thesis.

INDEX

INTRODUCTION

Over the past decade, conversational interfaces have evolved from novelty to near-ubiquity over the past decade, becoming an integral part of daily consumer routines. From voice assistants built into smartphones and vehicles to retailer-deployed chatbots, these systems now mediate a wide range of tasks: searching, comparing, even purchasing. The appeal is clear: hands-free operation, conversational ease, and frictionless input. But voice also brings structural constraints: information is delivered linearly, memory demands are higher, and users have fewer opportunities to scan, compare, or reverse actions. Meanwhile, visual interfaces have grown more sophisticated, offering dense, glanceable data: reviews, filters, side-by-side comparisons, and cues that help users build trust quickly. Academic discourse mirrors this push-pull. Some research highlights ease of use and effort expectancy; others zero in on how modality affects perceived control, cognitive load, and downstream satisfaction. However, the findings remain piecemeal. Many studies stop at usage intentions, without tracing what happens *during* the decision-making process. Others examine only one modality in isolation, sidestepping direct comparisons. Moderators like product type (e.g., utilitarian vs. hedonic) and internal states such as perceived control are often mentioned, but rarely tested together within a single experimental framework. Practitioners must make real choices: when to use voice, when to switch to screen, and how to create satisfaction. But there's little in the way of a coherent, causal map to guide them.

This study durectly addresses this gap. It asks how interaction mode, voice or screen, shapes consumer decision-making, and under what conditions these effects strengthen or fade. The goals are threefold: to test whether voice speeds up decisions, to assess whether perceived control mediates the relationship between mode and perceived decision quality, and to evaluate whether satisfaction improves when the mode aligns with the product type - using voice for utilitarian tasks and screen for hedonic ones, which is a popular claim in commercial circles. There's also a practical aim: to extract usable insights that can help orchestrate digital experiences, clarifying when voice should lead, when screens should take over, and how to design for perceived control, trust, and engagement so users walk away feeling confident in their decisions.

The first chapter lays the conceptual groundwork. It tracks the evolution of conversational tech in commerce and synthesizes what we know about how mode shapes decision-making. Key issues include how information is presented (sequentially vs. in parallel), the impact on cognitive load, and what types of control cues are available. Classic categories like hedonic vs. utilitarian goods are reframed not as static labels, but as indicators of task demand, some choices call for speed, others for verification or experience. Perceived control is explored as a core psychological mechanism through which interfaces influence perceived decision quality, while constructs like trust and engagement are positioned as experience-centric drivers of satisfaction. From this synthesis emerges a model in which interaction mode affects speed, perceived quality, and satisfaction, with control acting as a moderator and product type potentially conditioning the satisfaction response. This model is then translated into hypotheses that anchor the study.

Chapter two details the experimental design. A between-subjects layout assigns participants to one of four groups: Voice–Hedonic, Voice–Utilitarian, Screen–Hedonic, or Screen–Utilitarian. After exposure to their assigned condition, participants rate their experience on seven-point scales assessing satisfaction, trust, engagement, perceived control, usefulness, and both perceived and logged decision time. A composite index for decision quality is created by averaging four related dimensions, and its reliability is verified statistically.

Data preparation is meticulous: variables are coded and recoded for consistency, values standardized, and checks for missing data and outliers are conducted. Each hypothesis is matched with a tailored statistical test: t-tests for speed comparisons, a moderated regression for perceived quality (mode × control), and a two-way ANOVA for satisfaction (mode × product type). Additional diagnostics such as distribution checks, visualizations, correlations, help ensure robustness and context. A lean but informative set of extra analyses is also planned: regression modeling of satisfaction, and an exploratory link between actual decision time and perceived confidence.

Chapter three walks you through the results. Descriptive stats confirm sufficient variation and a balanced sample. The t-test results indicate no significant difference in decision speed between voice and screen: a finding that pushes back against assumptions that hands-free input is inherently faster when cognitive demands are high. The analysis of decision quality shows that perceived control is a solid predictor, but also reveals a key interaction: the effect of control on perceived quality is more pronounced for screen users. This suggests that visual affordances -

filters, reversibility, overviews - make users feel more in command, and that this sense of control is more easily converted into satisfaction.
Satisfaction itself is shaped mainly by mode, not by a match between product type and interaction style. Screens outperform voice regardless of whether the task is hedonic or utilitarian. Follow-up regressions clarify the drivers of satisfaction: trust and engagement matter most. Control fades in significance once these experience-centered constructs are accounted for. Faster decision times correlate with greater confidence - but not necessarily with higher satisfaction - pointing to an individual trait (decisiveness) rather than a mode effect.

These findings carry weight for both theory and practice. Theoretically, they challenge the assumption that voice is inherently more efficient. Instead, they suggest that its sequential flow can limit decision support. The work also adds nuance to the concept of perceived control, showing that it's not just the feeling of control that matters, but the *conversion* of that feeling into confidence and satisfaction, which depends heavily on interface design. Most importantly, the research shifts focus away from speed as the central metric. What seems to matter more are the experiential levers - trust, clarity, engagement - that shape how users feel about their choices.

On the managerial side, the implications are direct. When consumers need to explore, compare, or verify, screens are the better choice. Voice is better suited for early-stage filtering, simple tasks, or hands-free contexts - but should quickly hand off to screen as complexity rises. Voice flows must be designed intentionally, with control cues made explicit - "I found three options," "You can go back at any time," "Here's what we've got so far." Trust signals and engagement tools should be embedded throughout the journey, in both modes. Above all, teams should stop treating speed as the gold standard. What counts more is how empowered, engaged, and confident users feel.

Ultimately, this study delivers a clear message: screen-based interfaces better support high-quality, satisfying decisions, not because they’re faster, but because they surface the right cues at the right time. Voice still has a role to play, especially early in the journey. The future lies in well-designed, flexible multimodal experiences that blend voice for convenience and screen for confidence.

Chapter I: Description of the Phenomenon - The Impact of Voice Assistants on Consumer Decision-Making Processes

1.1 Introduction and Overview

1.1.1 Overview of the Voice Assistant Phenomenon

Voice assistants (Vas) have become a pervasive element of contemporary life, fundamentally reshaping consumer behavior and interaction with technology. Systems such as Amazon Alexa, Apple Siri, and Google Assistant, leveraging advances in artificial intelligence, process spoken language to execute a range of functions, from answering questions to providing product suggestions. This technological shift signals not only the emergency of a new user interface but also marks a transformation in consumer attitudes toward digital engagement.

But what are the definition of Vas? A voice assistant is intelligent software that responds to voice commands and runs on various devices like smartphones, speakers, computers, tablets, wearables, gaming consoles, TVs, Virtual Reality headsets, cars, and internet of things devices *(Wohr, 2025)*. Unlike traditional interfaces that rely on visual cues or physical inputs, voice assistants enable communication through natural language allowing users to issue commands, ask questions, and receive recommendations conversationally (*Flavián et al., 2022*). This shift fosters seamless multitasking: users can set reminders, control smart home devices, or conduct purchases with hands-free ease.

The functionality of voice assistants now extends beyond simple task execution. Increasingly, these systems serve as decision-making agents, personalizing recommendations and anticipating user needs based on contextual cues. A recent study highlights this evolution from basic functionality to context-aware personalization, positioning voice assistants as both practical tools and engaging companions in daily routines *(Acikgoz, Perez-Vega, 2023)*. Moreover, voice assistants represent a new form of intelligent recommender system, complementing established platforms such as electronic word-of-mouth and online reviews. The immediacy and naturalness of voice-based recommendations can uniquely influence consumer behavior.

Empirical research suggests that recommendations delivered via voice are often more effective than those provided in text, likely due to the inherent warmth, emotional resonance, and social presence conveyed by the auditory modality *(Flavián et al., 2022).*
These features have been found to enhance perceived credibility, usefulness, and the overall quality of user interactions. There is salso a growing social and emotional dimension to human interaction with voice assistants. Many users engage with these devices as if they were human interlocutors, exhibiting politeness, expecting empathy, and establishing trust-based relationships. Such anthropomorphism not only strengthens the assistant's impact on user behavior but also shapes attitudes, expectations, and perceptions of associated brands *(Moriuchi, 2019).*

Voice assistants are thus becoming deeply integrated into the consumer experience, influencing both decision-making processes and brand perceptions. Given their rapid proliferation across global markets, voice assistants are poised to become an increasingly strategic and influential presence. By 2023, the number of voice-enabled devices was projected to surpass the global population, indicating their pervasive integration in households, mobile devices, and vehicles *(Wohr, 2025).* In this context, it is essential for both researchers and practitioners to investigate the psychological and behavioral mechanisms underpinning voice-based interactions, with the aim of designing user experiences that are effective, engaging, and trustworthy.

1.1.2 Global Adoption and Diffusion of Voice Assistants

Over the past decade, the proliferation of voice assistants has transformed them from a niche feature embedded in smartphones to an integral component of daily digital life. These systems are now embedded throughout a range of devices - smart speakers, wearables, vehicles, household appliances, and broader IoT ecosystems. The rate of expansion has been substantial; industry estimates projected an average annual growth of approximately 28% between 2021 and 2023, with the remarkable milestone of 8.4 billion voice-enabled devices globally by 2022, surpassing the global human population. In the United States alone, the number of users is forecasted to reach 154.3 million in 2025, with Google Assistant leading the market (92.4 million users), followed by Apple's Siri (87.0 million) and Amazon's Alexa (77.6 million), and is expected to grow further to 170.3 million by 2028 *(Wohr, 2025).*

Significant advancesin artificial intelligence, particularly in natural language processing, cloud computing and machine learning, inextricabily link to this exponential growth. As a result, modern voice assistants exhibit enhanced capacity for interpreting user intent, recognizing context, and adapting to diverse linguistic and accent variations. These technological improvements have rendered voice interactions more accurate and accessible, thus broadening their appeal across different demographics and levels of technological literacy.

Demographically, millennials in the United States represent the most active users of voice assistants, accounting for over a third of national usage. All age groups are increasingly adopting voice assistants, and this trend i salso evident internationally, with countries like China, the United Kingdom, and Germany witnessing significant uptake. Adoption patterns also reveal a "barbell effect," with Gen Z and seniors driving growth - Gen Z as digital natives and older users due to the accessibility advantages offered by voice interfaces *(Wohr, 2025).* The integration of voice assistants into travel behaviors is evident, with approximately half of travelers utilizing these tools for real-time information and navigation.

The widespread adoption of voice assistants can be attributed to their seamless integration into everyday routines and digital ecosystems since their functionalities span a broad spectrum, even applications related to mental health. Furthermore, integration with IoT infrastructures allows users to control multiple devices via voice commands, thereby streamlining digital interactions and reducing user friction. It also is noteworthy that voice assistants are not confined to a singular use case or platform. Their continuous availability and embedding into various digital services have positioned them as ubiquitous digital companions, influencing user preferences and behaviors in both private and commercial settings. The impact of voice technology is also evident within social media environments. As the prevalence of voice assistants continues to rise, their influence on consumer decision-making, commerce, and media engagement is expected to deepen, fundamentally reshaping patterns of digital interaction and daily life.

1.1.3 Growing Importance in Consumer Behavior and Marketing Strategies

Artificial intelligence has become deeply rooted in modern consumer behavior and marketing strategies, particularly through the integration of voice assistants such as Alexa, Siri, and Google Assistant. These devices are no longer merely technological novelties; rather, they have become essential components of daily routines. Consumers now rely on voice assistants for a diverse array of tasks, ranging from information-seeking to product evaluation and purchase decisions. The shift from traditional web-based browsing to voice-driven interactions fundamentally alters the decision-making process. Where consumers previously sifted through extensive lists of options online, voice interfaces often present only a single response or very limited choices, thereby expediting decisions and shaping brand visibility in new ways (*Flavián et al., 2022).*

This streamlined process can result in consumers making faster, more habitual choices, particularly in the context of routine or low-stakes purchases *(Dellaert et al., 2020).* For marketers, this feature presents both significant opportunities and notable challenges. Voice interfaces enable brands to deliver highly personalized, context-aware messages tailored to individual preferences, behaviors, and even specific languages *(Acikgoz, Perez-Vega, 2023).* At the same time, the competitive nature of voice search necessitates alignment with AI algorithms and the development of effective voice strategies. Brands that fail to adapt risk diminished visibility and influence.

Consumer engagement with voice assistants extends beyond simple usage frequency. It encompasses emotional, cognitive, and behavioral investment. Research suggests that this more profound engagement is a more robust predictor of brand loyalty than traditional satisfaction metrics *(Moriuchi, 2019).* Many users interact with voice assistants in ways that mirror social relationships, including the use of polite language and reciprocal turn-taking. This social dimension promotes confidence and emotional closeness, positioning the voice assistant as a relational intermediary between the consumer and the brand *(Moriuchi, 2019).*

As user trust increases, so does reliance on the assistant, reinforcing patterns of behavior that ultimately promote brand loyalty. With continued advancements in machine learning, voice assistants are becoming increasingly sophisticated and integral to the consumer journey. They now function as both practical tools and strategic brand touchpoints, influencing everything from product discovery to post-purchase engagement. This omnichannel presence provides marketers with new avenues to shape consumer preferences, enhance brand loyalty, and deliver consistent value across the consumer lifecycle.

1.2 The Communicative Power of Voice: Cognitive and Emotional Foundations

1.2.1 The Psychological Effect of the Human Voice in Communication

The communicative power of the human voice is both profound and multifaceted, grounded in evolutionary, psychological, and social mechanisms. From infancy, individuals demonstrate a remarkable sensitivity to vocal cues such as tone, rhythm, and pitch, factors which communicate emotional states, intentions, and relational warmth. The prosodic qualities of the voice make it a multidimensional tool, not simply transmitting information but also facilitating emotional exchange, fostering social cohesion, and shaping perceptions. In the arena of human-technology interaction, these dynamics become especially salient: the vocal characteristics of digital assistants, for example, can significantly influence user engagement and attitudes toward both the agent and its associated brand.

Spoken language, as compared to written communication, is marked by increased spontaneity and subjectivity. Speech unfolds in real time, lacking the opportunity for extensive revision or reflection, which thus prompts greater emotional expressiveness and freedom. This immediacy is evident in the frequent use of personal pronouns and emotionally charged vocabulary, resulting in interactions that are perceived as more authentic and informal. Empirical studies, such as *Epley and Schroeder (2014)* provide evidence that speech tends to elicit affective reasoning, enhancing the sense of authenticity and personalization in voice-based exchanges, particularly when emotionally salient topics are discussed.

A notable psychological phenomenon in this context is the "feeling right" effect, or preference fluency. When the mode of communication, such as speaking, aligns with the nature of the decision at hand (e.g., decisions about experiential or pleasure-oriented products), individuals report a metacognitive sense of appropriateness and confidence. This congruence facilitates smoother and more satisfying decision-making, especially in emotionally significant situations. As a result, consumers may perceive their vocalized thoughts as more trustworthy and informative, which enhances both decision confidence and overall satisfaction (*Epley, Schroeder, 2014)*.

Voice interfaces also heighten the user's sense of social presence, the impression that a social entity is present during the interaction. This is especially significant in digital commerce, where the assistant is not human but is designed to simulate human-like responses. Scholars such as *Flavián et al. (2022)*, argue that this emotional richness can foster parasocial bonds, leading users to interact with voice assistants as if they were genuine social partners. These bonds may increase user satisfaction, perceived intelligence of the assistant, and brand affection, thereby reinforcing trust and habitual engagement. These psychological mechanisms are particularly influential in emotionally-driven or low-stakes decisions, where intuition often takes precedence over rational analysis. Voice interfaces can streamline such decisions, sometimes promoting more impulsive or habitual consumer behavior. While such behavior may benefit brands, especially in contexts involving repeat purchases or emotionally charged products, it does present ethical considerations related to transparency and informed consent.

1.2.2 How Voice Creates Social Presence and Trust

Voice as a communication channel, holds a unique capacity for conveying emotional and social cues, which is particularly significant in the realm of human-technology interaction - especially with voice assistants. When users engage with these systems, there is a notable tendency to attribute social presence to the assistant, perceiving it as more than a mere tool. This phenomenon, well-documented in the Computers are Social Actors (CASA) paradigm *(Nass, Steuer, 1994),* demonstrates that when technology mimics human conversational patterns, individuals unconsciously apply the social rules and expectations typically reserved for human interaction. Unlike text-based interfaces, which lack immediacy and vocal nuance, voice interactions facilitate real-time dialogue. The inclusion of elements such as intonation, warmth, and pacing introduces verbal and nonverbal cues that clarify intent and reduce ambiguity - factors that are critical for fostering relational trust.

Empirical research, such as that *Moriuchi (2019)* - has shown that users perceive voice assistants with warmer and more enthusiastic vocal tones as not only more trustworthy but also more competent. This perception, in turn, increases the likelihood that users will rely on these systems for advice and decision-making. Moreover, voice-mediated communication enables a higher degree of emotional expressiveness from users. Individuals are more inclined to disclose personal information or discuss emotionally charged topics, as the medium itself supports affective engagement. This dynamic reduces psychological distance, enhances perceived warmth, and establishes a foundation for trust, both toward the assistant and the brand it represents.

While recent scholarship, including work by *Acikgoz et al. (2023*), highlights the significant of social intelligence in voice assistants. Systems that can adapt to user mood, maintain conversational continuity, and tailor responses to context are more likely to foster emotional connection and long-term behavioral loyalty. In these instances, the relationship between user and assistant shifts from functional to relational, with the assistant being regarded as a reliable partner rather than a simple tool. In this sense, the implications for user experience and commercial strategy are significant. Higher perceived social presence leads users to accept product recommendations from voice assistants with greater confidence, often foregoing additional research. This trust extends to the associated brand, amplifying satisfaction and affinity. Brands that humanize their voice interfaces through natural dialogue, friendly tone, and cultural sensitivity can generate more robust trust and loyalty.

1.2.3 Cognitive Processing Differences Between Spoken and Written Input

The way people interact with technology, whether by speaking or writing, directly shapes how they process information, make judgments, and make decisions. If you look at something like product recommendations from voice assistants, the differences between talking and typing become even more obvious and crucial. Spoken input is usually spontaneous and loaded with emotion. People don’t really plan out what they’re about to say to Alexa or Siri; they just blurt it out, using personal pronouns and expressing feelings. This is what psychologists would label “feeling-based” or System 1 thinking: it’s quick, intuitive, and, let’s face it, sometimes impulsive *(Epley, Schroeder, 2014)*. You acquire more narrative, personal language, and decisions are often driven by what feels right in the moment rather than by careful analysis. On the flip side, writing things out gives users space and time to think. Typing out a query or response is more structured and deliberate, there’s room for editing and for organizing thoughts. This is where System 2, or “reason-based” processing, comes in. It’s slower and more analytical. Written messages tend to be more formal and logically structured, prioritizing clarity and depth over raw emotion. Actual studies back this up. Using tools like LIWC (Linguistic Inquiry and Word Count), researchers have found that people who interact by voice are more emotional and less analytic, while those using text are more logical and organized. *(Koschate-Fischer, Hoyer, Schindler, 2024).* So, it's not just a gut feeling; it's measurable.

These differences have real consequences for how consumers handle product recommendations. Spoken information comes one after the other, you hear something, and then it's gone unless the assistant repeats it. This makes it hard to compare options, so people often just go with whatever's suggested first or second. That fleeting nature can encourage impulsive choices and default bias, because, who really remembers all choices the VAs just rattled off? And there's more: voice-based interactions come with nonverbal cues like tone and pacing. As already said previously, these can make the voice assistant seem more relatable or trustworthy, which sometimes boosts the sense of connection or even persuasion *(Guha et al., 2023).* Yet, this can also mean users feel a subtle nudge instead of real autonomy, and analytic depth might take a hit.

Written input, in contrast, is persistent. You can scroll through all the recommendations, compare them side-by-side, and revisit details as needed. Visual memory kicks in, helping you recall specifics, which is especially handy for utilitarian products where you care about specs, price, or features. This setup gives users more control over the decision process and strengthens their sense of being the one in charge *(Dellaert et al., 2020).*

There's also a fit between input style and product type. If you're evaluating hedonic items—stuff that's about pleasure, like entertainment or food—spoken input actually works well because the emotional, quick-thinking style matches the context. But for practical, utilitarian products, written input supports more careful analysis and boosts user confidence *(Flavián et al., 2022), (Dellaert et al., 2020), (Koschate-Fischer et al., 2024).*

1.3 Evolution of Voice Assistants in the Digital Landscape

1.3.1 History and Development of Voice Technologies (e.g., Alexa, Siri, Google Assistant)

Voice technologies have experienced a profound evolution, progressing from rudimentary speech recognition tools to complex, AI-driven virtual assistants now woven into the fabric of modern digital life. The early days were marked by extremely limited functionality, with systems constrained by the era's modest computing power and primitive linguistic models. It was not until the 2000s, following significant advancements in artificial intelligence and natural language processing, that these technologies began to move beyond mere sound-to-word mapping toward understanding and interacting with human intent.

The introduction of Apple's Siri in 2011 signaled a pivotal shift, making conversational computing accessible to mainstream consumers. Siri's launch, soon followed by Google Now (which evolved into Google Assistant) and Amazon's Alexa, marked the transition from voice as a simple command interface to a platform for proactive digital companionship. Each major assistant brought distinct advantages: Alexa's open "skills" ecosystem encouraged third-party development; Google leveraged its search capabilities for highly contextual responses; Apple prioritized privacy and seamless integration across its devices. Microsoft's Cortana and Samsung's Bixby, though less widely adopted, contributed to the competitive landscape.

As foundational technologies matured, voice assistants became far more sophisticated. Today, they can manage multi-turn conversations, recognize individual users by voice, even infer emotional states—shifting from reactive tools to anticipatory systems that can predict and respond to user needs in real time. This transformation is the result of ongoing progress in machine learning, speech recognition, and biometrics. It is also important to clarify terminology. While "chatbots," "voice assistants," and "intelligent speakers" are often used interchangeably, they denote different aspects of the ecosystem. Chatbots primarily operate via text; voice assistants like Siri and Google Assistant are designed for spoken interaction; intelligent speakers (such as Amazon Echo) are physical devices housing these assistants, often serving as hubs for smart home control.

The convergence of hardware and software has driven remarkable growth. By 2020, more than four billion voice-enabled assistants were in use globally, with forecasts suggesting this could double by 2024. Their presence is now ubiquitous—embedded in smartphones, vehicles, appliances, and wearables—enabled by ongoing improvements in accuracy, contextual awareness, and responsiveness. Yet, as these technologies proliferate, so too do concerns regarding privacy, data security, and user trust. The always-listening capabilities of many devices have prompted scrutiny over data collection and usage practices. Accordingly, the field's trajectory now requires not only technical innovation but also a robust commitment to transparency, user control, and ethical stewardship.

1.3.2 Integration into Everyday Devices (Smartphones, Vehicles, Smart Homes)

The widespread adoption of voice assistants represents a significant transformation in the way individuals engage with technology. No longer relegated to the realm of novelty, these digital assistants—such as Amazon's Alexa, Apple's Siri, and Google Assistant—have become deeply integrated into daily routines and a wide variety of devices. What began as a feature primarily associated with smartphones has now expanded into automobiles, smart home devices, wearables, and even kitchen appliances, reflecting a marked shift toward ubiquitous, ambient computing.

Major technology firms, including Amazon, Google, Apple, Microsoft, and Meta, are continuously advancing their voice technologies and embedding them across their respective ecosystems. Smartphones remain the primary platform for voice assistant interactions, but their presence in other domains—particularly automotive and home automation—has grown rapidly. In vehicles, voice-enabled platforms like Apple CarPlay, Android Auto, and specialized systems from automakers such as Mercedes-Benz and BMW have made voice control integral for enhancing driver safety and convenience. These systems enable users to manage navigation, communication, and entertainment hands-free, thereby reducing distractions and aligning with contemporary expectations for in-car technology.

Within domestic environments, voice assistants have become central to smart home management. Devices powered by Alexa or Google Assistant facilitate voice-based control of lighting, climate, security, and connected appliances. Open APIs and integration with third-party devices extend their functionality, allowing users to automate tasks and streamline daily routines. This frictionless, voice-driven control has become an expectation for many consumers seeking efficiency and convenience in their digital environments.

Furthermore, the expansion of voice assistant capabilities into wearable technology—such as smartwatches, wireless earbuds, and augmented reality glasses—underscores a broader movement toward ambient computing. These wearables enable discreet, immediate access to digital assistance, supporting users as they navigate various activities throughout the day without disruption. Empirical data corroborates this integration: millions of users regularly interact with voice assistants for tasks ranging from information retrieval and list management to setting reminders and controlling smart devices. The hands-free nature of these interactions is particularly advantageous in multitasking scenarios, such as cooking or driving, where manual operation is impractical.

From a strategic perspective, the integration of voice assistants into everyday life presents new opportunities for brands and marketers. Each interaction serves as a potential touchpoint for building brand affinity, fostering trust, and delivering personalized experiences. As a result, voice assistants are not merely functional tools but also vehicles for deepening consumer engagement and loyalty. The proliferation of voice assistants across an expanding array of devices signals a paradigm shift in human–technology interaction. As these systems become more natural, anticipatory, and omnipresent, they are reshaping user expectations, behaviors, and the broader digital landscape, solidifying their role as central figures in the contemporary consumer experience.

1.3.3 Differences Between Voice Interfaces and Graphical Interfaces (Websites, Mobile Apps)

Voice and graphical interfaces genuinely diverge in how they shape user experiences, cognitive processing, and decision-making. When you interact with a voice assistant, it's essentially a linear, auditory exchange—information arrives one piece at a time, and, bluntly, there's no "rewind" or "let me check that again" option. This format leans hard on short-term memory, nudging users toward quick, intuitive choices (Koschate-Fischer et al., 2024). The spoken content disappears almost instantly, which often makes users settle for the first or second option simply because it's there and easy. There's limited space for deep analysis or methodical comparison, so while the process feels smooth and sometimes engaging, it subtly limits user agency *(Flavián et al., 2022), (Dellaert et al., 2020).*

Switch gears to graphical interfaces, and the difference is night and day. Here, information is persistent, visual, and laid out spatially. Users can compare, reflect, and revisit options at their own pace. The visual format naturally leads to more deliberate, analytical thought—logic and structured reasoning take the front seat *(Epley, Schroeder, 2014)*. This visual "permanence" reinforces users' sense of control and autonomy; you're not just reacting, you're actively shaping your own path *(Koschate-Fischer et al., 2024).*

Socially, there's a clear divide as well. Voice interfaces bring in tone, inflection, even a hint of personality, which can make interactions feel more social—almost like dealing with a person, not just a gadget *(Koschate-Fischer et al., 2024), (Moriuchi, 2019)*. That extra human touch can build trust and rapport *(Guha et al., 2023).*

On the flip side, graphical interfaces are pretty neutral. They lack those emotional cues, relying instead on design and structure to communicate credibility. Despite their differences, both types of interfaces have their niches. Voice is ideal for multitasking or hands-free situations, making technology feel accessible and natural *(Wohr, 2025), (Nass, Steuer, 1994)*. Graphical interfaces, though, shine when accuracy, transparency, and depth are needed—they support detailed, confident decision-making *(Flavián et al., 2022).*

Voice and graphical interfaces are not rivals; rather, they complement each other. Each has strengths that, when understood and leveraged, can create more satisfying and effective digital experiences. Designers, marketers, and brands who recognize this balance can better align their products with the cognitive habits and expectations of their users (*Acikgoz et al., 2023), (Luchtenberg et al., 2025)*

1.4. Impact of Voice Assistants on the Consumer Decision-Making Process

1.4.1 Influence on Purchase Decisions and Search Behavior

Voice assistants have rapidly transitioned from simple transactional tools to influential intermediaries in consumer purchasing and product search. Integrated into a wide range of devices - smartphones, vehicles, smart speakers, even household appliances - these systems have made shopping almost effortlessly embedded within daily life *(Wohr, 2025).* The immediacy and convenience of hands-free, conversational interfaces mean that consumers can now initiate purchases while multitasking, fundamentally changing how, when, and why transactions occur *(Simss, 2019)***.** Yet, this shift comes at a cost. Unlike the traditional e-commerce experience, where individuals might browse, compare, and critically evaluate a broad set of options, voice assistants typically present a limited number of recommendations, often just one or two.

This sequential, auditory mode of interaction, coupled with natural constraints of human memory, results in most users accepting initial suggestions rather than seeking further alternatives *(Koschate-Fischer et al., 2024), (Epley, Schroeder, 2014).* The process is streamlined and efficient, but it reduces agency, narrows the scope of consideration, and can obscure transparency regarding how recommendations are selected *(Dellaert et al., 2020).*

Personalization sits at the heart of this transformation. Voice assistants leverage vast amounts of contextual data, past behaviors, and stated preferences to deliver recommendations that feel uniquely relevant to each user *(Acikgoz et al., 2023).* This individualized approach reduces decision fatigue and accelerates the purchasing process, but it also encourages increased reliance on the assistant as a trusted advisor *(Flavián et al., 2022).*

Over time, such habitual dependence can reinforce particular purchasing patterns and drive brand loyalty that is rooted more in convenience than in conscious deliberation *(Moriuchi, 2019).* Moreover, the emotional dimension of voice interactions should not be underestimated. Recommendations delivered through speech can convey warmth, immediacy, and even a sense of empathy, strengthening user trust and compliance *(Guha et al., 2023).* When users perceive these systems as reliable and secure, they are more likely to accept advice without significant

scrutiny. Nevertheless, this trust remains fragile, particularly given persistent concerns about privacy, data usage, and the opacity of underlying algorithms *(Luchtenberg et al., 2025)*.
For brands, these developments present both significant opportunities and new challenges. Securing a position as the default or first recommendation within a voice ecosystem can dramatically increase the likelihood of selection, effectively redefining the criteria for visibility in the digital marketplace *(Simss, 2019)***.** Consequently, brands must adapt by prioritizing transparency, meaningful personalization, and emotionally intelligent design to foster lasting trust and satisfaction. Voice assistants are restructuring consumer search and purchase behavior by compressing exploration, amplifying the persuasive power of curated recommendations, and elevating themselves from passive tools to active architects of consumer choice *(Flavián et al., 2022)*. This evolution offers convenience and efficiency, but it also raises important questions about autonomy, transparency, and the future dynamics of consumer-brand relationships *(Dellaert et al., 2020)*.

1.4.2 Effects on Perceived Control and Satisfaction

The growing ubiquity of VAs, has introduced a complex interplay of psychological factors, particularly in relation to perceived control, user satisfaction, and trust *(Flavián et al., 2022)*. Understanding these nuanced effects is essential for comprehending both immediate user experiences and the long-term implications for consumer loyalty *(Moriuchi, 2019)*.

A central theme in this dynamic is the user's sense of perceived control. Individuals report greater satisfaction when they understand the rationale behind a voice assistant's recommendations and feel empowered to influence or override these outcomes (Dellaert et al., 2020). This sense of agency is especially pronounced in decision-making contexts, where the ability to request alternatives or ask clarifying questions transforms the user from a passive recipient to an active participant *(Acikgoz et al., 2023)*. Conversely, when suggestions are delivered without clear reasoning or transparency, users may experience feelings of disempowerment or suspicion, regardless of the objective quality of the outcome *(Guha et al., 2023)*.

Context plays a substantial role in shaping perceptions of control. For straightforward, low-stakes tasks - such as checking the weather or setting reminders - users tend to value automation and reduced cognitive effort. In such cases, satisfaction is largely driven by efficiency and predictability *(Simss, 2019).* In contrast, more complex or unfamiliar tasks can leave users uneasy when the assistant provides limited feedback or lacks visible options. The opacity of underlying algorithms in these situations can erode trust, as users are left uncertain about the reasoning behind specific recommendations *(Luchtenberg et al., 2025).*

Additionally, perceived ease of use and usefulness remain central mediators of user experience, as outlined by the Technology Acceptance Model[1] (TAM). When a voice assistant is intuitive and demonstrably beneficial, users are more likely to adopt the technology, feel empowered, and remain loyal to the associated brand *(Moriuchi, 2019).* This cycle of satisfaction and engagement serves to strengthen brand affiliation and encourage sustained usage *(Flavián et al., 2022).*

Personalization and localization further influence perceptions of control and satisfaction. When voice assistants adapt to individual preferences, communicate in a user's local language or dialect, and provide contextually relevant recommendations, the overall interaction feels more attuned to user needs *(Acikgoz et al., 2023).* This not only enhances emotional connection but also reinforces a sense of agency and command, which can deepen loyalty to the brand.

Transparency remains a critical concern. Users consistently report higher satisfaction not only when a voice assistant performs well, but also when the logic guiding its suggestions is accessible *(Guha et al., 2023).* The opportunity to follow up, inquire about rationale, or receive multiple options fosters a sense of fairness and collaboration. As noted by Acikgoz et al. *(2023),* the quality of user experience is determined not solely by technical performance but also by perceptions of agency, justice, and partnership.

[1] An information systems theory that explains how users adopt and engage with new technologies.

1.5. The Role of Artificial Intelligence and Personalization

1.5.1 How AI Personalizes the Shopping Experience

Artificial intelligence is at the heart of how modern voice assistants offer personalization, shifting them from basic tools into more nuanced, context-aware companions. These AI-driven systems leverage machine learning and natural language processing to collect data from user interactions - not just explicit commands, but also subtler cues such as user location, time, emotional tone, and behavioral patterns. This level of data integration enables the assistant to generate highly relevant, individualized recommendations, making user interactions feel tailored and responsive.

A key process here is dynamic user modeling. As users continue to interact with voice assistants, the systems build evolving profiles that reflect changing preferences and behaviors. For instance, if a user consistently purchases vegetarian products or demonstrates brand loyalty, the assistant recognizes these patterns and adjusts its recommendations over time. This continuous learning reduces the user's cognitive load, streamlining routine decisions and building trust and familiarity. As Flavián et al. *(2022)* note, being recognized in this way tends to strengthen emotional connection and engagement.

Natural language processing is especially central to the fluidity of these interactions. Rather than relying solely on rigid, keyword-based commands, users can speak conversationally. The assistant interprets these more natural requests - such as "What should I make for dinner?" - and responds with contextually appropriate suggestions. This conversational flexibility mimics human interaction, making the technology more intuitive and socially engaging. As use continues, the assistant's understanding grows more sophisticated, enhancing both personalization speed and accuracy.

This high degree of personalization often leads users to delegate routine decisions to the assistant, particularly in low-involvement contexts. Over time, consumers may become comfortable letting the assistant handle certain tasks or product selections. This habitual delegation not only streamlines daily routines but can also increase brand loyalty, as users become accustomed to both the assistant and the brands it recommends *(Moriuchi, 2019).*

From a marketing perspective, these developments offer significant opportunities. AI-powered voice assistants allow brands to micro-target consumers in real time, aligning promotions with individual contexts in a way that feels less intrusive and more relevant. Brands can deliver timely reminders, offers, or prompts based on historical data and behavior, deepening engagement without requiring users to search or browse manually. However, the effectiveness of these strategies depends on user trust and perceived value.

Importantly, the extent of personalization depends on the user's willingness to share personal information. Users who provide access to purchase histories, preferences, or payment information enable the assistant to offer tailored experiences, such as one-click reordering or personalized promotions. Yet, this exchange carries risks; overly opaque or intrusive personalization can raise concerns about privacy, surveillance, or manipulation. Therefore, ethical design—including transparency about data usage, robust privacy controls, and clear user agency—is essential for maintaining trust.

Finally, personalization also contributes to emotional resonance. Voice assistants that recall past interactions, anticipate user needs, and communicate in a warm, human-like manner foster a sense of intimacy and social presence. This dynamic encourages sustained engagement and greater user satisfaction, particularly when the assistant relieves decision fatigue or provides reassurance in uncertain situations *(Acikgoz et al., 2023).*

1.5.2 The Value of Voice Assistant Recommendations in Consumer Behavior

Voice assistants have assumed a notably influential position in shaping consumer decision-making, owing to their unique method of interaction - namely, conversational delivery, perceived impartiality, and real-time personalization. In contrast to traditional digital platforms, which present a multitude of options for users to compare and evaluate, voice assistants engage in a linear, sequential exchange, typically offering one suggestion at a time. This approach, reminiscent of a guided conversation rather than a static selection menu, heightens their persuasive impact within the decision process.

Empirical studies indicate that consumers frequently interpret recommendations from voice assistants as more neutral and genuinely supportive than those delivered through conventional web interfaces. The language adopted by these systems - phrases such as “I suggest” or “You might like” - conveys an impression of personal advice, which fosters trust and strengthens the perceived relationship between user and assistant *(Flavián et al., 2022), (Acikgoz et al., 2023).* When the assistant’s tone is warm and empathetic, users are especially likely to perceive the recommendations as tailored rather than generic or commercially motivated.

A further advantage of voice assistants is their capacity to mitigate cognitive overload. In today’s digital environment, consumers are often confronted with an overwhelming number of choices, particularly in the early phases of the purchase journey. Voice assistants address this by narrowing the set of options and providing contextually relevant, streamlined suggestions. This not only expedites decision-making in routine or low-stakes situations but also enhances users’ sense of control and satisfaction by rendering the process more manageable. Additionally, the integration of personalization and localization features significantly increases the perceived value of recommendations provided by voice assistants. When these systems leverage prior user behavior, stated preferences, or local context, users tend to feel recognized and understood. Context-aware and timely suggestions—such as recommending a nearby store or a product relevant to the user’s habits—are regarded as actionable and meaningful. Such relevance is central to perceived usefulness, an important component of the TAM, which in turn influences trust and long-term adoption.

Nevertheless, the trust associated with voice assistants is not without complexity. Because these systems typically offer a single suggestion at a time, users may accept initial recommendations without critically considering their origins or weighing alternatives. At times, these suggestions may be shaped by undisclosed commercial partnerships or algorithmic biases that favor sponsored content. The opacity of these mechanisms raises ethical concerns regarding fairness, impartiality, and the integrity of consumer decision-making. Over-reliance on voice assistants, particularly when users presume neutrality, may subtly compromise consumer autonomy.

Despite these challenges, consistently accurate and helpful recommendations from voice assistants foster robust psychological connections between users, the technology, and the associated brand. Literature suggests that such engagement, driven by high-quality recommendations, is a more effective predictor of brand loyalty than conventional loyalty programs. Engaged users are more likely to engage in repeated interactions, advocate for the brand, and integrate these technologies into their daily routines, both for transactional and informational purposes.

1.6 Voice Assistants and Consumer Engagement

1.6.1 Effects on Consumer Engagement and Loyalty

Voice assistants—such as Amazon's Alexa and Apple's Siri—are fundamentally reshaping how consumers engage with brands, transforming digital interactions from isolated, transactional events into continuous, emotionally resonant conversations *(Flavián et al., 2022).* Rather than navigating traditional visual interfaces or manually inputting text, users interact naturally through spoken language, integrating these exchanges seamlessly into their daily routines and across multiple devices *(Wohr, 2025).* This evolution signifies a new era of consumer engagement: dynamic, habitual, and notably personal *(Acikgoz et al., 2023).*

A central driver of this shift is the voice assistant's capacity to deliver interactive, responsive, and personalized experiences. Users make requests, pose questions, and receive immediate, tailored feedback, cultivating a unique sense of co-presence and reciprocity. When voice assistants employ natural language, adopt an appropriate tone, and accommodate cultural nuances, these interactions increasingly mirror human conversations *(Guha et al., 2023).* This resemblance often leads users to perceive such exchanges as meaningful, reinforcing both emotional connection and behavioral engagement with the assistant—and, by extension, the brand it represents *(Moriuchi, 2019).*

This heightened engagement is maintained through habitual use. Whether checking the weather, managing their schedules, or making purchases, consumers regularly interact with voice assistants as part of everyday life. Repeated positive experiences foster familiarity and trust, positioning the assistant as more than a mere tool, but rather as an integral component of the user's environment *(Simss, 2019).* Empirical research substantiates this; sustained, positive interactions with voice assistants foster brand attachment—a psychological bond that predicts loyalty, particularly within digital contexts where emotional resonance can otherwise be elusive *(Acikgoz et al., 2023).*

Anthropomorphic features of voice assistants, such as their ability to converse, respond, and recall prior interactions, play a pivotal role in this process. When users perceive an assistant as intelligent, reliable, and socially attuned, they are more likely to experience a sense of being understood and supported *(Nass, Steuer, 1994), (Guha et al., 2023)*. This relational closeness extends beyond the device itself, positively influencing perceptions of the parent brand (such as Amazon or Apple). Satisfying interactions with the assistant promote continued use, advocacy, and emotional loyalty toward the brand *(Moriuchi, 2019)*.

Empirical evidence further underscores the strength of these relationships. Studies indicate that perceived usefulness and ease of use significantly increase engagement, which, in turn, robustly predicts brand loyalty—often surpassing traditional factors like satisfaction or product quality *(Flavián et al., 2022)*. Engagement encompasses not only frequency of use but also cognitive, emotional, and behavioral investment in the brand relationship. Whether in transactional contexts (e.g., shopping) or non-transactional contexts (e.g., information seeking), engagement exerts a statistically significant effect on loyalty.

1.6.2 Implications for Loyalty Strategies Using Voice Assistants

The emergence of voice assistants marks a pivotal shift in the strategies brands use to cultivate consumer loyalty, moving far beyond the transactional models that have historically dominated digital marketing. Rather than focusing solely on rewards or repetitive exposure, VA-enabled loyalty initiatives emphasize contextual relevance, personalization, emotional engagement, and the continuity of conversation. In this framework, voice assistants become more than mere communication channels; they act as facilitators of relationships, enabling dynamic and evolving interactions that foster deeper loyalty.

A central factor driving loyalty within this context is personalized conversational engagement. Utilizing artificial intelligence and machine learning technologies, VAs are capable of analyzing individual user behaviors, preferences, routines, and even emotional nuances. As a result, they can provide tailored recommendations and reminders that are contextualized to the user's unique circumstances. These interactions are typically perceived as supportive rather than intrusive, thereby increasing the likelihood of user acceptance and strengthening positive brand associations. For instance, a VA may suggest reordering a preferred product or highlight a relevant promotion based on previous interactions—subtly integrating brand engagement into everyday routines.

Localization further enhances the efficacy of VAs in building loyalty. When assistants adapt to local languages, colloquialisms, and cultural contexts, users experience a heightened sense of recognition and respect. This increased relevance is particularly impactful in nontransactional settings, such as information requests or customer support. By providing culturally and contextually sensitive responses, VAs deepen the psychological connection between the user and the brand, fostering loyalty that extends beyond mere transactional incentives.

The development of relational loyalty is also facilitated by VAs' ability to foster trust and emotional continuity. When assistants remember previous interactions, recognize returning users, and adjust to evolving user behaviors or moods, they create a sense of familiarity and care. This empathetic responsiveness, particularly when communicated through a conversational and warm tone, strengthens emotional engagement and brand attachment over time. Users may begin to perceive the assistant as a reliable, intelligent companion, and by extension, develop a stronger loyalty to the brand it represents.

However, effective implementation of these strategies requires that marketers ensure both technical proficiency and ethical integrity in their use of VAs. As assistants increasingly collect and act upon personal data, transparency is crucial. Users must be informed about how their data is utilized, the rationale behind the assistant's recommendations, and how they may control or revoke permissions. Brands that demonstrate responsible data stewardship and respect user autonomy are more likely to earn and sustain consumer trust—an essential component of long-term loyalty.

It is equally important to recognize that loyalty in the context of VAs is dynamic, not static. It is co-created and continually reshaped through ongoing interactions. Brands must be committed to iterative improvement, leveraging user feedback, monitoring engagement metrics, and refining personalization and localization approaches. Regular updates to the VA's capabilities and responsiveness help maintain relevance and prevent user disengagement.

Furthermore, VAs can serve as potent advocates for brand expansion. Satisfied and engaged users are more likely to recommend both the assistant and the associated brand to others. Facilitating this form of organic advocacy—whether through referral mechanisms or social sharing—can amplify the impact of loyalty initiatives, effectively transforming loyal users into brand ambassadors.

1.7 Reflections and Managerial Implications

1.7.1 Reflections on the Ongoing Transformation of Digital Marketing

The emergence of voice assistants has fundamentally transformed the landscape of digital marketing, marking a significant departure from previous communication modes between brands and consumers. Rather than representing a mere technological upgrade, this shift signals a profound change in consumer behavior and brand strategy. Where users once relied on typed searches and manual scrolling to find information, the prevalence of voice-enabled devices has normalized the expectation of immediate, spoken responses to direct queries *(Wohr, 2025).* This transition from a "search and select" paradigm to an "ask and receive" model places considerable pressure on brands, as voice assistants typically deliver a single, definitive recommendation, intensifying competition for consumer attention and necessitating precise, natural language optimization *(Simss, 2019), (Flavián et al., 2022).*

Crucially, the integration of voice assistants into everyday routines introduces not just technical, but also social and psychological dimensions. Increasingly, consumers perceive these systems as more than mere utilities; they are seen as companions or advisors, fostering trust and emotional engagement through conversational immediacy *(Moriuchi, 2019), (Guha et al., 2023).* This dynamic compels brands to move beyond episodic, campaign-driven outreach and adopt continuous, personalized engagement strategies that become ingrained in daily life. Advanced personalization, powered by artificial intelligence, enables voice assistants to anticipate user preferences and deliver tailored recommendations, enhancing both engagement and conversion *(Acikgoz et al., 2023).*

The "voice-first" paradigm also challenges traditional branding strategies. Absent visual cues, the auditory characteristics of a brand—tone, cadence, persona—become vital elements in communicating credibility and fostering connection *(Koschate-Fischer et al., 2024).*

Concurrently, marketing practices are shifting away from discrete campaigns toward the establishment of habitual, seamless touchpoints that are integrated into consumers' everyday routines.

Finally, traditional digital marketing metrics prove increasingly inadequate for capturing the nuances of voice-driven engagement. Accurate assessment now demands sophisticated analytics, including voice interaction tracking, sentiment analysis, and behavioral modeling. An interdisciplinary approach, drawing on technology adoption research, privacy scholarship, and behavioral theory, is essential for gaining a comprehensive understanding of consumer engagement in the voice-enabled marketplace *(Epley, Schroeder, 2014).*

1.7.2 Managerial Challenges and Opportunities in the Voice Era

Voice assistants have fundamentally reshaped managerial strategies, presenting both significant opportunities and distinct challenges. These technologies now enable brands to integrate into consumers' daily routines in unprecedented ways *(Simss, 2019)*. A substantial strategic advantage exists for brands that establish an early and prominent presence within these voice ecosystems. Given that voice assistants often provide only one or two recommendations per user query, occupying a default position becomes exceptionally advantageous *(Flavián et al., 2022)*. Brands that proactively invest in voice search optimization, develop custom skills, and form partnerships with platforms like Amazon Alexa or Google Assistant are positioned to capture a considerable portion of consumer attention *(Wohr, 2025).*

Personalization, powered by the analysis of user data and contextual information, is another pronounced benefit. Voice assistants are capable of delivering tailored suggestions, thereby increasing consumer satisfaction, fostering loyalty, and potentially improving conversion rates *(Moriuchi, 2019).* Yet, this strength introduces vulnerabilities. Over-personalization or lack of algorithmic transparency can erode user trust and generate ethical concerns *(Dellaert et al., 2020).* Managers must, therefore, balance personalization initiatives with a commitment to transparency, ensuring users comprehend the basis for recommendations and retain autonomy over their data *(Acikgoz et al., 2023).*

The design of voice interactions introduces further complexity. Achieving effective outcomes requires not only technical proficiency but also expertise in conversational design and behavioral psychology—skills that may be underdeveloped within traditional marketing teams *(Luchtenberg et al., 2025).* The construction of a brand's voice, including tone, accent, and prosody, must be carefully aligned with both the organization's identity and consumer expectations *(Guha et al., 2023).*

Beyond facilitating isolated transactions, voice assistants can contribute to long-term loyalty strategies. Brands may provide value-added features such as post-purchase support, cross-selling opportunities (e.g., recipes linked to grocery purchases), or exclusive rewards accessible via voice platforms *(Moriuchi, 2019).* In this manner, voice assistants evolve from transactional tools to ongoing assets in cultivating consumer relationships. Nevertheless, considerable challenges persist. Ethical risks tied to algorithmic opacity and restricted user agency raise concerns about reputation and regulatory compliance *(Flavián et al., 2022).*

Furthermore, designing voice systems that cater to diverse accents, languages, and cultural norms is essential for global applicability *(Koschate-Fischer et al., 2024).* Addressing these demands necessitates interdisciplinary collaboration, iterative development, and authentic user feedback. Measurement methodologies must also evolve. Traditional metrics like impressions and clicks do not adequately capture the nuances of voice interactions. Instead, managers should assess engagement quality, speed of decision-making, user trust, and sentiment—metrics that better reflect shifting consumer dynamics in a voice-first environment *(Epley, Schroeder, 2014).*

The rise of voice assistants compels managers to rethink strategic approaches, technological infrastructure, and required competencies. Success in this domain will depend upon innovation, ethical rigor, transparency, and the cultivation of emotionally intelligent user experiences *(Luchtenberg et al., 2025).* Brands that adapt swiftly and responsibly are best positioned to secure a competitive edge and shape the trajectory of digital marketing

1.8 Relevant Metrics

There are some metrics that support the study's psychological finding in managerial relevance:

- **Marketing ROI**: Quantifies incremental revenue per euro spent attributable to a given interface or channel; it translates modality effects – in this case - screen's higher satisfaction/quality, trust, engagement - into economic impact and guides budget allocation between voice and screen.

- **Customer Lifetime Value**: Estimates the net present value of a customer over time. It links experience variables identified in the study - trust, engagement, perceived quality - to long-term outcomes beyond a single session.
- **Lead metrics** (cost per lead, lead-to-customer conversion, lead quality): Capture top-of-funnel efficiency; they clarify where voice can reduce early friction and gather qualified prospects while screen drives evaluation and commitment, enabling a coherent multimodal strategy.

Chapter 2: Framework Analysis & Development of Research Hypotheses

2.1 Independent Variable: Interaction Mode (Voice vs. Traditional)

2.1.1 Definition and theoretical background of interaction mode

In the context of this thesis, “interaction mode” refers specifically to the primary channel through which consumers engage with, inquire about, and act upon information within digital marketplaces. To be clear, this isn’t just about the medium—whether you’re using a voice assistant like Alexa, Siri, or Google Assistant, or you’re navigating a traditional website or app. The distinction between these channels is far from trivial. It fundamentally shapes the cognitive, social, and experiential dynamics of decision-making in ways that are often underappreciated. To understand why this matters, it helps to zoom in on three foundational perspectives.

First, let’s talk about media richness and social presence. Voice-based interfaces are not merely about auditory delivery; they’re packed with subtle cues—tone, rhythm, pauses, even emotional inflection. These paralinguistic features heighten the sense of social presence, making the interaction feel more like a conversation with an actual person than simply reading static text on a screen. This isn’t just a matter of comfort; research suggests that heightened social presence can encourage trust and engagement, sometimes even nudging users toward decisions they might not make in less “human” environments *(Flavián et al., 2022)*. On the other hand, screen-based interfaces prioritize information density and permanence. You can review, compare, and revisit information at your own pace, which supports more deliberate and less socially influenced decision-making.

Secondly, the communication modality - whether you’re speaking or typing - directly shapes cognition. Speaking is generally spontaneous and flows sequentially, while writing (and reading) tends to be more deliberate, organized, and open to revision *(Epley, Schroeder, 2014)*.
This distinction isn’t minor. When people speak, they’re more likely to rely on intuition, gut reactions, and emotional responses. In contrast, when they engage with text, they have more room to analyze and reflect. As a result, the same individual might make different decisions depending

on whether they're interacting through voice or screen—sometimes faster, sometimes more impulsive, but not always with greater satisfaction.

Third, consider the choice architecture in AI-mediated environments. Voice assistants typically present options one at a time, or in very limited sets. In effect, they act as gatekeepers or curators, subtly steering users toward certain outcomes by filtering and sequencing information. This is a stark contrast to screens, where users can view multiple alternatives simultaneously, easily compare features, and exercise greater autonomy in their exploration *(Dellaert et al., 2020).* This difference is not simply about convenience; it fundamentally alters exploration versus acceptance dynamics. With voice, exploration is often constrained by design, which can speed up decision-making but might restrict perceived autonomy.

Synthesizing these perspectives, it becomes clear that interaction mode is not a mere technical detail—it is a structural property of the decision environment itself. The mode by which users engage with information systematically shapes not just the speed of their decisions, but also their perceptions of quality, satisfaction, and even trust in the platform. In other words, the interface is not neutral; it actively participates in shaping the consumer's cognitive and emotional journey.

If anything, these insights suggest that as digital marketplaces continue to integrate voice-based and AI-mediated interfaces, designers and policymakers must pay close attention to the subtle but powerful ways these modalities influence choice. Ultimately, the "how" of interaction is just as consequential as the "what," reshaping the contours of consumer experience in the digital age *(Flavián et al., 2022), (Dellaert et al., 2020), (Acikgoz et al., 2023).*

2.1.2 Why voice feels different: immediacy, agency, and curation

Voice assistants are now deeply woven into the fabric of daily life—popping up in smartphones, cars, and even the most unassuming home gadgets. Their draw is obvious: they offer a level of convenience that's hard to beat. With just a simple voice command, tasks that once required physical input—like typing or tapping—can be handled instantly. This hands-free access is especially valuable when multitasking is necessary or when physical movement is restricted, such as while driving or cooking.

A core feature driving the adoption of these assistants is their ability to provide immediate responses, often presenting a small set of options curated by algorithms. This immediacy, paired with the perception that the assistant "knows" what the user wants, leads many individuals to accept the first suggestion, particularly for routine or low-stakes decisions (think reordering dish soap or checking the weather). While this streamlines the decision-making process and reduces cognitive load, it can also limit the breadth of options considered, essentially compressing the comparison process that users might otherwise engage in if left to their own devices.

Another fascinating aspect is the anthropomorphic quality of these assistants. Their voices are designed to sound warm, competent, and personable—attributes that encourage users to trust them and perceive them as genuinely helpful. There's a psychological effect at play: people are more likely to feel a sense of rapport and comfort, even though the "personality" is entirely artificial. This social dimension can enhance the user experience, fostering a sense of partnership with the technology.

However, alongside these clear benefits, significant concerns arise. The decision-making process of the voice assistant is often opaque; users rarely understand why a particular option is presented over others. This lack of transparency can erode trust if users begin to suspect bias or manipulation, especially in contexts involving commercial recommendations or sponsored results. Furthermore, the very convenience that makes voice assistants appealing can also undermine the user's sense of agency. When alternatives are buried or hard to access, individuals may feel a loss of control over their choices, ultimately raising important questions about autonomy and informed decision-making in the age of algorithmic curation.

2.1.3 Operationalization in this study (voice simulation vs. screen interface)

To rigorously examine the impact of delivery mode on participant responses, I employed a between-subjects experimental design with careful attention to methodological detail and control. My manipulation consisted of two primary conditions:

- Voice Condition: In this simulated scenario, participants interacted with an assistant that communicated recommendations. Participants began by issuing a simple directive, following which the assistant offered a primary recommendation. If participants requested further options, up to two additional suggestions were delivered sequentially. This interaction was structured to parallel the stepwise information delivery characteristic of mainstream voice assistants, where recommendations typically unfold one at a time in response to user prompts.

- Traditional (Visual) Condition: Here, the interface simulated a more conventional digital shopping experience, presenting 3 product options simultaneously using a mock product card layout. Each option included succinct details such as price and user ratings. Participants were free to scroll, compare alternatives, and make selections at their own pace, emulating the parallel information processing typical of e-commerce platforms *(Dellaert et al., 2020).*

To validate that the manipulation was both salient and effective, I included several manipulation checks. Participants were asked to report their perceived mode of interaction (e.g., "I interacted mainly by voice/screen"), the conversational nature of the exchange, and whether they perceived information delivery as simultaneous or sequential—items adapted from established measures *(Epley, Schroeder, 2014), (Flavián et al., 2022).*
Crucially, I controlled for contextual confounds by holding the device context constant across all vignettes—participants imagined using the assistant in the same environment (e.g., always at home, on a phone, or in a car). This precaution minimized the risk of conflating delivery mode effects with contextual factors such as convenience *(Simss, 2019).*

Collectively, these methodological choices - detailed manipulation checks, stringent context and persona controls, and balanced presentation of options - were designed to isolate the effects of modality on participant perceptions and decisions. Through this approach, our study aims to contribute nuanced, empirically grounded insights into the psychology of human-technology interaction within both auditory and visual recommendation environments.

2.1.4 Boundary conditions and potential confounds

- Individuals with a history of frequent VA use tend to navigate tasks with notable efficiency, seemingly irrespective of situational variables *(Acikgoz et al., 2023).* To account for this, my methodology includes a measure of participants' prior VA experience, which is incorporated as a covariate in my analyses to control for habituation effects and ensure more precise interpretation of outcomes.

- The nature of the task itself appears to significantly modulate the impact of VA usage on decision speed. Routine, low-stakes decisions - such as setting reminders or managing shopping lists - are particularly conducive to accelerated completion via voice interface. In contrast, when faced with novel or high-stakes choices, users may exhibit greater deliberation, thereby diminishing or even negating the speed advantage typically associated with VA interactions *(Dellaert et al., 2020), (Simss, 2019)***.** This distinction underscores the importance of contextual factors in evaluating VA efficacy.

- Transparency, often operationalized as providing succinct rationales for recommendations, serves as a critical lever for fostering both user autonomy and acceptance. Brief explanations such as "here's why this option was suggested" can meaningfully enhance perceived control, which, in turn, may increase the likelihood of users following VA recommendations *(Flavián et al., 2022), (Luchtenberg et al., 2025).* This suggests that even minimal efforts to clarify system logic can yield substantial gains in user trust and engagement.

- Finally, I systematically assess perceived cognitive load to disentangle genuine improvements in task speed from potential artifacts of rushed or superficial processing. By capturing subjective reports of effort, I aim to ensure that observed efficiencies are not merely the byproduct of users feeling hurried or pressured, but rather reflect authentic enhancements in process fluency. This allows for more nuanced interpretations of the relationship between VA use and task performance.

2.1.5 Hypothesis for interaction mode: decision speed

Interactions within voice environments tend to speed up decision-making, and honestly, it's not hard to see why. Unlike traditional digital interfaces that bombard users with a million options and endless scrolling, voice assistants cut straight through the noise. The immediacy of conversation - just saying what you want out loud - minimizes the time spent dithering over choices *(Flavián et al., 2022), (Dellaert et al., 2020).* There's a kind of "get to the point" efficiency baked into the whole process.

What's really fascinating is how voice queries push users toward expressing clear goals. Instead of passively browsing, people tend to blurt out what they need: "Order me coffee," or "Find me the nearest gas station." This outcome-focused style not only shortens the decision path, but also leaves less room for second-guessing or distraction *(Epley, Schroeder, 2014).* It's like the system encourages users to be decisive, prioritizing immediate results over exhaustive comparison.

And there's a reality here - there's actually a cognitive relief in not having to weigh every possible alternative. The sequential, conversational format means you're less likely to get stuck analyzing every option, which lowers the mental load. In academic terms, voice environments seem to streamline cognitive processing by structuring choices as a dialogue, rather than an overwhelming display of options. This shift in interaction design can fundamentally alter how consumers approach decisions, nudging them toward faster, more confident actions. Thus:

H1. Customers interacting through voice channels will exhibit higher decision-making speed compared to customers interacting through traditional digital channels.

2.2 Dependent Variables: Decision Outcomes

2.2.1 Decision speed - construct definition and measurement

Decision speed, at its core, refers to the time taken from the moment an individual is presented with a set of options or a recommendation to the point where a final choice is made. To rigorously assess this, I employ a two-pronged approach: I incorporate a subjective measure, asking participants to self-assess how quickly they made their decision, using a 1–7 scale ranging from "not at all quickly" to "extremely quickly." This dual-method approach helps capture both actual behavior and perceived experience, offering a more nuanced understanding of decision latency.

What's particularly interesting - backed by recent research - is how the modality of interaction influences this decision-making process. There's growing evidence that voice-based interfaces, in contrast to traditional written or click-based formats, tend to streamline the decision pathway. When individuals vocalize their responses, they are less likely to engage in extensive exploration of every possible option. This vocal modality appears to encourage more rapid acceptance of options, effectively reducing the time taken to reach a choice *(Flavián et al., 2022), (Dellaert et al., 2020).*

One possible explanation ties back to dual-process theories of cognition: speaking out loud seems to activate what's known as "System 1" thinking - an intuitive, fast, and heuristic-driven process - rather than the slower, more analytical "System 2" mode that often predominates during written deliberation *(Epley, Schroeder, 2014).* In practical terms, this means that voice interfaces might not just make the process faster, but could also shape the kind of choices people make, potentially promoting decisions that are more instinctive rather than deeply contemplated.

This dynamic has significant implications for the design of decision environments and digital interfaces. If the goal is to facilitate quicker decisions - say, in high-velocity e-commerce contexts or when cognitive overload is a concern - integrating voice as a primary modality could prove particularly effective. At the same time, it's worth noting that speed isn't always synonymous with quality; the very factors that accelerate decision-making might also lead to less deliberate, and possibly less optimal, choices. Therefore, understanding the interplay between modality, decision speed, and outcome quality remains an important area for ongoing research.

2.2.2 Perceived quality of choice — definition and assessment

Perceived quality of choice, in essence, revolves around an individual's sense that their selected option is not only suitable but also the result of a thoughtful and informed process. This evaluation is far from superficial—it's shaped by a cluster of factors, such as how confident the person feels about their decision, whether they believe they've actually considered a wide enough array of alternatives, and if the overall decision-making process itself seemed appropriate for the context.

To put it more concretely, researchers often attempt to quantify these perceptions through tools like confidence scales (usually ranging from 1 to 7), and by asking participants to reflect on whether they truly explored enough possibilities or if the path they followed to reach a decision seemed fitting for the situation at hand. These metrics, while inherently subjective, are crucial for understanding the psychological underpinnings of satisfaction and regret in consumer and user experiences.

The situation gets even more layered when I factor in the influence of "voice" - that is, systems or interfaces that communicate choices through spoken or written prompts. Studies like *Flavián et al. (2022)* suggest that when these systems manage to deliver recommendations that feel personal and contextually relevant, individuals are likely to perceive their choices as higher quality. Personalization, here, acts as a signal that the system "gets" the user, which can foster a sense of trust and satisfaction with the decision.

However, the dynamic isn't always so straightforward. *Dellaert et al. (2020)* found that when users experience a lack of control - perhaps because the system doesn't make the range of available options transparent, or because the process feels too automated or opaque - perceptions of quality can quickly deteriorate. People generally want to feel like they're in charge, not just passive recipients of whatever the algorithm spits out. If that sense of agency is missing, even the most personalized recommendations might fall flat.

The perceived quality of a choice is not a fixed attribute; it's highly contingent on whether individuals feel empowered during the decision-making process. The more control and transparency they experience, the more likely they are to judge their choices as sound and wellfounded. This relationship between perceived control and choice quality is a critical consideration for designers of decision-support systems, and one that the literature continues to explore in greater depth, particularly in subsequent hypotheses and studies.

2.2.3 Customer satisfaction — definition and assessment

Satisfaction, in this context, isn't just a single fleeting feeling - it's a multi-layered evaluation of both how a decision was made and whether the end result actually hits the mark. It's not enough to just look at whether someone ended up happy with their choice (though, sure, outcome satisfaction matters). You really have to dig into how the decision-making experience felt. Was the process smooth, or did it feel like wading through mud? Did the steps make sense, or was it a confusing mess? That's where process satisfaction comes in. And then, pulling it all together, there's overall satisfaction, which we typically capture by asking folks to slap a number on it - one to seven, simple enough, but it gives a sense of the big picture.

When we talk about what drives satisfaction, several key factors keep showing up. For one, if the process is seamless - meaning, decisions don't get bogged down by unnecessary hurdles or confusion - people are a lot more likely to walk away content. Perceived helpfulness is another big player; if users feel that the assistance they receive genuinely improves their experience, satisfaction shoots up. Social intelligence matters too, especially with virtual assistants. If the interaction feels intuitive, responsive, and, dare I say, a bit human, people respond positively *(Moriuchi, 2019), (Acikgoz et al., 2023), (Luchtenberg et al., 2025).*

On the flip side, there are definite pitfalls. If a system comes off as too controlling - like it's twisting your arm or nudging you toward a choice in a way that feels unnatural - satisfaction drops. Similarly, if there's a lack of transparency, users get wary. They want to know why certain options are being suggested, or how their information is being used. When these elements are missing, trust erodes, and so does satisfaction *(Flavián et al., 2022), (Dellaert et al., 2020).*

Hence, satisfaction with decision-making isn't just about the final outcome. It's a holistic assessment shaped by both the journey and the destination, influenced by factors like friction, helpfulness, and transparency. These insights are particularly relevant as virtual assistants and AI-driven systems become more prevalent, making it critical to design interactions that feel both supportive and trustworthy.

2.2.4 Mechanisms linking mode to outcomes

- Preference fluency. When the mode of communication fits the user's natural processing style - say, spoken word for those who rely on intuition, or written text for more analytical types - the act of choosing just feels more seamless. This isn't just a small thing, either; there's a psychological comfort that comes from feeling "in sync" with the medium, which in turn boosts user confidence and satisfaction. When people feel like the process matches how their mind works, they're much more likely to trust their own decisions, and decision fatigue drops off considerably *(Epley, Schroeder, 2014).* Notably, this has implications for designing digital services: aligning interface with cognitive preference isn't just a UX bonus, it's central to how users assess the legitimacy of their own choices.

- Curated consideration sets. Voice interfaces typically provide a limited menu of options, which undeniably reduces complexity and cognitive load. While this streamlining can be beneficial, helping users avoid the paralysis of endless scrolling, it may also create a sense of missed opportunities or incompleteness. Users might doubt the thoroughness of the recommendations, especially if the rationale behind the curation isn't communicated clearly *(Flavián et al., 2022), (Dellaert et al., 2020).* To mitigate this, it's critical for designers to pair concise option sets with transparent explanations. Otherwise, the perceived simplicity can backfire, leading to skepticism about whether the best options were actually presented.

- Social presence & warmth. The tone and perceived personality of a digital assistant or interface play a significant role in shaping user perceptions of service quality. When delivery is marked by both warmth and competence, users report higher satisfaction and are more likely to view the service as trustworthy *(Moriuchi, 2019), (Acikgoz et al., 2023).* This social dimension is not trivial; even subtle cues of friendliness or expertise can humanize the interaction, fostering a sense of connection that goes beyond mere transactional efficiency. In other words, the "vibe" of the interaction matters - a robotic or cold delivery can undermine even the most accurate recommendations.

- Transparency & control. Users increasingly expect not just recommendations, but also clear explanations for why those recommendations are made. When interfaces provide accessible rationales and allow users to explore further options if they wish, perceptions of quality and fairness are enhanced *(Flavián et al, 2022), (Luchtenberg et al., 2025).* This transparency is central to building trust, as it reassures users that the system is acting in their best interest, not just pushing predetermined outcomes. Additionally, offering the ability to "see more" or override default suggestions empowers the user, reinforcing a sense of agency and ownership over the decision process.

2.2.5 Measurement strategy and validity safeguards

Speed in this study is captured through two complementary approaches: I record the actual time it takes participants to complete tasks (objective timer data), but I'm also interested in their subjective experience, so I directly ask them how fast they felt the process was. This dual approach recognizes that people's perceptions of speed don't always match the clock, and both can influence their overall evaluation.

When it comes to perceived quality, I dig into several dimensions: confidence (are participants sure about the information they received?), comprehensiveness (does the information feel complete and thorough?), and fit (how well does the information meet the specific needs or context?). These indices let me move beyond basic satisfaction, giving a more nuanced picture of how participants judge the quality of their experience.

Satisfaction is not treated as a monolith; instead, I break it down into satisfaction with the process itself, the outcome, and the overall experience. This layered approach acknowledges that someone might be happy with the end result but frustrated with how they got there, or viceversa. By teasing these apart, I aim for a more granular understanding of user attitudes.

2.2.6 Hypothesis on perceived control as a boundary condition for quality

When individuals perceive that they can actively shape or influence the recommendations an AI system provides—say, by requesting additional alternatives or adjusting selection criteria—they tend to view the end result as more credible and of higher quality. This isn't just speculation; researchers like *Dellaert et al. (2020)* and *Flavián et al. (2022)* have demonstrated that the feeling of control isn't some trivial add-on, but actually central to how people evaluate AI-mediated decisions. Delving deeper, this sense of agency can act as a buffer, enhancing overall satisfaction with the process. If users are allowed to interact, experiment, or even challenge the system's suggestions, their trust in the outcome increases, and the whole experience feels less mechanical and more collaborative.

This effect is especially pronounced with voice-based interfaces, where the default set of choices is typically quite limited—think of smart speakers giving only a handful of responses. In these contexts, perceived control becomes even more essential; the ability to expand the choice set or refine the criteria can transform a potentially frustrating or narrow interaction into one that feels empowering. Ultimately, fostering a sense of user control isn't just a nice-to-have feature—it's foundational to making AI-mediated decisions feel legitimate, meaningful, and tailored to individual needs. Giving these insights, the second hypothesis is:

H2. The relationship between interaction mode (voice vs. traditional) and perceived quality of choice will be moderated by perceived control, such that higher perceived control leads to higher perceived quality in both interaction modes.

2.3 Moderators: Perceived Control and Product Type

2.3.1 Perceived control in technology-mediated choice

Perceived control, at its core, refers to an individual's subjective sense of agency, that underlying belief that they can actively shape conversations, direct interactions, and influence outcomes within a given system. In the context of voice-based technology, this sense of control can be a bit of a double-edged sword. On one hand, the immediacy and simplicity of voice commands ("Do X for me") can empower users, making them feel as though they're effortlessly steering the process. On the other, however, when the system imposes its own sequencing or offers limited, opaque responses, users may feel sidelined or even disempowered *(Dellaert et al., 2020).*

This tension is particularly acute among younger users - especially those from Gen Z - who tend to have heightened expectations for both transparency and personalization. Rather than passively accepting a single, algorithm-selected outcome, these users frequently demand explanations ("Why did you pick this?") and variety ("Show me three alternatives"). The availability of such customization levers not only restores agency but also fosters a greater sense of trust in the technology *(Luchtenberg et al., 2025).* In fact, the ability to request alternatives or probe the rationale behind a recommendation has become a baseline expectation for this demographic.

Moreover, the delivery style of these systems plays a critical role in shaping user perceptions. Socially intelligent communication—marked by politeness, warmth, and responsiveness—signals respect for user autonomy. When users perceive the system as attentive and personable, engagement and satisfaction levels increase *(Moriuchi, 2019), (Acikgoz et al., 2023).*
This suggests that the technical functionality of a voice assistant cannot be separated from its communicative style; both dimensions are integral to the user experience.

To empirically assess perceived control, I rely on a subjective self-report measure. Participants are asked to rate the extent to which they felt in control of the decision-making process, using a 1–7 Likert scale (e.g., "How much control did you feel you had during the choice?"). This approach captures the participant's perceived agency and provides a direct indication of how the interaction mode shapes feelings of autonomy within the decision experience.

2.3.2 Product type: utilitarian vs. hedonic

Utilitarian purchases - think stuff like detergent - are pretty much all about getting things done quickly and efficiently. There's really not much excitement in picking out the best detergent ever, right? For these kinds of products, voice assistants are a natural fit. The whole point is to minimize effort and time: you just tell your device what you need, and it handles the rest. There's no need to scroll through endless lists or read product descriptions. It's the digital version of "grab and go." This lines up with prior research *(Simss, 2019),* which points out that voice interfaces are designed for speed, convenience, and automating routine purchases.

On the flip side, hedonic products - things like fragrances - are all about the experience. These purchases are more personal, expressive, and honestly, a bit indulgent. People actually enjoy browsing, comparing, and discovering new options in these categories. The visual and interactive aspects of screen-based shopping make it a better channel for this kind of exploration. Screens let consumers engage with images, reviews, and a broader sense of discovery, which is crucial for products tied to self-expression and pleasure.

To empirically test these ideas, I manipulate the type of product within our scenarios, using sets of stimuli to make sure they come across as either utilitarian or hedonic. This isn't just theoretical I've validated that the scenarios genuinely reflect the intended product orientation from the perspective of typical consumers. This methodological step is essential because it ensures that any observed effects can be attributed to the product type rather than confounding variables.

Ultimately, this approach allows me to rigorously examine whether there is a true fit between the shopping channel (voice versus screen) and the nature of the product (utilitarian versus hedonic). If my predictions hold, it would provide valuable insight into how digital interfaces can be optimized to better meet the specific needs of different types of shopping experiences.

2.3.3 Dual-moderation logic and analytic plan

Building on the earlier discussion, the interplay between perceived control and product type emerges as a critical factor influencing the relationship between shopping mode and consumer outcomes. This dynamic is far from trivial; rather, it underscores the nuanced ways in which context and consumer psychology shape the overall shopping experience.

In utilitarian purchasing scenarios, think everyday necessities or functional products, the alignment of high perceived control with voice-based shopping methods appears particularly advantageous. When consumers believe they are firmly in control, leveraging voice technology not only streamlines the process, yielding greater efficiency, but also seems to enhance perceptions of quality and satisfaction. This suggests that, in contexts where speed and efficacy are at a premium, technology-enabled autonomy can reinforce positive consumer evaluations.

Conversely, hedonic contexts - where the purchase is more about enjoyment, self-expression, or indulgence - present a markedly different picture. Here, traditional shopping methods, coupled with a strong sense of personal control, seem to better support the desire to browse and compare options. This approach enables consumers to extract greater pleasure from the process itself, thereby sustaining or even elevating perceived quality and satisfaction. In other words, the act of leisurely exploration becomes integral to the value derived from the experience.

To empirically test these propositions, I propose a moderated model in which the mode of shopping (e.g., voice versus traditional) serves as the independent variable, predicting key outcomes such as speed, perceived quality, and satisfaction. Critically, this relationship is hypothesized to be contingent upon both perceived control and product type, which act as moderators. By examining conditional effects and conducting planned contrasts - drawing on methodological precedents established by *Dellaert et al. (2020)* and *Flavián et al. (2022)* - I can more precisely delineate the circumstances under which different shopping modalities optimize consumer outcomes.

Overall, this approach acknowledges that consumer behavior is not monolithic; rather, it is shaped by the complex interplay between technological affordances, psychological perceptions, and the situational context of the purchase. Such insights hold significant implications for both theory and practice, offering guidance for retailers seeking to tailor their strategies to diverse consumer needs and expectations.

2.3.4 Hypothesis on channel–product fit and satisfaction

Research consistently underscores a compelling link between user satisfaction and the alignment of interaction modes with a product's intended purpose. When individuals engage with utilitarian products - think mundane tasks like setting reminders or checking the weather - they overwhelmingly prefer interfaces that prioritize speed and efficiency. In such contexts, simplicity isn't just appreciated; it's expected. A voice assistant that lets users accomplish a task in seconds, with minimal effort or cognitive load, can transform a routine interaction into a genuinely positive experience.

Conversely, hedonic products - those designed for pleasure, exploration, or entertainment - demand a richer, more immersive approach. Users navigating music discovery platforms, digital art galleries, or even online shopping environments often seek dynamic engagement and opportunities for curiosity-driven exploration. Here, an interface that encourages playful interaction or provides personalized recommendations can foster deeper satisfaction and even cultivate brand loyalty.

What's particularly intriguing is how this matching of interaction mode to user intention is supported across both empirical research and industry guidance *(Flavián et al., 2022), (Luchtenberg et al., 2025).* The distinction isn't merely academic; it has real-world implications for product designers aiming to optimize user experience. Ultimately, it becomes clear that a "one-size-fits-all" approach falls short. Instead, tailoring the user journey - whether through streamlined processes for utilitarian tasks or more layered, exploratory pathways for hedonic ones - is crucial for maximizing satisfaction and, by extension, product success. Giving these insights, the thirs hypotheses is:

H3. Customer satisfaction will be higher when interaction mode is aligned with product type, such that voice yields higher satisfaction for utilitarian products, while traditional interaction yields higher satisfaction for hedonic products.

2.4 Practical design implications (cross-cutting)

To enhance both internal validity and practical significance, the approach deliberately aligns with established best practices in the literature, but also pushes further to ensure the voice assistant experience is both empirically sound and user-centric.

First, the design intentionally curbs the risk of over-narrow curation. Rather than boxing users into a limited set of options, I proactively offer accessible alternatives—"Would you like to see a couple more choices?" This approach is grounded in *Dellaert et al. (2020),* who demonstrate that expanding the choice set helps counteract algorithmic tunnel vision and supports user autonomy. In practice, this means users are less likely to feel manipulated or cornered by the system's suggestions.

Second, the assistant's tone was treated as a critical design dimension. Rather than implementing synthetic speech, the study used a written scenario that described the simulation and explicitly specified the assistant's communicative style. Participants were told that the assistant would respond in a socially intelligent manner—polite, warm, and efficient—friendly and respectful but not overly familiar or sales-oriented. This standardized description set a clear expectation for the interaction and kept tone constant across respondents, helping us isolate modality effects. Prior work underscores the relevance of affective tone as a differentiator—especially as users grow more discerning about digital interactions (Moriuchi, 2019), (Acikgoz, 2023).

Third, I tailor prompts to fit the context of the user's needs. For utilitarian tasks—think reordering household essentials—the assistant keeps language concise and action-oriented. For hedonic or exploratory scenarios—like browsing for entertainment—the assistant offers richer descriptions and multiple avenues to discover new content. This context-sensitive prompting, as supported by *Simss (2019),* ensures the system feels relevant and adaptive, rather than one-size-fits-all.
Finally, I employ measurement triangulation to strengthen our findings: combining objective metrics (such as response timers) with subjective self-reports and relevant covariates like prior experience with voice assistants and degree of user involvement. This multi-pronged approach is crucial for capturing the complexity of user experience and ensuring robust causal inference.

Hence, these design guardrails serve a dual purpose: they not only sharpen the rigor of my experimental tests, but also reflect, and even extend, current best practices in voice assistant design. The goal is to bridge the gap between controlled research and the unpredictable realities of real-world deployment, ensuring that our findings have genuine managerial and user-facing relevance.

2.5 The conceptual framework

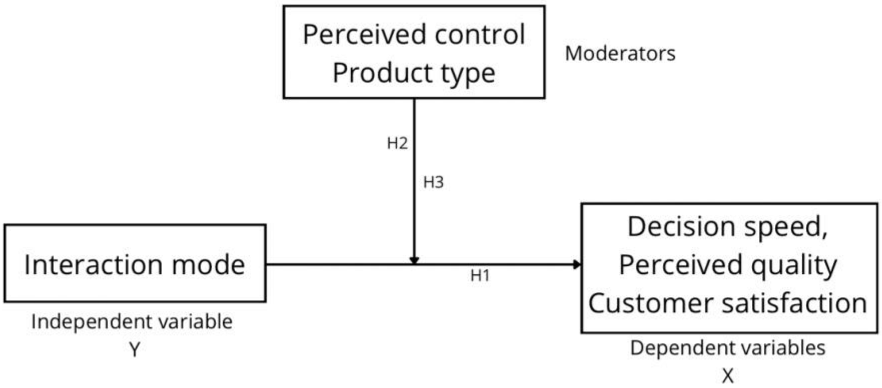
This thesis centers on a conceptual framework that connects interaction mode - specifically, voice-based versus traditional screen-based interfaces - with key outcomes in consumer decision-making. The primary independent variable, interaction mode, distinguishes between voice assistants (like Alexa, Siri, or Google Assistant) and conventional digital interfaces such as websites or mobile apps. This variable essentially captures how the structure of the decision environment and the channel of communication shape consumer experiences.

Three dependent variables are examined: decision speed (how quickly a choice is made), perceived quality of choice (the consumer's sense of making an appropriate, well-informed decision), and overall customer satisfaction (which includes both the decision process and the final outcome).

The influence of interaction mode on these outcomes is not assumed to be uniform. Two moderating factors are considered: perceived control (the extent to which consumers feel agency in the process) and product type (distinguishing functional, utilitarian products from more experiential, hedonic ones). The framework posits that these contextual and psychological variables condition the effects of interaction mode.

Specifically, it is proposed that voice-based interfaces will increase decision speed relative to traditional ones. The impact of interaction mode on perceived quality, however, is hypothesized to depend on the degree of perceived control - higher control is expected to yield higher perceived quality, regardless of interface. Lastly, alignment between interaction mode and product type is anticipated to maximize customer satisfaction; voice interfaces are expected to perform better for utilitarian purchases, while traditional interfaces may be superior for hedonic ones.

This framework offers a structured lens through which to examine how digital communication channels, alongside psychological and contextual moderators, shape consumer decision outcomes. It underpins the empirical analyses conducted in the thesis and advances theoretical understanding of the evolving relationship between technology and consumer behavior.



Chapter III – Methodology and Data Analysis

3.1 Research Design

3.1.1 Research Objectives

The main objective of this study is to investigate how interaction mode influences consumer decision-making processes when using intelligent voice assistants compared to traditional digital interfaces. Specifically, the research aims to assess whether voice interaction leads to faster decision-making (H1), how perceived control moderates the relationship between interaction mode and perceived quality of choice (H2), and whether product type moderates the effect of interaction mode on customer satisfaction (H3). In doing so, the study contributes both theoretically and managerially by clarifying under which conditions voice assistants can enhance or hinder consumer decision-making.

3.1.2 Experimental Design

The research is based on a quantitative experimental methodology with a between-subjects design. Participants were randomly assigned to scenarios in which they interacted either through a voice-based interface or a traditional digital interface. Each participant was exposed to product-related decision tasks that varied by product type. This design allows the identification of causal relationships between interaction mode and consumer responses, while also testing the moderating role of perceived control and product type. Random assignment ensures internal validity and reduces the risk of systematic bias across conditions.

3.1.3 Research Variables

To test the hypotheses, the following variables were included in the study:

- **Independent Variable**
    - Interaction Mode: two experimental conditions, Voice vs. Screen
- **Dependent Variables**
    - Decision speed
    - Customer satisfaction

- Perceived quality of choice

- **Moderators**
  - Perceived Control: measured as the extent to which participants felt in control of the decision process during the simulation.
  - Product Type: two categories, utilitarian product (functional, practical) vs. hedonic product (pleasurable, experiential).

3.2 Questionnaire Structure

3.2.1 Introduction and Preliminary Questions

The questionnaire began with a introductory section aimed to introduce respondents with the objective of the work and to gather fundamental information regarding their prior use of voice assistants and their buying habits when shopping online. Participants were inquired if they had ever used a voice assistant (such as Siri, Alexa, Google Assistant, etc.) and if so, for what reasons. Collecting and grouping all answers, profiles the sample and support external validity, while prior experience with Vas can shape trust, engagement and perceived quality. The most recurring themes were:

- Listening to music / media control.
- Timers/alarms and quick commands (e.g., while cooking, reminders).
- Quick information requests (weather, quick web lookups, directions).
- Hands-free calls/messages (often while driving).
- Basic smart-home control (lights/TV).

There was also a non-negligible group of non-users (“No / I don’t use it”).

Hai mai utilizzato un assistente vocale (Es. Siri, Alexa, Google Assistants)? Se sì, per cosa lo usi di solito?

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido		26	9.3	9.3	9.3
	No	1	.4	.4	9.7
	Acquisti online	1	.4	.4	10.0
	Alexa	5	1.8	1.8	11.8
	Alexa per ascoltare musica	1	.4	.4	12.2
	Alexa per musica di sottofondo	1	.4	.4	12.5
	alexa, ascoltare musica in doccia	1	.4	.4	12.9
	Alexa, per ascoltare musica	1	.4	.4	13.3
	Cercare su Internet quando non ho voglia di scrivere	1	.4	.4	13.6
	chiamare le persone quando ho le mani occupate, cercare canzoni, mettere la musica	1	.4	.4	14.0
	Chiamate vocali in auto	1	.4	.4	14.3
	Chiedere info quando sono di fretta	1	.4	.4	14.7
	Cucinare (timer) e tenere il tempo in generale, rendere domotica la casa	1	.4	.4	15.1
	Di. Solo qualche volta x provarla	1	.4	.4	15.4

Respondents were further inquired regarding factors that most affect their buying decisions when shopping online. Reveals the **decision cues** participants rely on (reviews, comparisons, detailed specs, images, delivery info). These cues are inherently **visual and parallel**, which Screen delivers better than Voice. Most frequently mentioned:

- Reviews / other customers' opinions (the single most cited element).
- Price and value for money.
- Website/seller reliability.
- Product photos and detailed specifications/descriptions.
- Delivery times and reliability.

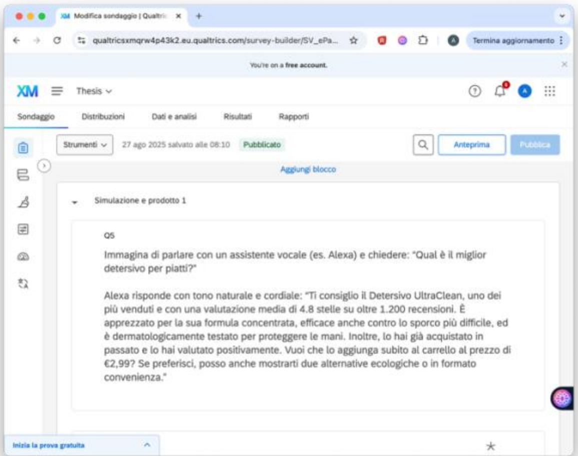
Quando fai acquisti online, cosa ti aiuta di più a prendere una decisione?

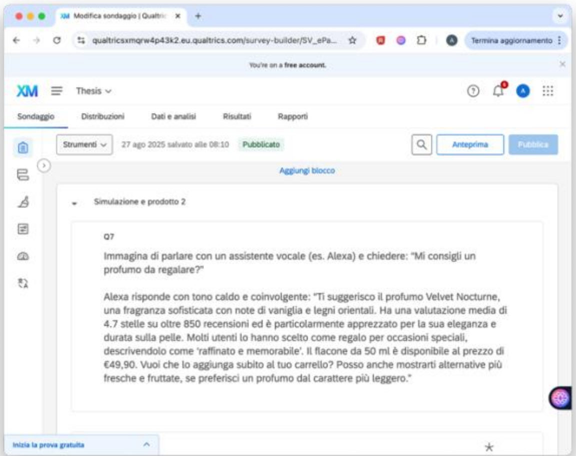
		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido		41	14.7	14.7	14.7
	Non faccio acquisti on line	1	.4	.4	15.1
	A parità di prodotto, il prezzo più economico se la piattaforma mi dà garanzia di sicurezza	1	.4	.4	15.4
	Affidabilità del rivenditore, prezzo	1	.4	.4	15.8
	Affidabilità del sito e recensioni	1	.4	.4	16.1
	affidabilità del sito, mancanza di prodotti simili in negozi fisici in vicinanza, tipologia di prodotto	1	.4	.4	16.5
	Affidabilità del sito, recensioni degli altri clienti, confronto con altri articoli in vendita, rapporto qualità prezzo, modalità di consegna	1	.4	.4	16.8
	Affidabilità sito	1	.4	.4	17.2
	aspettare del tempo, e poi ritornaci sopra con più calma trovaprezzi tiktok	1	.4	.4	17.6
	Attendibilità venditore e prezzo a parità di prodotto	1	.4	.4	17.9
	Avere a disposizione le recensioni di altri utenti	1	.4	.4	18.3

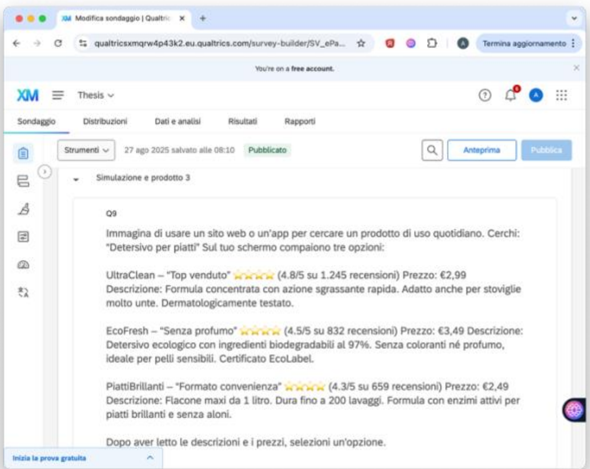
3.2.2 Experimental Simulations

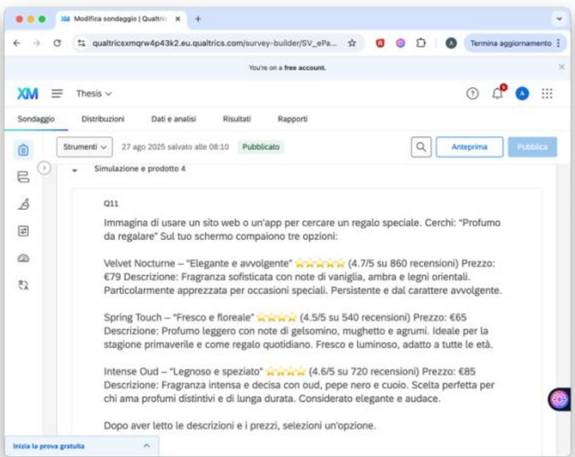
Participants received a collection of experimental situations. These simulations were randomly assigned in such a way that participants would do decision-making tasks either via a voice assistant or a traditional digital interface.

- Voice condition: Participants were asked to picture conversing with Alexa, where a product was recommended, and it inquired if it could add the product to a cart. Two different products were shown:
    - Utilitarian product: dishwashing detergent.
    - Hedonic product: a perfume as a gift to buy.
- Screen condition: Participants were asked to assume viewing a site or app in search of similar products. In this experiment, several choices were shown visually, such as product details, prices, and reviews. The same two categories of products were used.









This experimental design produced four experimental conditions: Voice–Utilitarian, Voice–Hedonic, Screen–Utilitarian, and Screen–Hedonic. Each participant was in one of these conditions through random assignment.

3.2.3 Post-Simulation Evaluation Questions

After completing the simulation, participants answered a series of evaluation questions regarding their experience. These questions measured the core constructs of the study. Responses were collected on a 7-point Likert scale (1 = not at all, 7 = very much).

- Satisfaction with the decision process: the majority of responses were 'almost satisfy' (34%), 'a lot satisfy' (24%) and 'neutral' (20%). Hence, it can be said that the majority of the respondants were sufficiently satisfied during the purchasing experience provided by the simulations.

Tabella delle frequenze

Customer satisfaction

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	1 = Per niente soddisfatto/a	5	1.8	2.2	2.2
	2 = Poco soddisfatto	4	1.4	1.7	3.9
	3 = Leggermente soddisfatto	14	5.0	6.1	10.0
	4 = Neutro	46	16.5	20.1	30.1
	5 = Abbastanza soddisfatto	77	27.6	33.6	63.8
	6 = Molto	55	19.7	24.0	87.8
	7 = Completamente soddisfatto/a	28	10.0	12.2	100.0
	Totale	229	82.1	100.0	
Mancante	Sistema	50	17.9		
Totale		279	100.0		

- Perceived control over the decision: the majority of responses were 'neutral' (23%), 'full control' (17%) and 'moderated-high control' (15%). Hence, it can be said that the majority of respondants perceived almost full control during the purchasing experience provided by the simulations.

Perceived control

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	1 = Nessun controllo	28	10.0	12.3	12.3
	2 = Poco controllo	17	6.1	7.5	19.7
	3 = Moderato-basso controllo	26	9.3	11.4	31.1
	4 = Neutro	52	18.6	22.8	53.9
	5 = Moderato-alto controllo	34	12.2	14.9	68.9
	6 = Molto controllo	33	11.8	14.5	83.3
	7 = Pieno controllo	38	13.6	16.7	100.0
	Totale	228	81.7	100.0	
Mancante	Sistema	51	18.3		
Totale		279	100.0		

- Perceived usefulness of the interface/assistant: the majority of responsens were ‘neutral’ (24%), ‘almost usefull’ (21%) and ‘very usefull’ (14%). Hence, it can be said that the majority of respondants perceived almost full usefullness during the purchasing experience provided by the simulations.

Quanto ti è sembrato utile l'assistente/interfaccia durante la simulazione?

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	1 = Per niente inutile	32	11.5	14.1	14.1
	2 = Poco utile	19	6.8	8.4	22.5
	3 = Moderatamente utile	21	7.5	9.3	31.7
	4 = Neutro	55	19.7	24.2	55.9
	5 = Abbastanza utile	48	17.2	21.1	77.1
	6 = Molto utile	32	11.5	14.1	91.2
	7 = Estremamente utile	20	7.2	8.8	100.0
	Totale	227	81.4	100.0	
Mancante	Sistema	52	18.6		
Totale		279	100.0		

- Decision speed (perceived decision time): the majority of responsens were ‘little’ (29%), ‘istant’ (28%) and ‘slightly’ (22%). Hence, it can be said that the majority of respondants perceived rapid decision time during the purchasing experience provided by the simulations.

Decision time

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	1 = Immediato	64	22.9	28.2	28.2
	2 = Poco	66	23.7	29.1	57.3
	3 = Leggermente	49	17.6	21.6	78.9
	4 = Neutro	30	10.8	13.2	92.1
	5 = Abbastanza	11	3.9	4.8	96.9
	6 = Molto	6	2.2	2.6	99.6
	7 = Troppo	1	.4	.4	100.0
	Totale	227	81.4	100.0	
Mancante	Sistema	52	18.6		
Totale		279	100.0		

- Confidence in the decision made: the majority of responsens were ‘enough secure’ (23%), ‘neutral’ (18%) and ‘much secure’ (16%). Hence, it can be said that the majority of respondants perceived high security in the decision made during the purchasing experience provided by the simulations.

Perceived security in choice

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	1 = Per niente sicuro/a	8	2.9	3.5	3.5
	2 = Poco sicuro	12	4.3	5.3	8.8
	3 = Leggermente sicuro	22	7.9	9.7	18.6
	4 = Neutro	40	14.3	17.7	36.3
	5 = Abbastanza sicuro	52	18.6	23.0	59.3
	6 = Molto sicuro	37	13.3	16.4	75.7
	7 = Completamente sicuro/a	55	19.7	24.3	100.0
	Totale	226	81.0	100.0	
Mancante	Sistema	53	19.0		
Totale		279	100.0		

- Engagement during the simulation: the majority of responsens were ‘neutral’ (24%), ‘slightly engage’ (20%) and ‘enough engage’ (19%). Hence, it can be said that the majority of respondants perceived slightly involvement during the purchasing experience provided by the simulations.

Perceived involvement in choice

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	1 = Per niente coinvolto	26	9.3	11.5	11.5
	2 = Poco coinvolto	27	9.7	11.9	23.5
	3 = Leggermente coinvolto	46	16.5	20.4	43.8
	4 = Neutro	55	19.7	24.3	68.1
	5 = Abbastanza coinvolto	42	15.1	18.6	86.7
	6 = Molto coinvolto	21	7.5	9.3	96.0
	7 = Completamente coinvolto	9	3.2	4.0	100.0
	Totale	226	81.0	100.0	
Mancante	Sistema	53	19.0		
Totale		279	100.0		

- Trust in the interface/assistant for future real-world use: the majority of responsens were ‘neutral’ (25%), ‘enough trust’ (20%) and ‘slightly trust’ (14%). Hence, it can be said that the majority of respondants were slightly confident during the purchasing experience provided by the simulations.

Trust in usage

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	1 = Per niente fiducioso	44	15.8	19.5	19.5
	2 = Poco fiducioso	22	7.9	9.7	29.2
	3 = Leggermente fiducioso	31	11.1	13.7	42.9
	4 = Neutro	57	20.4	25.2	68.1
	5 = Abbastanza fiducioso	46	16.5	20.4	88.5
	6 = Molto fiducioso	21	7.5	9.3	97.8
	7 = Completamente fiducioso	5	1.8	2.2	100.0
	Totale	226	81.0	100.0	
Mancante	Sistema	53	19.0		
Totale		279	100.0		

3.2.4 Demographic Section

Finally, the questionnaire included demographic questions and usage habits to profile respondents and provide potential covariates for the analysis. Information was collected on:

- Gender: male (44%), female (54%), non-binary (1%), prefer not to say (1%). Hence, the gender of the respondats were female.

Tabella delle frequenze

Gender

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Maschio	99	35.5	44.0	44.0
	Femmina	121	43.4	53.8	97.8
	Genere non-binario / Terzo genere	3	1.1	1.3	99.1
	Preferisco non dirlo	2	.7	.9	100.0
	Totale	225	80.6	100.0	
Mancante	Sistema	54	19.4		
Totale		279	100.0		

- Age: the majority of responses were 24 years old (10%), 23 years old (8%) and 25 years old (7%). Hence, the age of the respondants were mainly between 23 and 25 included.

Age

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	19	3	1.1	1.4	1.4
	20	7	2.5	3.2	4.6
	21	16	5.7	7.3	11.9
	22	8	2.9	3.7	15.5
	23	18	6.5	8.2	23.7
	24	21	7.5	9.6	33.3
	25	16	5.7	7.3	40.6
	26	4	1.4	1.8	42.5
	27	9	3.2	4.1	46.6
	28	3	1.1	1.4	47.9
	30	3	1.1	1.4	49.3
	33	2	.7	.9	50.2
	35	2	.7	.9	51.1
	37	1	.4	.5	51.6
	38	2	.7	.9	52.5
	39	1	.4	.5	53.0
	40	1	.4	.5	53.4
	41	1	.4	.5	53.9
	46	1	.4	.5	54.3

- Frequency of online purchases: the majority of responses were rarely (37%), several times in a month (26%) and 'once a month' (23%). Hence, the respondants were 'neutral' online shoppers.

Frequency of online purchases

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Mai	9	3.2	4.0	4.0
	Raramente (meno di una volta al mese)	82	29.4	36.6	40.6
	Circa una volta al mese	51	18.3	22.8	63.4
	Più volte al mese	58	20.8	25.9	89.3
	Circa una volta a settimana	11	3.9	4.9	94.2
	Più volte a settimana	10	3.6	4.5	98.7
	Quasi ogni giorno	3	1.1	1.3	100.0
	Totale	224	80.3	100.0	
Mancante	Sistema	55	19.7		
Totale		279	100.0		

These sections ensured that the sample characteristics could be described and that demographic variables could be included as covariates in additional analyses if necessary.

3.3 Data Preparation

3.3.1 Coding and Recoding of Variables

Each analysis variable was properly prepared to ensure consistency as well as clarity in the data set. The interaction mode variable (INTERACT_MODE) was created to indicate the experimental manipulation of the decision interface: it was coded as a binary factor, 0 = Screen condition, 1 = Voice condition. This enabled a direct test of the first hypothesis, in which differences in outcomes in decision-making in the two interaction modes were compared.

The product type variable (PRODUCT_TYPE) was also created as a categorical factor to distinguish between the two categories of products used in the scenarios: 0 = Hedonic product and 1 = Utilitarian product. The above coding enabled measurement of product type as a moderator, specifically, in testing the third hypothesis relating to congruence of product type and interaction mode.

The moderator variable perceived control was assessed along a 7-point Likert scale and retained in its continuous form for subsequent analysis. For convenience in testing moderation, a centered version of the variable was computed. Additionally, an interaction variable between perceived control and interaction mode was computed (MODE_X_CONTROL), and also an interaction variable between product type and interaction mode was computed (MODE_X_PRODUCT). Accordingly, a standardized scale of perceived control (ZCONTR) was computed, enabling comparisons irrespective of scale measures as well as robustness when entering the variable in regression models. Dependent variables—satisfaction, perceived speed of decision, and choice confidence—were likewise measured through 7-point Likert scales.

The values and labels were systematically verified and corrected over the entire dataset such that categorical variables like gender, interaction mode, and type of product were clearly defined and coded in a similar manner. This facilitated the interpretation of results and clarified output readability. Lastly, demographic variables like gender, age, and shopping frequency online were recoded when necessary in an effort to create uniform categories so that descriptive statistics could be included and covariates in subsequent analyses.

3.3.2 Creation of Composite Indices

I built a composite index of perceived quality of choice (QUALITY_INDEX) by averaging responses to a series of survey items, each of which dealt different aspects of this construct: choice utility, choice security, trust, and perceived involvement. The index attempts to give a richer, subtler measure of perceived quality of decision than would a solitary item by aggregating these several aspects in a single metric.

This method aligns with rigorously tested research, which highlights that the quality of a decision is multidimensional. Notably, constituent items for this index at the individual level were retained in the dataset as well, such that in supplementary analyses, these could serve to give a richer description of subtler effects.

3.3.3 Reliability of Scales (Cronbach's Alpha)

Reliability analysis was employed to investigate internal consistency in the quality index, which in turn was formed by averaging four items: perceived involvement in choice, trust in usage, satisfaction from a customer's perspective, and perceived usefulness of the interface during simulation. The reliability analysis yielded a Cronbach's Alpha of 0.749 derived from 226 usable cases, well above the 0.70 minimum and so establishes appropriate internal reliability. The item statistics reported that each of the four components substantially contributed to the scale, with corrected item–total correlations of 0.30 or higher, the minimum recommended. Inter-item correlations were moderate to strong (.208 to .574), where items were related but not redundant.

Statistiche di affidabilità

Alpha di Cronbach	Alpha di Cronbach basata su elementi standardizzati	N. di elementi
.749	.744	4

Statistiche degli elementi

	Media	Deviazione std.	N
Perceived involvement in choice	3.70	1.599	226
Trust in usage	3.54	1.689	226
Customer satisfaction	5.02	1.284	226
Quanto ti è sembrato utile l'assistente/interfaccia durante la simulazione?	4.08	1.808	226

Matrice di correlazione tra gli elementi

	Perceived involvement in choice	Trust in usage	Customer satisfaction	Quanto ti è sembrato utile l'assistente/interfaccia durante la simulazione?
Perceived involvement in choice	1.000	.547	.360	.504
Trust in usage	.547	1.000	.345	.564
Customer satisfaction	.360	.345	1.000	.208
Quanto ti è sembrato utile l'assistente/interfaccia durante la simulazione?	.504	.564	.208	1.000

3.3.4 Data Cleaning and Management

Data preparation took a systematic approach so that the dataset was ready for hypothesis testing. First, frequency tables along with descriptive statistics in SPSS were generated for categorical and continuous variables, respectively. This enabled identification of missing responses as well as values outside the desired Likert scale range of 1 to 7. Of the initial 279 questionnaire forms received, a few included partially filled response forms or missing values in essential dependent variables such as satisfaction, decision time, and perceived control. These cases were removed through listwise exclusion, and a final respondent sample of 226 was used in subsequent analyses.

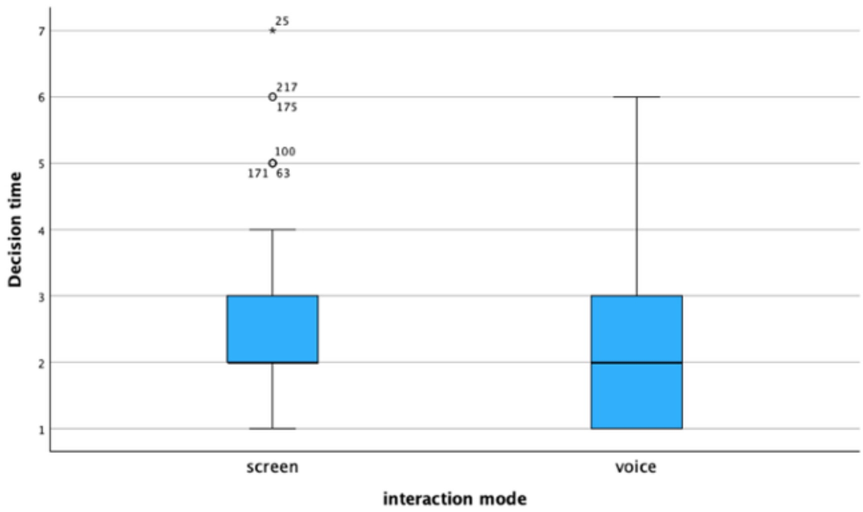
Frequenze

Statistiche

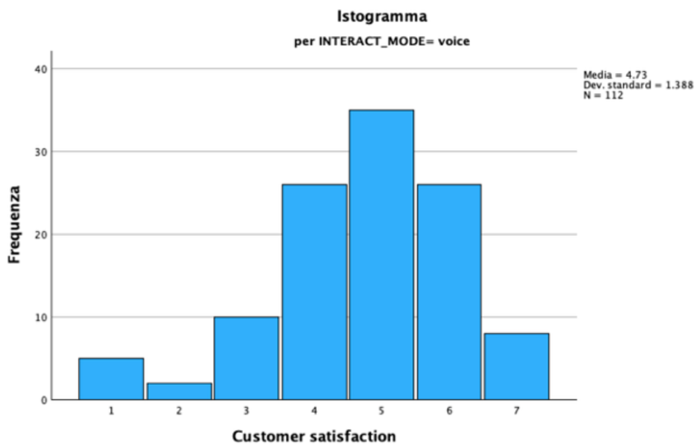
		Gender	Age	Frequency of online purchases	Customer satisfaction	Perceived control	Quanto ti è sembrato utile l'assistente/interfaccia durante la simulazione?	Decision time	Perceived security in choice	Perceived involvement in choice	Trust in usage
N	Valido	225	219	224	229	228	227	227	226	226	226
	Mancante	54	60	55	50	51	52	52	53	53	53
Media		1.59	41.03	3.10	5.02	4.31	4.07	2.47	4.98	3.70	3.54
Mediana		2.00	33.00	3.00	5.00	4.00	4.00	2.00	5.00	4.00	4.00
Modalità		2	24	2	5	4	4	2	7	4	4
Deviazione std.		.568	18.523	1.263	1.299	1.915	1.809	1.335	1.661	1.599	1.689
Varianza		.323	343.119	1.596	1.688	3.667	3.273	1.781	2.760	2.556	2.854
Intervallo		3	69	6	6	6	6	6	6	6	6
Minimo		1	19	1	1	1	1	1	1	1	1
Massimo		4	88	7	7	7	7	7	7	7	7

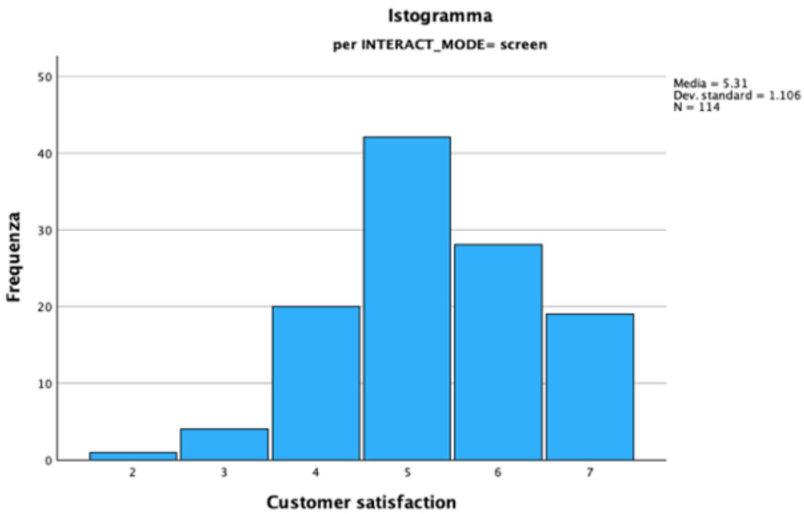
Second, I used SPSS *Explore* to generate boxplots and histograms for the key continuous variables—satisfaction, trust, perceived control, decision time, and the quality index—split by interaction mode and product type. The procedure flagged a few statistical outliers, mostly scale endpoints on the 7-point items (1 or 7). Because these values fell within the legitimate response range, they were considered plausible and retained to preserve natural variability; no miscoded values outside 1–7 were detected. The observed distributions were broadly suitable for parametric analyses, and the visual patterns anticipated the inferential results, such as:

- Distribution of decision time for both interaction mode are quite similar, resulting in a not significant difference of decision speed based on interaction mode.

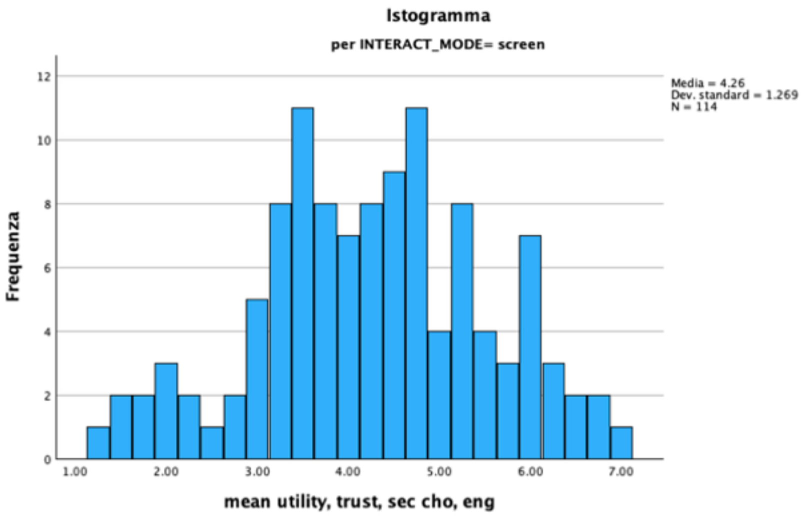


- It is notably that the mean of customer satisfaction for interaction mode is higher in screen interaction mode (5.31), rather than in voice interaction mode (4.73), not supporting the hypothesis three.

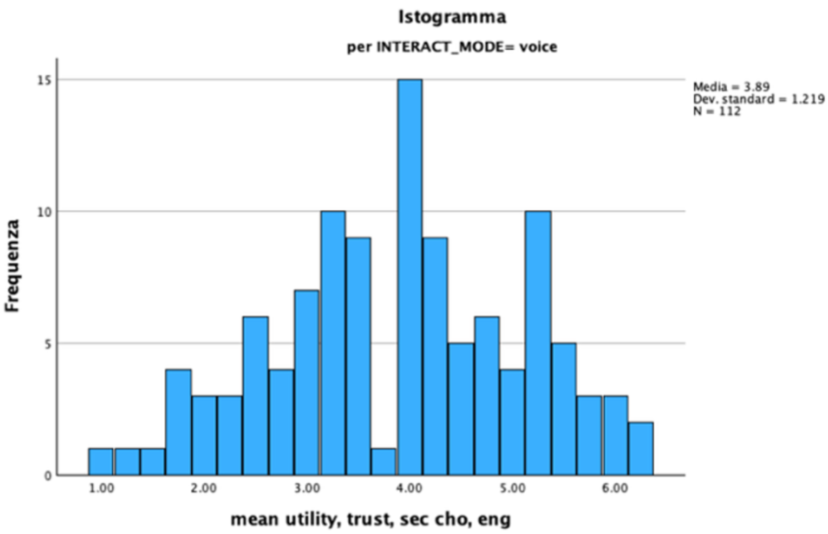




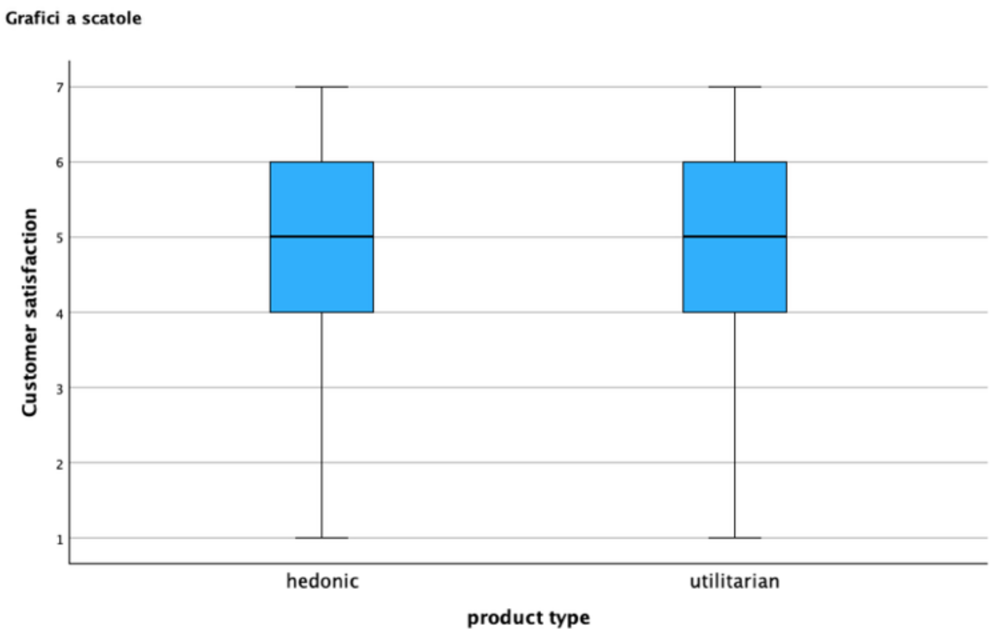
- It is notably that the mean of quality index for interaction mode is higher in screen interaction mode (4.26), rather than in voice interaction mode (3.89), not supporting the hypothesis two.



mean utility, trust, sec cho, eng



- It is notably that there’s no difference in terms of customer satisfaction in considering one type of product rather than the other one, suggesting the absence of a main effect of product type.



Finally, a filter was applied so that only participants who completed all required experimental sections were retained. This guaranteed analysis based upon full and consistent cases, yielding a final dataset of 226 participants used to test hypotheses.

3.4 Data Analysis and Hypotheses Testing

3.4.1 Descriptive Analysis

After the cleaning, 226 answers are used to represent the final sample. There are four experimental conditions: Voice/Utilitarian, Voice/Hedonic, Screen/Utilitarian, and Screen/Hedonic. The number of respondents of each group was roughly the same for these four conditions. More changes are in the direction of balanced experimental factors and the easier comparability between groups. At the overall level, 58.8% of subjects belonged to the Screen condition and 41.2% belonged to the Voice condition, 59.5% belonged to the Hedonic product condition, and 40.5% belonged to the Utilitarian product condition.

The four block were randomized such that 19.7% of valid respondents were assigned to voice-utilitarian simulation, 22.6% were assigned to voice-hedonic simulation, 21.2% were assigned to screen-utilitarian simulation, 21.9% were assigned to screen-hedonic simulation.

FL_10 - Block Randomizer - Display Order Simulazioneeprodotto1

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	selected	55	19.7	100.0	100.0
Mancante	Sistema	224	80.3		
Totale		279	100.0		

FL_10 - Block Randomizer - Display Order Simulazioneeprodotto2

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	selected	63	22.6	100.0	100.0
Mancante	Sistema	216	77.4		
Totale		279	100.0		

FL_10 - Block Randomizer - Display Order Simulazioneeprodotto3

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	selected	59	21.1	100.0	100.0
Mancante	Sistema	220	78.9		
Totale		279	100.0		

FL_10 - Block Randomizer - Display Order Simulazioneeprodotto4

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	selected	61	21.9	100.0	100.0
Mancante	Sistema	218	78.1		
Totale		279	100.0		

Tabella delle frequenze

interaction mode

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	screen	164	58.8	58.8	58.8
	voice	115	41.2	41.2	100.0
	Totale	279	100.0	100.0	

product type

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	hedonic	166	59.5	59.5	59.5
	utilitarian	113	40.5	40.5	100.0
	Totale	279	100.0	100.0	

In the next step, descriptive statistics were calculated for all dependent and moderating variables, that is satisfaction, perceived control, perceived usefulness, perceived speed of decision, trust, and engagement. On a 7-point scale, the respective means ranged from 3.9 to 5.4, and the standard deviations ranged from 1.26 to 1.91. The minimum and maximum values corresponded with predictions (1–7), so no errors were detected in coding, and there were no illegal values.

Statistiche

		Customer satisfaction	Perceived control	Quanto ti è sembrato utile l'assistente/interfaccia durante la simulazione?	Decision time	Perceived security in choice	Perceived involvement in choice	Trust in usage	mean utility, trust, sec cho, eng
N	Valido	229	228	227	227	226	226	226	227
	Mancante	50	51	52	52	53	53	53	52
Media		5.02	4.31	4.07	2.47	4.98	3.70	3.54	4.0672
Mediana		5.00	4.00	4.00	2.00	5.00	4.00	4.00	4.0000
Modalità		5	4	4	2	7	4	4	4.00
Deviazione std.		1.299	1.915	1.809	1.335	1.661	1.599	1.689	1.26064
Varianza		1.688	3.667	3.273	1.781	2.760	2.556	2.854	1.589
Intervallo		6	6	6	6	6	6	6	6.00
Minimo		1	1	1	1	1	1	1	1.00
Massimo		7	7	7	7	7	7	7	7.00

Descriptive analysis, in general, endorsed that the dataset was suitably cleaned, balanced in terms of conditions, had enough variability for all constructs, and that there were no ceiling and floor effects. Such outcomes provided the basis for a dependable and generalizable sample in the subsequent hypothesis testing.

3.4.2 Hypothesis Testing

- *H1. Those who communicate in voice channels will have higher speed in decision-making compared to those who communicate in traditional digital channels.*

  An independent-samples t-test compared decision time between the Screen (n = 114, M = 2.54, SD = 1.28) and Voice (n = 113, M = 2.41, SD = 1.39) conditions. Because Levene's test was non-significant (F = 1.124, p = .290), equal variances were assumed. The mean difference was not significant, t(225) = 0.72, p = .471; the 95% confidence interval for Screen minus Voice ranged from −0.221 to 0.477.
  Effect size was negligible (Cohen's d = 0.096; Hedges' g = 0.095; Glass's Δ = 0.092). Although lower scores indicate faster decisions and the Voice mean was slightly lower, there is no evidence of a reliable difference in decision time between the two modes.

  Thus, H1 was not supported.

Test t

Statistiche gruppo

	interaction mode	N	Media	Deviazione std.	Errore standard della media
Decision time	screen	114	2.54	1.284	.120
	voice	113	2.41	1.386	.130

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie							
						Significatività				Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	P unilaterale	P bilaterale	Differenza della media	Differenza errore std.	Inferiore	Superiore
Decision time	Varianze uguali presunte	1.124	.290	.722	225	.236	.471	.128	.177	-.221	.477
	Varianze uguali non presunte			.722	223.377	.236	.471	.128	.177	-.222	.478

Dimensioni effetto campioni indipendenti

		Standardizzatore[a]	Stima del punto	Intervallo di confidenza 95%	
				Inferiore	Superiore
Decision time	D di Cohen	1.336	.096	-.165	.356
	Correzione di Hedges	1.341	.095	-.164	.355
	Delta di Glass	1.386	.092	-.168	.353

a. Il denominatore utilizzato per stimare le dimensioni dell'effetto.
La d di Cohen utilizza la deviazione standard raggruppata.
La correzione di Hedge utilizza la deviazione standard raggruppata, piu un fattore di correzione.
Il delta di Glass utilizza la deviazione standard del campione del gruppo di controllo (ovvero il secondo).

- *H2. Perceived control will mediate the interaction mode-effect of voice versus conventional interaction mode, such that higher perceived control will lead to higher perceived quality of choice in both interaction modes.*

The dependency was regressed in a moderation analysis where quality index was a dependent variable, interaction mode was an independent variable, perceived control was a moderator variable, in its standardized form, and its interaction was entered in a model. The regression equation was significant, $F(3, 222) = 11.21$, $p < .001$, explaining 13.2% of perceived quality's variability ($R^2 = .132$, adjusted $R^2 = .120$).

Results showed that perceived control was a strong positive predictor of perceived quality during interactions ($B = 0.575$, $\beta = .458$, $t = 4.94$, $p < .001$), such that a higher level of control was associated with higher perceived quality. Interaction mode was found to have a significant impact as well ($B = –0.353$, $\beta = –.141$, $t = –2.25$, $p = .025$), such that, compared to the Screen condition, there was a slightly lower level of perceived quality in the Voice condition. Importantly, the interaction term between interaction mode and perceived control was significant as well ($B = –0.377$, $\beta = –.222$, $t = –2.39$, $p = .018$), which means that there was a different association between perceived control and quality depending upon interaction mode.

Riepilogo del modello

Modello	R	R-quadrato	R-quadrato adattato	Errore std. della stima	Statistiche delle modifiche: Modifica R-quadrato	Modifica F	gl1	gl2	Sign. Modifica F
1	.363[a]	.132	.120	1.17824	.132	11.208	3	222	<.001

a. Predittori: (costante), MODE_X_CONTROL, interaction mode, Punteggio Z: Perceived control

ANOVA[a]

Modello		Somma dei quadrati	gl	Media quadratica	F	Sign.
1	Regressione	46.680	3	15.560	11.208	<.001[b]
	Residuo	308.191	222	1.388		
	Totale	354.871	225			

a. Variabile dipendente: mean utility, trust, sec cho, eng

b. Predittori: (costante), MODE_X_CONTROL, interaction mode, Punteggio Z: Perceived control

Coefficienti[a]

Modello		Coefficienti non standardizzati: B	Errore standard	Coefficienti standardizzati: Beta	t	Sign.	95,0% Intervallo di confidenza per B: Limite inferiore	Limite superiore	Statistiche di collinearità: Tolleranza	VIF
1	(Costante)	4.246	.110		38.453	<.001	4.028	4.463		
	interaction mode	-.353	.157	-.141	-2.253	.025	-.662	-.044	.999	1.001
	Punteggio Z: Perceived control	.575	.117	.458	4.938	<.001	.346	.805	.454	2.200
	MODE_X_CONTROL	-.377	.158	-.222	-2.392	.018	-.688	-.066	.455	2.199

a. Variabile dipendente: mean utility, trust, sec cho, eng

Specifically, the positive effect of perceived control over perceived quality was larger in the Screen condition relative to the Voice condition. For high perceived control, both modes afforded higher perceived quality, whereas the advantage of Screen over Voice was larger.

Thus, contrary to initial expectation, only in part was there support for H2: perceived control significantly influenced perceived quality, although there was an interaction in moderation different from that hypothesized.

- *H3. Customer satisfaction is higher when interaction mode is consistent in terms of product type, such that voice produces higher satisfaction for utilitarian products and traditional interaction produces higher satisfaction for hedonic products.*

Two-way ANOVA was employed where interaction mode and product type were fixed factors and satisfaction from customers was a dependent variable. Levene's test confirmed assumptions of homogeneity of variances, $F(3,225) = 1.24$, $p = .296$.

The results revealed a significant main effect for interaction mode, $F(1,225) = 10.41$, $p = .001$, $\eta^2 = .044$, where participants in the Screen condition reported higher satisfaction ($M = 5.29$, $SD = 1.15$) than participants in the Voice condition did ($M = 4.74$, $SD = 1.39$). The main effect of product type was not significant, $F(1,225) = 2.28$, $p = .132$, $\eta^2 = .010$.

Importantly, the interaction effect between interaction mode and product type was not significant, $F(1,225) = 0.08$, $p = .783$, $\eta^2 < .001$, providing no support for the hypothesized alignment effect. However, pairwise tests revealed significant simple effects in each product category: For hedonic products, satisfaction was higher in the Screen condition ($M = 5.14$) than in the Voice condition ($M = 4.64$), $\Delta M = 0.50$, $p = .037$. For utilitarian products, satisfaction was greater in the Screen condition ($M = 5.44$) than in the Voice condition ($M = 4.85$), $\Delta M = 0.59$, $p = .015$.
Thus, surprisingly, consumers consistently showed greater satisfaction in the Screen condition in both product categories, and hypothesized alignment effect (Voice–Utilitarian, Screen–Hedonic) did not come to pass.

Thus, H3 was not supported

Fattori tra soggetti

		Etichetta valore	N
interaction mode	.00	screen	116
	1.00	voice	113
product type	.00	hedonic	116
	1.00	utilitarian	113

Statistiche descrittive

Variabile dipendente: Customer satisfaction

interaction mode	product type	Medio	Deviazione std.	N
screen	hedonic	5.14	1.187	57
	utilitarian	5.44	1.103	59
	Totale	5.29	1.150	116
voice	hedonic	4.64	1.436	59
	utilitarian	4.85	1.338	54
	Totale	4.74	1.387	113
Totale	hedonic	4.89	1.337	116
	utilitarian	5.16	1.250	113
	Totale	5.02	1.299	229

Test di Levene di eguaglianza delle varianze dell'errore[a,b]

		Statistica di Levene	gl1	gl2	Sig.
Customer satisfaction	Basato sulla media	1.240	3	225	.296
	Basato sulla mediana	.659	3	225	.578
	Basato sulla mediana e con il grado di libertà adattato	.659	3	214.239	.578
	Basato sulla media ritagliata	.966	3	225	.410

Verifica l'ipotesi nulla che la varianza dell'errore della variabile dipendente sia uguale tra i gruppi.

a. Variabile dipendente: Customer satisfaction

b. Disegno: Intercetta + INTERACT_MODE + PRODUCT_TYPE + INTERACT_MODE * PRODUCT_TYPE

Test di effetti tra soggetti

Variabile dipendente: Customer satisfaction

Origine	Somma dei quadrati di tipo III	df	Media quadratica	F	Sig.	Eta quadrato parziale	Parametro noncent.	Potenza osservata[b]
Modello corretto	21.131[a]	3	7.044	4.357	.005	.055	13.070	.866
Intercetta	5761.577	1	5761.577	3563.766	<.001	.941	3563.766	1.000
INTERACT_MODE	16.830	1	16.830	10.410	.001	.044	10.410	.895
PRODUCT_TYPE	3.690	1	3.690	2.283	.132	.010	2.283	.325
INTERACT_MODE * PRODUCT_TYPE	.122	1	.122	.076	.783	.000	.076	.059
Errore	363.760	225	1.617					
Totale	6160.000	229						
Totale corretto	384.891	228						

a. R-quadrato = ,055 (R-quadrato adattato = ,042)

b. Calcolato utilizzando alfa = .05

Confronti pairwise

Variabile dipendente: Customer satisfaction

product type	(I) interaction mode	(J) interaction mode	Differenza della media (I-J)	Errore std.	Sig.[b]	95% intervallo di confidenza per differenza[b] Limite inferiore	Limite superiore
hedonic	screen	voice	.496*	.236	.037	.031	.962
	voice	screen	-.496*	.236	.037	-.962	-.031
utilitarian	screen	voice	.589*	.239	.015	.117	1.061
	voice	screen	-.589*	.239	.015	-1.061	-.117

Basato sulle medie marginali stimate

*. La differenza della media è significativa al livello .05.

b. Adattamento per confronti multipli: differenza meno significativa (equivalente a nessun adattamento).

Test univariati

Variabile dipendente: Customer satisfaction

product type		Somma dei quadrati	df	Media quadratica	F	Sig.	Eta quadrato parziale	Parametro noncent.	Potenza osservata[a]
hedonic	Contrasto	7.140	1	7.140	4.417	.037	.019	4.417	.553
	Errore	363.760	225	1.617					
utilitarian	Contrasto	9.776	1	9.776	6.047	.015	.026	6.047	.687
	Errore	363.760	225	1.617					

Ogni F verifica gli effetti semplici di interaction mode all'interno di ciascuna combinazione di livello degli altri effetti visualizzati. Questi test si basano sui confronti pairwise linearmente indipendenti tra le medie marginali stimate.

a. Calcolato utilizzando alfa = .05

3.4.3 Additional Analyses

To complement the hypothesis tests (H1–H3) and provide a fuller picture of the data, three additional, theory-consistent analyses were conducted: (a) a correlation matrix among the study variables, (b) a multiple regression identifying the best predictors of satisfaction, and (c) a correlation between objective decision time and perceived confidence in one's choice. These analyses were selected to (i) verify convergent patterns among constructs used in the main models, (ii) explain variance in satisfaction beyond experimental factors (useful given H3's lack of interaction), and (iii) relate behavior (time) to attitudinal outcomes (confidence), adding interpretation to H1.

a) *Pearson correlations among key constructs*

Pearson correlations were computed among satisfaction, trust, engagement, the quality index, perceived control, decision time, and perceived security in choice. This provides convergent/ discriminant evidence across constructs and helps interpret the direction and strength of associations relevant to H2–H3.

Satisfaction correlated positively with trust ($r = .345$, $p < .001$), engagement ($r = .360$, $p < .001$), and the quality index ($r = .419$, $p < .001$). A weaker—yet significant—positive correlation emerged with perceived control ($r = .192$, $p = .004$). Satisfaction was not related to objective decision time (r = .029, p = .665). Trust and engagement were strongly associated with the quality index ($|r| \geq .795$, $p < .001$).

These patterns are consistent with the main results: satisfaction aligns more with relational/experiential dimensions (trust, engagement, perceived quality) than with speed (helping explain why H1 was not supported). The modest link with perceived control also fits the H2 finding that control matters but does not straightforwardly strengthen the voice vs. screen effect.

Correlazioni

		Customer satisfaction	Trust in usage	Perceived involvement in choice	mean utility, trust, sec cho, eng	Perceived control	Decision time	Perceived security in choice
Customer satisfaction	Correlazione di Pearson	1	.345**	.360**	.419**	.192**	.029	.369**
	Sign. (a due code)		<.001	<.001	<.001	.004	.665	<.001
	N	229	226	226	227	228	227	226
Trust in usage	Correlazione di Pearson	.345**	1	.547**	.804**	.167*	.121	.274**
	Sign. (a due code)	<.001		<.001	<.001	.012	.070	<.001
	N	226	226	226	226	226	226	226
Perceived involvement in choice	Correlazione di Pearson	.360**	.547**	1	.795**	.239**	.113	.337**
	Sign. (a due code)	<.001	<.001		<.001	<.001	.091	<.001
	N	226	226	226	226	226	226	226
mean utility, trust, sec cho, eng	Correlazione di Pearson	.419**	.804**	.795**	1	.302**	.059	.599**
	Sign. (a due code)	<.001	<.001	<.001		<.001	.374	<.001
	N	227	226	226	227	227	227	226
Perceived control	Correlazione di Pearson	.192**	.167*	.239**	.302**	1	.077	.300**
	Sign. (a due code)	.004	.012	<.001	<.001		.248	<.001
	N	228	226	226	227	228	227	226
Decision time	Correlazione di Pearson	.029	.121	.113	.059	.077	1	-.196**
	Sign. (a due code)	.665	.070	.091	.374	.248		.003
	N	227	226	226	227	227	227	226
Perceived security in choice	Correlazione di Pearson	.369**	.274**	.337**	.599**	.300**	-.196**	1
	Sign. (a due code)	<.001	<.001	<.001	<.001	<.001	.003	
	N	226	226	226	226	226	226	226

**. La correlazione è significativa a livello 0,01 (a due code).

*. La correlazione è significativa a livello 0,05 (a due code).

b) *Multiple regression: predictors of satisfaction*

A linear regression was estimated with satisfaction as the dependent variable and trust, engagement, and perceived control as predictors. This model tests which attitudinal factors best account for satisfaction, beyond experimental manipulations (useful given H3's absent interaction).

The model was significant, $F(3, 222) = 15.255$, $p < .001$, explaining 17.1% of variance ($R^2 = .171$; adj. $R^2 = .160$). Trust ($B = 0.156$, $\beta = .206$, $p = .005$) and engagement ($B = 0.180$, $\beta = .224$, $p = .003$) were positive, significant predictors; perceived control was not significant when entered with the other predictors ($B = 0.068$, $\beta = .102$, $p = .106$). Multicollinearity was not a concern (VIFs ≤ 1.47).

Satisfaction is chiefly explained by trust and engagement, not by perceived control once the former are accounted for. This aligns with H2's outcome (moderation effects not in the hypothesized direction) and with H3's pattern (screen > voice) by suggesting that the screen interface likely fosters higher trust/engagement, which in turn elevates satisfaction.

Riepilogo del modello

					Statistiche delle modifiche				
Modello	R	R-quadrato	R-quadrato adattato	Errore std. della stima	Modifica R-quadrato	Modifica F	gl1	gl2	Sign. Modifica F
1	.413[a]	.171	.160	1.177	.171	15.255	3	222	<.001

a. Predittori: (costante), Perceived control, Trust in usage, Perceived involvement in choice

ANOVA[a]

Modello		Somma dei quadrati	gl	Media quadratica	F	Sign.
1	Regressione	63.390	3	21.130	15.255	<.001[b]
	Residuo	307.499	222	1.385		
	Totale	370.889	225			

a. Variabile dipendente: Customer satisfaction

b. Predittori: (costante), Perceived control, Trust in usage, Perceived involvement in choice

Coefficienti[a]

		Coefficienti non standardizzati		Coefficienti standardizzati			Statistiche di collinearità	
Modello		B	Errore standard	Beta	t	Sign.	Tolleranza	VIF
1	(Costante)	3.509	.251		13.974	<.001		
	Trust in usage	.156	.056	.206	2.813	.005	.700	1.429
	Perceived involvement in choice	.180	.060	.224	3.015	.003	.679	1.473
	Perceived control	.068	.042	.102	1.621	.106	.941	1.063

a. Variabile dipendente: Customer satisfaction

c) *Objective decision time vs. perceived security in choice*

To connect behavior to felt certainty, I correlated decision time with perceived security/confidence in choice. If quicker decisions associate with greater confidence, this offers a behavioral lens on the null result of H1. Decision time correlated negatively with perceived security, $r = -.196$, $p = .003$ ($N = 226$): faster decisions were associated with higher reported confidence. Although voice did not make decisions faster on average (H1 not supported), individuals who decided more quickly—regardless of condition—tended to feel more confident about their choice. This helps dissociate "speed" as a personal/within-group tendency from the between-group (mode) comparison.

Correlazioni

Correlazioni		Decision time	Perceived security in choice
Decision time	Correlazione di Pearson	1	-.196**
	Sign. (a due code)		.003
	N	227	226
Perceived security in choice	Correlazione di Pearson	-.196**	1
	Sign. (a due code)	.003	
	N	226	226

**. La correlazione è significativa a livello 0,01 (a due code).

Taken together, these additional analyses show that (i) satisfaction clusters with trust, engagement, and perceived quality rather than with speed; (ii) trust and engagement are the most informative predictors of satisfaction; and (iii) quicker decisions relate to higher confidence, but this relationship does not depend on voice vs. screen. These insights are coherent with the hypothesis tests: H1 (no speed difference) and H3 (no alignment effect) were not supported, while H2 highlighted that perceived control does not straightforwardly amplify the benefit of voice relative to screen.

Conclusions

The study's objectives were, in large part, achieved. The experiment ran as planned: measures were collected and validated (including a reliable quality index, $\alpha = .749$), the data were cleaned down to a robust analytical sample ($N = 226$), and hypotheses were tested with the appropriate methods—an independent-samples t-test for decision speed, a moderated regression for perceived quality, and a two-way ANOVA for satisfaction. The results gave clear answers: voice did not speed up decisions compared to screen; perceived control was a strong predictor of perceived decision quality but interacted with mode in the opposite direction of what had been anticipated, with control translating more effectively into quality under screen conditions; and satisfaction was consistently higher under screen across both product types, with no evidence for the predicted voice–utilitarian/screen–hedonic alignment. Complementary analyses added depth, showing that satisfaction clustered with trust, engagement, and perceived quality rather than with decision time, and that faster individual decisions correlated with greater confidence regardless of modality. In doing so, the study not only met its goals but also refined them, shifting the role of "speed" to a secondary factor compared to the experiential levers of satisfaction.

Viewed through the lens of the central research question—when and how interaction mode shapes consumer decision processes and outcomes—the thesis offers a coherent response. In the tested purchase scenarios, screens systematically outperformed voice on satisfaction and on the conversion of perceived control into perceived decision quality. Voice did not deliver the expected speed advantage. The evidence suggests that modality effects are contingent on the affordances demanded by the task: when users need to compare options in parallel, scan information quickly, and backtrack with ease, visual interfaces dominate. The strongest contributors to satisfaction were trust and engagement rather than efficiency. Thus, the answer is both empirical and conceptual: voice is most useful as a friction-reducer early in a journey or for simple tasks, while screens remain superior for evaluation-heavy decisions; designing for perceived control is universally beneficial, but its impact is magnified when options and progress are made visible.

Arriving at these conclusions involved both methodological and conceptual challenges. Operationalizing "perceived quality" required the careful construction and validation of a composite index, balancing theoretical fidelity with parsimony. Data cleaning demanded nuanced handling of missing values and outliers without stripping natural variation, while assumption checks had to be reconciled with the robustness of parametric tests in larger samples. Interpreting non-significant results—such as the null for H1 and the lack of alignment in H3—required triangulation through complementary analyses to avoid simplistic readings of the null. Constraints typical of simulation-based research also applied: voice and screen interactions were standardized rather than tailored, perceived control was measured rather than manipulated, the product set was deliberately narrow, and outcomes were captured as self-reports tied to single episodes rather than longitudinal behavior. These design choices safeguarded internal validity but limited generalizability, warranting cautious interpretation beyond the tested contexts.

Future research can push these findings forward on several fronts. On the design side, experiments that manipulate control cues directly—adjusting menu depth in voice, adding state summaries, or varying reversibility options—would help establish the precise conditions under which control enhances decision quality. At the journey level, testing multimodal flows (e.g., voice-to-screen handoffs) could show how orchestration across modes, rather than reliance on one channel, shapes outcomes for simple versus complex tasks. At the model level, mediated and sequential frameworks (for example, Mode → Trust/Engagement → Satisfaction, with Control as a moderator or mediator) could be tested using structural models and richer behavioral logs, such as latency, repair tu

Bibliography

- Flavián, Akdim, Casalò - Effects of voice assistant recommendations on consumer behavior, Psychology & Marketing, 2022
- Dellaert, Arora, Bode - Consumer decisions with artificially intelligent voice assistants, Marketing Letters, 2020
- Epley, N., & Schroeder, J. - Mistakenly Seeking Solitude, Journal of Experimental Psychology*:* General, 2014
- Guha, Grewal, Schweiger - How artificiality and intelligence affect voice assistant evaluations, Journal of the Accademy of Marketing Science, 2023
- Koschate-Fischer, Hoyer, Schindler - How speaking versus writing to conversational agents shapes consumers' choice and choice satisfaction**,** Journal of the Accademy of Marketing Science, 2024
- Moriuchi - Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty, Psychology & Marketing, 2019
- Acikgoz, Perez-Vega, Stylos - Consumer engagement with AI-powered voice assistants: A behavioral reasoning perspective, Psychology & Marketing, 2023
- Luchtenberg, Blackader, Buesing - The evolution of customer care: AI and the Gen Z effect, McKinsey & Company, 2025
- Simss - How Voice Assistants Could Change the Way We Shop, Harward Business Review, 2019
- Nass & Steuer – Computers are Social Actors, Human Factors in Computing Systems, 1994
- Wohr, Voice assistants: How they are evolving and what they offer marketing and commerce, Emarketer.com, 2025