A Mixed-Methods Approach to Analysing Brand Positioning Through Electronic Word-of-Mouth: Combining Thematic and Correspondence Analysis for Marketing Insights

Table of Contents

Abstract

This research explores the relationship between brand positioning and electronic word-of-mouth through a mixed-methods approach. With online reviews becoming increasingly influential in shaping consumer perceptions, analyzing the vast amount of unstructured textual data presents a challenge for marketers. To address this, the study integrates qualitative thematic analysis with quantitative correspondence analysis, aiming to develop a methodology that overcomes many of the limitations of machine-led text mining.

The research has three main objectives: (1) evaluating the effectiveness of thematic analysis in extracting key psychological, behavioural, and sentiment factors from electronic word-of-mouth, (2) determining the necessary volume of reviews to achieve data saturation and establish a stable understanding of a brand's position, and (3) assessing the utility of correspondence analysis in visually representing brand relationships. Analysing Amazon product reviews, the study found that thematic saturation can be reached with approximately 50 reviews, while stable brand-theme relationships become evident with around 300 reviews.

The findings demonstrate the importance of combining qualitative insights with quantitative tools to better understand consumer behaviour. This mixed-methods approach offers a nuanced understanding of brand positioning, providing actionable insights that can be easily applied by marketing professionals. Ultimately, the study highlights the need for accessible, effective tools to analyse electronic word-of-mouth, bridging the gap between theoretical research and its practical application in brand management.

1. Positionality Statement

1.1 Educational and Professional Background

I am currently pursuing an MSc in Management with Marketing, focusing on a dissertation exploring brand positioning and eWOM. My academic background, including an MSc in Applied Quantitative Methods and a BSc in Psychology, provides a foundation in psychological insight, statistical rigor, and marketing theory. With over seven years of marketing experience, such as roles as Market Intelligence Analyst and Senior CX and Marketing Analyst, my research is influenced by the practical application of theories and research methodologies.

1.2 Researcher's Role and Potential Bias

While striving for objectivity, my quantitative methods background may introduce biases, as I prioritize methods that are accessible and effective in professional settings. Though I prefer qualitative approaches to understanding text, I maintain objectivity, avoiding vested interest in proving their superiority or in specific themes.

1.3 Mixed-Methods Approach Perspective

My scepticism about purely machine-led text mining, which often misses context and emotion, drives my focus on a human-led approach, balancing qualitative insights with my quantitative strengths to capture the complexity of human behaviour in marketing research.

1.4 Impact of Positionality on Research

My professional background influences my methodological choices, as my preference for human-led analysis may limit alternative approaches. My focus on actionable insights could skew the study toward practical applications rather than theoretical contributions.

2. Research Context and Rationale

Increasing numbers of adverts have led to an overall lower recall and lower perception of editorial quality as product volumes continue to grow and compete for consumer attention (Webb and Ray, 1979; Pieters and Bijmolt, 1997; Riebe and Dawes, 2006). Brand positioning aims to counteract such issues by finding a unique place within the minds of consumers (Trout and Ries, 1986), highlighting the importance of understanding how brands are positioned in a crowded marketplace.

2.1 Challenges in Analysing Unstructured Textual Data

Monitoring electronic word-of-mouth (eWOM) offers a unique opportunity to monitor a brand's position, but it comes with two main challenges. Firstly, text data rarely fits into neat and testable models like statistical hypothesis tests do. Secondly, the volume of such unstructured data is increasing rapidly (Liu et al., 2024), thus leading to challenges in analyse such large quantities of such data, and even the question of how much data is enough. This may often lead marketers and brand managers with a dilemma: whether to use qualitative methods to handle the nuances of such textual data, but not the volume or text-mining approaches to handle the volumes of data, but perhaps lose the nuance.

2.2 A Mixed-Methods Approach to Brand Positioning

The present study aims to test a mixed-methods approach which seeks to harness the benefits both qualitative and quantitative approaches, while favouring accessible methods to improve its real-world accessibility. Research from Lilien (2011) shows that marketing managers did not understand complex machine learning methods due to their complexity, leading to lack of trust. There is also a considerable gap between the demand and supply of analytically trained professionals capable of handling complex data analysis (Leeflang et al., 2014). This shortfall creates obstacles for companies trying to utilise data for decision-making. The scarcity of skilled experts who can interpret and use data effectively impedes firms' ability to compete in the digital marketplace. While innovations in technology offer powerful tools for brand positioning, their complexity may hinder practical application in everyday business contexts. For these tools to be effectively adopted, they must be easy to use and demonstrate clear, perceived value (Davis, 1989). Many analytical methods remain underutilised because they are not accessible or easily understood by practitioners (King and He, 2006; Venkatesh et al., 2003). Bridging this gap requires developing methods that are both sophisticated in insight and simple in application. Only by making these tools more user-friendly and valuable can they truly impact marketing practice (Turner et al., 2010).

2.3 Research Objectives

This research addresses the challenges marketers face in analysing unstructured textual data, particularly online reviews, by using a mixed-methods approach that integrates qualitative thematic analysis with quantitative correspondence analysis (CA). The study aims to develop accessible methods that bridge the gap between academic research and practical application, balancing the nuanced understanding of text data with statistical rigor. It focuses on three key areas: developing themes through thematic analysis, determining the necessary volume of reviews for reliable brand positioning, and evaluating the effectiveness of CA in visualizing brand relationships. Ultimately, the research provides actionable tools for practitioners to enhance brand positioning strategies.

2.4 Overview of the Report Structure

The rest of this report is structured as follows. It begins with a literature review, exploring the current research in the field to establish a theoretical foundation and highlight gaps that this study aims to address. The Methods section outlines the mixed-method approach used to analyze textual data, detailing the qualitative, quantitative and resampling techniques. Following this, the Findings and Discussion sections will present and examine the findings from the analysis, comparing new findings to existing knowledge. Finnaly, conclusion and future recommendations are discussed.

3. Literature Review

Unwittingly, brands establish a position in consumer minds via eWOM when potential customers read the reviews and social media posts of existing customers. This has become a critical element in shaping brand reputation and influencing purchasing behaviour. On this basis, such data provides valuable insights for brand managers to understand and enhance their brand's position, find areas for product and service improvement, and benchmark the competition.

3.1 The Role of eWOM

eWOM refers to exchanging information about products or services via online reviews, social media, and forums (Hu and Ha, 2015; Dwivedi et al., 2017). Ratings and reviews have become ubiquitous in the digital environment. This digital feedback significantly influences consumer behaviour as potential buyers rely on these reviews to make informed decisions. As modern consumers face an overwhelming volume of eWOM (Liu et al., 2024), understanding how this effects consumer behaviour is essential.

Given that most consumers only glance at a handful of reviews, they are likely to rely on heuristics to form their impressions. Kahneman (2017)'s dual-system theory offers insight into this: when overwhelmed with information, people often default to System 1—fast, automatic, and emotional thinking—over the slower, more logical System 2. This is likely to influence how people handle the ever increasing volume of reviews. Another heuristic some people use is negativity bias which infuences decision-making, with customers often giving more weight to negative reviews than positive ones Kanouse (1984). This is supported by Lee et al. (2008) who found negative reviews to generally be more influential.

Jiang et al. (2021) demonstrated that consumers often seek additional online information, such as reviews, before making a purchase. Furthermore, a meta-analysis of 69 studies conducted by Ismagilova et al. (2020), identified which eWOM factors as influential in shaping purchase decisions. Key findings identify factors with the strongest predictive power, such as argument quality and valence, and less effective ones, including volume and source credibility, on purchase intentions. However, valance has been found to be less important with lower number of reviews Maslowska et al. (2017). Additionally, a systematic review by Kutabish et al. (2023) shows that social learning theory, which emphasizes learning behaviours and information through observation, plays a significant role in how consumers process and act upon eWOM.

Expectation confirmation theory explains eWOM's impact on consumer behaviour (Aurier and Guintcheva, 2014), showing how post-purchase behaviour is influenced by whether a product meets or exceeds expectations. Exceeding consumer expectations, as explained by Expectation Confirmation Theory, increases satisfaction and the likelihood of positive reviews. Reading reviews before purchase shapes these expectations, affecting post-purchase evaluations. ECT suggests satisfaction depends on how well a product meets expectations, influencing repurchase intention and loyalty. Along with affecting the likelihood of repatronage, it also effects the likelihood that consumers will write reviews. For example, @tata2020examining showed that those whose expectations were exceeded were more likely to write a review, while @vijay2018intention found that for online shoppers whose expectations were met and who were satisfied, the likelihood of writing reviews increased.

3.2 Brand Positioning

Positioning refers to the strategies brands use to establish a unique place in the minds of consumers Trout and Ries (1986) and is considered a fundamental component of marketing strategy (Kotler et al., 2019, p 320). By highlighting key differences from competitors and aligning with consumer needs, positioning helps brands create a clear value proposition, enabling consumers to easily recognize which products best meet their expectations.

Research in brand positioning dates back to the 1960s where the concept was popularised by pioneers such as Alpert and Gatty (1969). Using factor analysis, they advocated for the use of behavioural life-styles. Trout and Ries (1986), on the other hand, states that positioning involves shaping the perception of a product or service and is more what is done in the consumer's minds rather than the deliberate activities carried out by the company. These contrasting views or exemplified by Bhat and Reddy (1998), noting that its definition is not fully agreed upon. For example, some researchers stat that positioning is the process of creating and modifying how customers perceive a product (e.g., Crawford, 1985; Wright, 1997). While Palmer (2013) argue that it's organization's active effort to differentiate its offerings and gain a competitive advantage. In their review of 152 papers, Saqib (2020) found that positioning research can be categorized into five core perspectives. These were focusing on creating a product image distinct from competitors; finding and filling an empty slot in consumers' minds; establishing changes in consumer perception about the product; Differentiation, creating a unique position separate from competitors; finally, achieving a competitive advantage through unique strategies not employed by competitors. Other papers have focussed on a brand's efforts to reposition themselves, and develop ways to strike a chord with newer generations (Simms and Trott, 2007).

Brands must remain constantly aware by monitoring shifts in consumer behaviour. For example, research in behavioural economics and consumer culture theory demonstrates that emotions, experiences, and psychological factors play a greater role than previously recognized (Malter et al., 2020) when making purchase decisions. Furthermore, consumer attitutudes shift, like the recent shift towards green issues (White et al., 2019). Finally, huge global events can shift consumer behaviours, such as where and when consumers go shopping (Tao et al., 2022). This research suggests that brand positions can naturally shift in accordance with consumer behaviour shifts alone, even when the brand does not change.

3.3 Understanding Brand Positioning and Reputation Through eWOM

Both Prihananto et al. (2024) and Alzate et al. (2022) applied LIWC and PCA mixed methods to analyse brand positions ofmobile phone and make-up brands respectively. LIWC is a dictionary-based method which organizes words into specific themes. For instance, words like "sexual", "lover", and "sexy" are categorized under the "body" theme, while terms such as "acne" and "allergy" are grouped within the "health" theme. Dressler and Paunovic (2021) examined German wine producers by analysing the most common key attribute terms used on their "about us" pages (brand identity) and in customer reviews (brand image). Through a clustering approach, the researchers identified differences between the brand's intended image and customers' perceptions. Sun et al. (2023) used a keyword-based sentiment analysis to explore how different cruise attributes influenced tourist emotions, identifying patterns across various market segments. This method involved extracting and analysing high-frequency words related to cruise features, providing insight into the aspects that most strongly impact tourist sentiments. Karadağ et al. (2022) analyzed user-generated and brand-generated content from Turkish universities' social media platforms, coding the data according to pre-defined brand personality key words such as prestige, sincerity, and cosmopolitanism. Taecharungroj (2022) used LDA-topic modelling to discover topics on a review website for Thai beach resorts. They then aggregated the number of topics per beach and performed clustering to identify resorts that were most similar in terms of topic similarity. This approach allowed the researchers to explore thematic patterns and similarities across different resorts. critique

While machine-led text mining can be a powerful tool, one major issue is the accessiblily of large volume of data for marketing professionals (Ankers and Brennan, 2002). Another shortcoming is that "bag-of-words" methods experience a loss of context, word order, and nuance. For example, in the case of the makeup brand study, where "nude" could refer to a lipstick colour or product names, the method fails to distinguish this from the concept of nudity, which has a vastly different meaning. Similarly, words like "body" may refer to cosmetics for body care, rather than physical attributes, and this lack of distinction can lead to misleading categorization. LIWC also obscures sentiment, e.g., whether a product causes or conceals acne has vastly different implications. This can lead to misinterpretation, where words are grouped similarly despite conveying different sentiments in varying contexts, further obscuring the true meaning of the text (Le and Mikolov, 2014). In the wine study, Dressler and Paunovic (2021) highlighted that 'about us' pages often emphasize heritage to attract customers, a theme that is less likely to appear in reviews, which may be a contributing factor to the mismatch between intended brand identity and customer perceptions. Finally, Karadağ et al. (2022) used pre-defined terms from a UK-based study (Rutter et al., 2016), which may introduce validity issues when applied to the context of Turkey. The paper did not assess whether these themes were appropriate or relevant in the new cultural setting.

These critiques highlight the importance of involving human judgement in text mining. While tools like LDA-topic modelling, as demonstrated by Hacking et al. (2023) in their study of interview data on long-term care, can complement analysis, they should not replace human expertise. Text mining lacks the nuanced understanding of emotions and context (Machová et al., 2023) that is essential for fully grasping a brand's position.

By combining qualitative and quantitative approaches, mixed methods offer a flexible solution to these challenges. They help bridge the gap between complex analytical tools and actionable insights for brand managers, providing protection against the rigidity of relying solely on one type of analysis (Harrison and Reilly, 2011). However, mixed methods also demand significant time, resources, and expertise, presenting challenges for researchers (Molina-Azorin, 2016), and therefore marketing professionals. Thus, this paper seeks understand: a) the clarity of themes derived from thematic analysis in understanding consumer sentiments; b) the number of online reviews required to achieve data saturation for brand positioning; c) the effectiveness of CA in providing actionable insights to improve brand positioning based on consumer perceptions.

3.4 Research Questions

Given the shortcomings of previous studies, particularly their over reliance on machine-based methods for understanding textual data, the first aim is to reintroduce a human-based approach to analysing reviews. Thematic analysis is a widely used qualitative method for identifying, analysing, and reporting themes within data (Braun and Clarke, 2006). It is particularly effective in making sense of unstructured text data, such as online reviews, which are increasingly important for understanding consumer sentiments and behaviours. It is simple, flexible and therefore a very accessible method for people wanting to use qualitative methods that do not have extensive training in social science research methods, thus a good choice for marketing professionals. Given the shortcomings in previous text mining-based research, it is important to assess this against the alternative of human-based approaches. Thus, research question 1 is formed as:

Research Question 1: How useful are themes derived from thematic analysis in extracting psychological, behavioural, and sentiment factors from online reviews?

It is unrealistic to expect brand managers to read all available online reviews. Thus, understanding the point at which quantity of online reviews provide sufficient insights for brand positioning is crucial. This research aims to explore the number of reviews needed to achieve data saturation, where no new themes emerge, and how many reviews are required to establish a stable understanding of a brand's position position from research question 3. By employing a bootstrapped resampling approach, the study seeks to determine these thresholds, providing actionable insights for both researchers and practitioners in brand management. Thus, research question 2 is formed as:

Research Question 2: How many online reviews are necessary to achieve data saturation and establish a stable brand position using mixed-methods analysis?

CA is a powerful multivariate graphical technique that reduces high-dimensional data to sub-dimensional space, making complex relationships between categorical variables more interpretable (Greenacre, 2017). In the context of brand positioning, CA could potentially visually represent the relationships between brands and the themes identified through thematic analysis of online reviews. This visual representation is crucial for marketers and brand managers, simplifying the positioning of brands based on key consumer behaviour dimensions revealed by the analysis. Thus, research question 3 is formed as:

Research Question 3: How effectively can CA visualize key dimensions of consumer behaviour and brand-theme relationships from online reviews to offer practical insights for improving brand positioning?

4. Methodology

4.1 Research Philosophy

To overcome the limitations of previous research and develop a more nuanced understanding of online reviews, this study utilises mixed methods. The first qualitative component employs thematic analysis (Braun and Clarke, 2006), focusing on online reviews to uncover essential consumer insights that shape perception and reputations of brands. Subsequently, CA and a heatmap is used to map the relationships between brands and identified themes (Greenacre, 2017), simplifying the data and making the results more interpretable. Resampling is used to better understand how many reviews is needed to achieve data saturation and a stable view of brand position. This pragmatic, flexible methodology adapts to the complexities and uncertainties of real-world industry settings, making it more applicable for brand managers than rigid approaches (Miles and Huberman, 1994, 5). This balances the need for scientific rigor with the practicalities of industry application, while aligning with the broader view that mixed methods can provide both robust and nuanced insights (Greene, 2008).

The use of online reviews as a data source for analysing brand positioning is justified by their exponential growth and their rich nature of consumer opinions, which can provide deep insights into a brand's reputation and position in consumer minds Liu et al. (2024). Additionally, Ismagilova et al. (2020) emphasize that eWOM is a critical factor influencing purchase decisions, making it an invaluable resource for brand managers.

4.2 Data

For the study, the hugging face repository of amazon reviews were used (McAuley Lab, 2023), these encompasses individual reviews across all of Amazon's product categories. The dataset includes information such as the product name, a unique product identifier and a data-time stamp. Since the dataset was missing the brand names, these were matched via the amazon website using the product identifier. The data sets are split into product categories and for the present study, the Home and Kitchen product category, which contains 23 million reviews, were the starting point[1]. Following this, a search was conducted on www.amazon.com to filter out the target products of interest. Filters on the websites were used to find espresso and coffee pod machines between 200-500USD. This is assumed that this is a rough price range that people work in when considering their first venture into the coffee machine world and thus

[1] Breville in this context refers to the Australian Breville brand, which is marketed as "Sage" in the UK and Europe, and not Breville UK, a completely different brand.

likely competitors products are all found within this price range. It is important to note that some of the prices have changed since writing this so if this study is recreated, it may result in a very slightly different set of products. This yielded 128 unique products across 13 different brands. Following this, the unique product identifier was collected from each of the 13 product's amazon.com page. This created a list of individual asins for products who are all likely competing for space in the same market. Only the top 5 brands in terms of number of reviews were kept to ensure that they were genuine competitors of each other and that there were a sufficient number of reviews for each product to creat. a meaningful analysis. The dataset was further reduced to only those reviews during 2023 producing a final dataset of 872 reviews taken froward to thematic analysis.

4.3 Thematic Analysis

This study conducted an exploratory thematic analysis using the method proposed by Braun and Clarke (2006), using MAXQDA24 software. The analysis aimed to identify potential emergent themes within the reviews, allowing themes to arise organically from the data rather than imposing existing theories. While existing consumer psychology themes, such as the 3Es (Park et al., 2016), could have guided the analysis, an inductive, emergent method was preferred. Enabling the researcher to engage with the data, bridging open coding with the development of new themes and theories. (Charmaz, 2008, 155).

The choice of an emergent method over confirmatory or deductive methods is twofold. First, the primary goal of the study is to assess the usefulness of these methods rather than to validate established consumer psychology theories. Second, since this approach is relatively unexplored, using pragmatism allows for flexibility in discovering and explaining new properties within the data, making it particularly suitable for a novel topic or method (Saldaña, 2021, 141). This flexible, data-driven technique ensures emergent themes reflect authentic perspectives (Morgan, 2007).

The thematic analysis followed six stages:

Familiarization with Data: In-depth review of the data to gain a comprehensive understanding and note initial impressions.

Generating Initial Codes: Systematic development of codes from the data, annotating the text with labels that summarize relevant information.

Searching for Themes: Collation of different codes into potential themes without fitting them into a pre-existing framework.

Reviewing Themes: Iterative refinement of themes to ensure they relate accurately to the codes and the overall dataset, discarding any that do not hold up.

Defining and Naming Themes: Refining the specifics of each theme and exploring their dimensions and interrelations, without finalizing them prematurely.

Producing the Report: Compiling the findings to reflect the exploratory nature of the analysis, linking the emergent themes back to the data without confirming fixed categories.

The identified themes were aggregated at the brand level, creating a contingency table of brands against their themes.

4.4 Correspondence Analysis

To assess the relationships between brands and themes, CA was employed using *ca* package (Nenadic and Greenacre, 2007) in R (R Core Team, 2024). The primary goal was to create a bi-plot that visually clarifies these relationships, making relationships more apparent than they would be on the original contingency table. Typically, this visualization is achieved on either two principal axes for 2D visualization or three principal axes for 3D visualization.

However, a key limitation of CA is the potential loss of information when reducing data complexity to lower-dimensional space. Therefore, it is crucial to balance improvements in interpretability (e.g., fewer dimensions) with information retention. Whether a plot has retained enough information is often context-dependent and left to the discretion of the researcher or the specific needs of the situation. In this study, the decision to focus on two principal axes with a following discussion on how much information has been retained and its consequences.

While CA has been chosen for its simplicity and ease of interpretation for managers, this paper also provides a brief explanation of the underlying mathematics of the brand-theme relationships to ensure a better understanding of the method. Although the mathematical complexities of CA can be handled by computer software, those interested in exploring the theoretical foundations further are encouraged to consult works by Greenacre (2017) and Hjellbrekke (2018).

To achieve this, the inner product was calculated to create a new vector A for each brand where A produces a set of values such that the larger the value, the more closely related the brand is to that given theme:

$$1. \qquad \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix}_n = x_{brand_n} \cdot \begin{pmatrix} x_{theme_1} \\ x_{theme_2} \\ \vdots \\ x_{theme_n} \end{pmatrix} + y_{brand_n} \cdot \begin{pmatrix} y_{theme_1} \\ y_{theme_1} \\ \vdots \\ y_{theme_n} \end{pmatrix}$$

4.5 Resampling Approach

Because of concerns over the workload of reading large sets of reviews, and to better assess the applicability in a real-world setting, an iterative bootstrapped resampling approach was employed (figure 1). This involves drawing random samples from the data at growing sample sizes. The aim is to test the effect of sample size for data saturation and brand-theme relationship stability. Methods are justified because they provide a powerful framework for assessing the stability and reliability of results in statistical contexts (James et al., 2021).

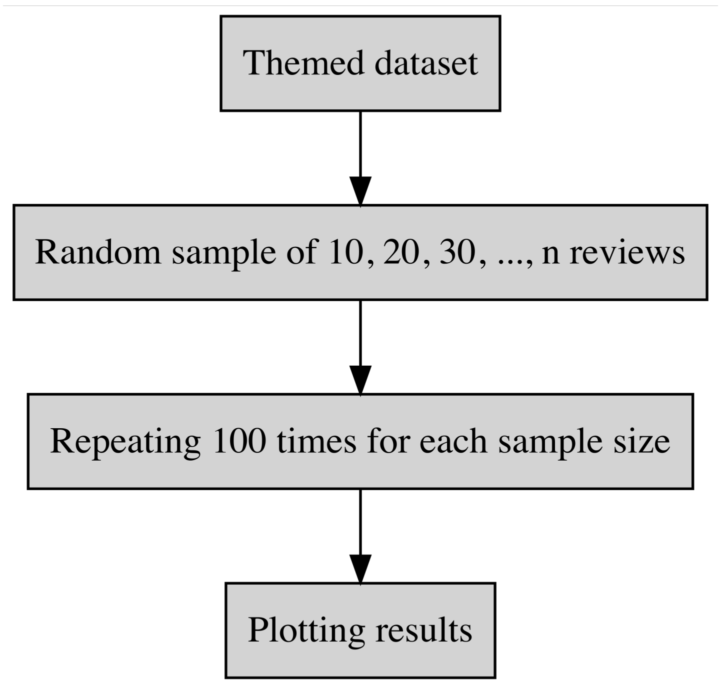

Figure 1. Resampling workflow

Data saturation. Resampling was to evaluate the progression of thematic development in the reviews and identify the point of data saturation. This aids in the situation where thematic analysis may be the only goal. This approach involved repeatedly drawing random samples from the already themed dataset, starting with subsets of 1 review and increasing the sample size incrementally by 50 reviews until the full dataset was reached, each sample size subset size was resampled 500 times. During each resample, the number of discovered themes in that sample was counted and placed in a table. The median number of themes for each sample size will be plotted.

Bi-plot map stability. The stability of the relationships between brands and themes, as represented in the bi-plot from the CA, was also evaluated using resampling. The goal was to determine the minimum sample size necessary to produce reliable and stable brand positioning in the bi-plot visualization. This was done by calculating the distance between each brand and every theme on the bi-plot (see equation 1, above), with this process repeated 100 times for each sample size, beginning with 50 reviews and increasing in increments of 50 until the entire dataset was analyzed. The standard deviation of these distances was calculated to assess the variability at each sample size, which was then plotted to determine how sample size impacts the variance and, consequently, the stability of the bi-plot. To further motivate this, see figure 2. These have been created from synthetic data designed such that the top-two plots represent a lower sample size and, therefore, greater variance; while the bottom 2 show a simulation of greater sample size and greater similarity.

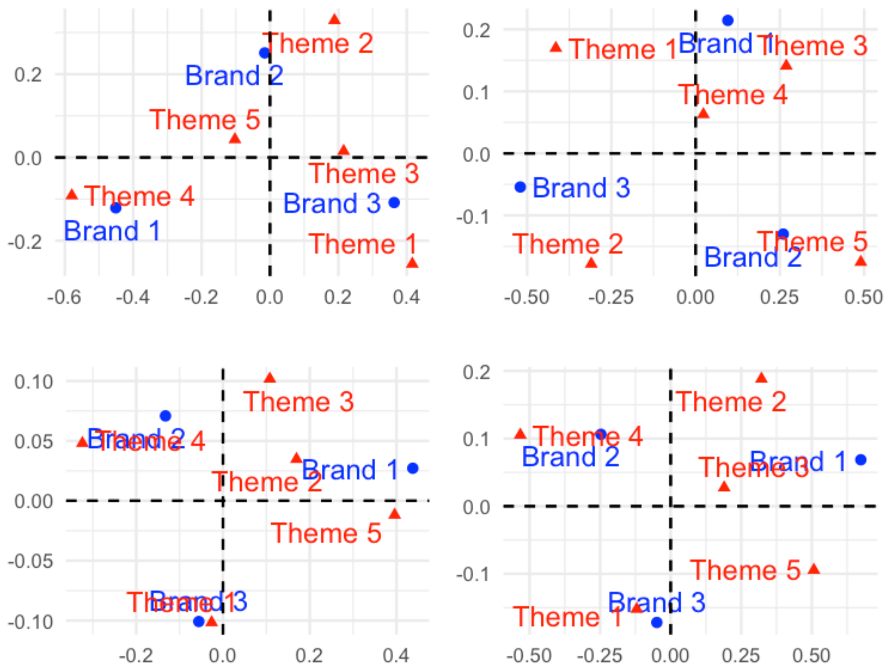

Figure 2. Simulated bi-plots

4.6 Ethics

Since these reviews are posted in public forums where visibility to anyone is expected, explicit informed consent is not typically required. The University of Lincoln's ethical guidelines confirming no need for ethics approval for research conducted on secondary public data. The dataset used did not include personal identifiers (e.g., names, usernames), thus respecting privacy in line with general ethical guidelines from a psychology perspective The British Psychological Society (2021), and from a market research perspective (Market Research

Society, 2023). The data set includes all reviews from the Amazon.com domain, filtered only by price and date to fit the study's scope, ensuring no additional selection criteria were applied. In terms of research bias and concerns about misuse of data, such as fabricating themes, are unfounded, as the researcher has no vested interest in the outcomes since study's focus is on the method, not the specific themes that emerge.

5. Findings

5.1 Thematic Analysis Results

The thematic analysis highlights diverse user experiences with the coffee machines, encompassing both positive and negative aspects. While ease of use, taste, and design were praised, challenges such as poor durability, complex setup, and unsatisfactory customer support were also noted. This analysis deliberately separates positive and negative aspects within each theme to provide a clearer understanding of the distinct factors that influence brand reputation, recognizing that consumers seek detailed insights into both challenges and benefits when reading reviews.

Theme 1: Ease of Use - Positive. This theme captures how the coffee machine's ease of use enhances the user experience. Reviews highlighted several sub-themes, including the simple and intuitive interface and the quick setup process. For example, one user stated "it is very beginner friendly, easy to understand," and another stated, "the DeLonghi TrueBrew automatic coffee machine is so much easier to use than my previous coffee machine."

Theme 2: Ease of Use - Negative. This theme encapsulates the challenges users encountered, including sub-themes like complex navigation and a difficult setup process. One reviewer shared, "The learning curve on this machine is far too steep," while another commented, "The water reservoir is difficult to see, so be careful when adding water and do not overfill, otherwise it will make a colossal mess in your kitchen".

Theme 3: Investment and Value for Money. This theme captures reviewers' assessments of the product's financial value. Some deemed it worth the investment, particularly when compared to the cost of buying coffee out. For example, one user stated "such a bargain at the price (paying for itself repeatedly over the years) and an essential household item for my family" Another user noted that the machine is expensive but worth the cost "…easy to clean (so far! customizable beverages. Overall, it's a big investment and one that we've been happy to make. No regrets!"

Theme 4: Good Taste. This theme highlights positive aspects of the taste of coffee from the product. One review noted, "And let me tell you, the Breville Infuser does not disappoint. From the optimal flavor in every cup…" while another commented, "There are so many great reviews here that I highly recommend reading for tips and tricks. I am just chiming in to say that my machine made beautiful espresso from thefirst pull…"

Theme 5: Poor Taste and Performance. This theme addresses negative feedback on taste and performance, highlighting that the machine, though functional, did not meet customer expectations. Sub-themes include disliked flavor and unpleasant taste. Reviewers commented, "…I bought this Phillips 3200 approximately a week ago and immediately started receiving very watery coffee. I had gone through all the proper preparation, and I understood that it might take some time… The grounds were inconsistent and the coffee was undrinkable." and "Like many of the low reviews mention, this machine does NOT make even decent espresso. It consistently comes out watery and there is no setting combination to fix that…". Regarding poor performance, one review stated, "If you want to use the steam function, expect to wait a looooong time. I wish we would have splurged (even further) on the model with the second pump, or just bought a simple espresso puller." while another mentioned, "The pressure gauge on this machine is utterly useless. Yeah, the needle moves around and stuff, but it doesn't read anything useful, and doesn't read at all when steaming."

Theme 6: Poor Customer Support. Reviews highlighted unresponsive andineffective customer support. One reviewer noted, "…3 days away I came back and the steam and hot water did not function. I called customer support, their training extends to if it doesn't function send it in for $264 to get repaired, they have no technical knowledge. I thought that I would de scale it again, the water and steam function worked fine during the descaling and also when I made a coffee again …." and another noted, "…Red lights keep going on and nothing will work. Called customer service and very unhelpful for a 500 dollar machine…".

Theme 7: Durability - Negative. Users highlighted poor durability and reliability, noting weak build quality and easy breakage. One reviewer noted, "After having for under 2 years, the machine already broke despite keeping up with maintenance." while another commented, "Wont suggest anyone else buy one. It leaks all over my bar!!"

Theme 8: Durability - Positive. This theme emphasizes the product's durability and reliability, with many reviews mentioning long-term daily use. For example, one review stated, "Great quality espresso for two, twice a day, and trouble-free for seven years. Easily the best, most durable appliance I have yet purchased. I owned three pump-driven espresso machines before this one, including two Italian-made models, and none of them came close." and another remarked, "The milk frother is a game changer. Even the durable construction and good quality espresso are impressive. While there are areas for improvement…"

Theme 9: Aesthetic Appeal. Reviewers provided positive comments about the aesthetics of the product, many of them stating how it "upgraded" their kitchen. One reviewer shared, "My initial interest for this coffee maker, was sparked due to the sleek and absolutely beautiful design! While the size of it is rather large, I was actually really happy with the display of it on my kitchen counter!" and another noted its beautiful design, "After going through a number of coffee makers - we have finally hit the jackpot! Beautiful design, great function, and easy to use!"

Theme 10: Lifestyle. This theme highlights how the coffee machine enhances daily routines, offering convenience, enjoyment, and barista-level quality at home. One reviewer noted, "I did not buy to replace a Starbucks expense (I rarely go) but rather to enjoy as an everyday enjoyable ritual producing near expert barista level results and it does this very well". Two reviewers stated this was, or should be, part of their work break routine, now "It is so easy now to make great coffee with a flick of a button. Mid-morning breaks with custom espresso latte right in the kitchen are a cinch (and a must!)." and "This would make a great addition to any home; it would also be a really useful asset for a small office breakroom (hint, hint!)."

5.2 Descriptive Quantitative Results

Figure 3 shows the total themes from user reviews for each coffee machine brand. Breville has the most reviews, followed by De Longhi, Philips, Casabrews, and Gaggia. This may indicate that either Breville has higher sales or that its users provide more detailed feedback. The data alone does not clarify if more themes result from higher sales or greater user engagement.

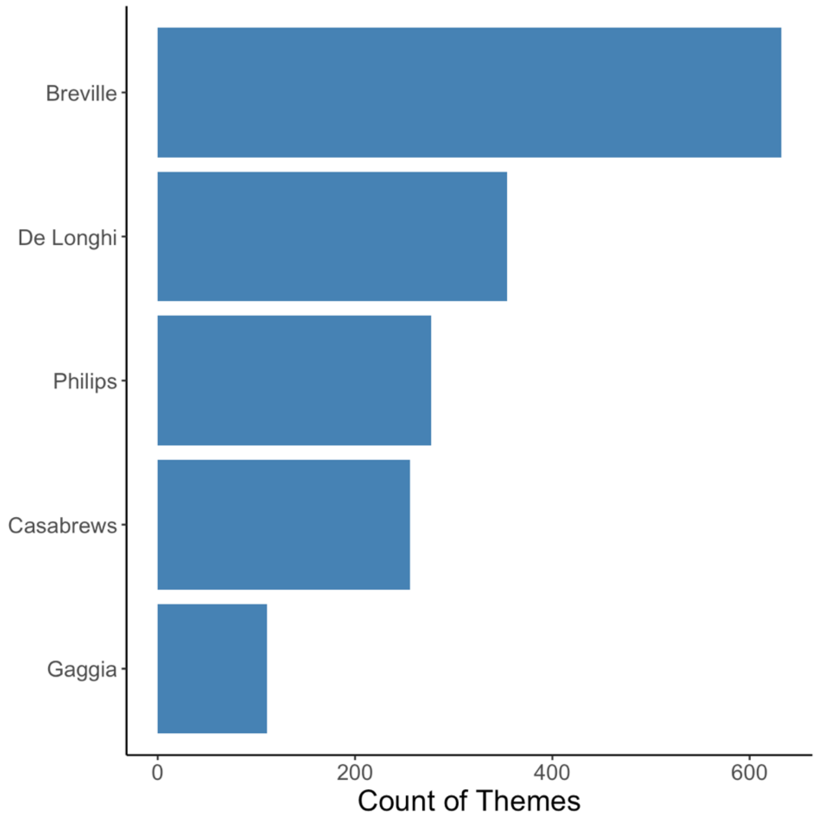

Figure 3. Brand review counts

Figure 4 shows total themes from user reviews of coffee machine features. “Ease of Use - Positive” and “Investment and Value for Money” are the top themes, showing key user concerns. “Poor Customer Support” and “Negative Taste and Performance” are less mentioned, indicating fewer concerns.

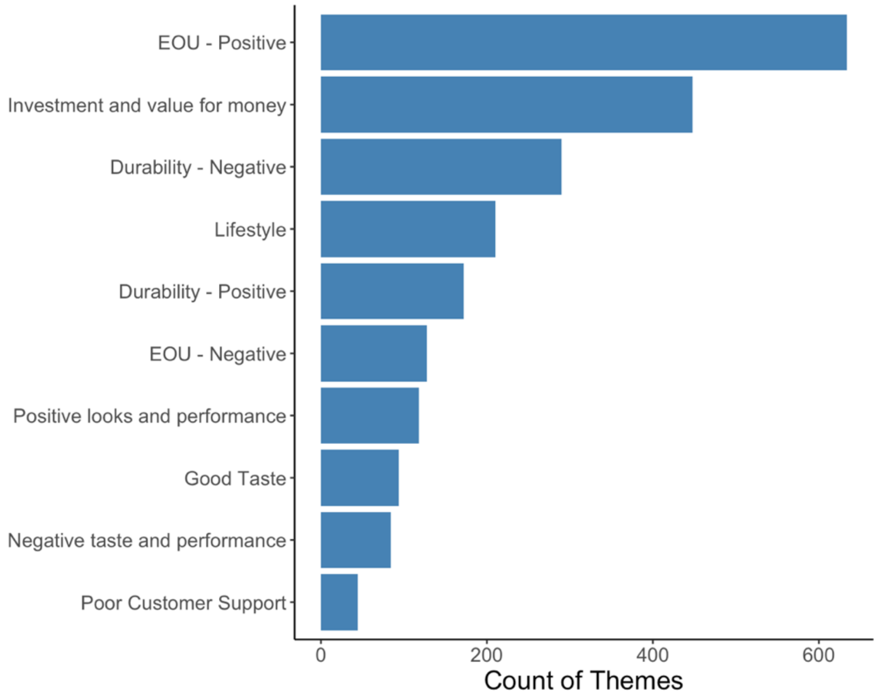

Figure 4. Counts of themes

Table 1 summarizes themes in user reviews for coffee brands with the rows representing the emergent themes, and the columns representing the brands. Each cell contains a count of how many times a particular theme was mentioned for a given brand. This data provides an overview of user feedback distribution across different themes for each brand. The brand-theme results will be examined in the following.

	Breville	Casabrews	De Longhi	Gaggia	Philips
EOU - Positive	96	60	87	20	54
EOU - Negative	31	7	12	6	8
Investment and value for money	97	49	23	13	42
Taste Positive	19	5	8	6	9
Negative taste and performance	15	7	6	3	11
Poor Customer Support	11	2	7	1	1
Durability - Negative	46	9	46	11	33
Durability Positive	36	4	14	12	20
aesthetic appeal	31	4	18	6	0
Lifestyle	58	11	18	7	11

Table 1. Brand-theme data from amazon reviews

5.3 Resampling Results

Data saturation. To assess theme saturation, 100 bootstrapped samples were taken, with sizes from 1 to the full dataset, in intervals of 10 (figure 5). Results show that after a sample size of 50, a median of 14 themes are found, indicating the total themes in the dataset. Thus, 50 reviews would have been sufficient to achieve data saturation in the present study.

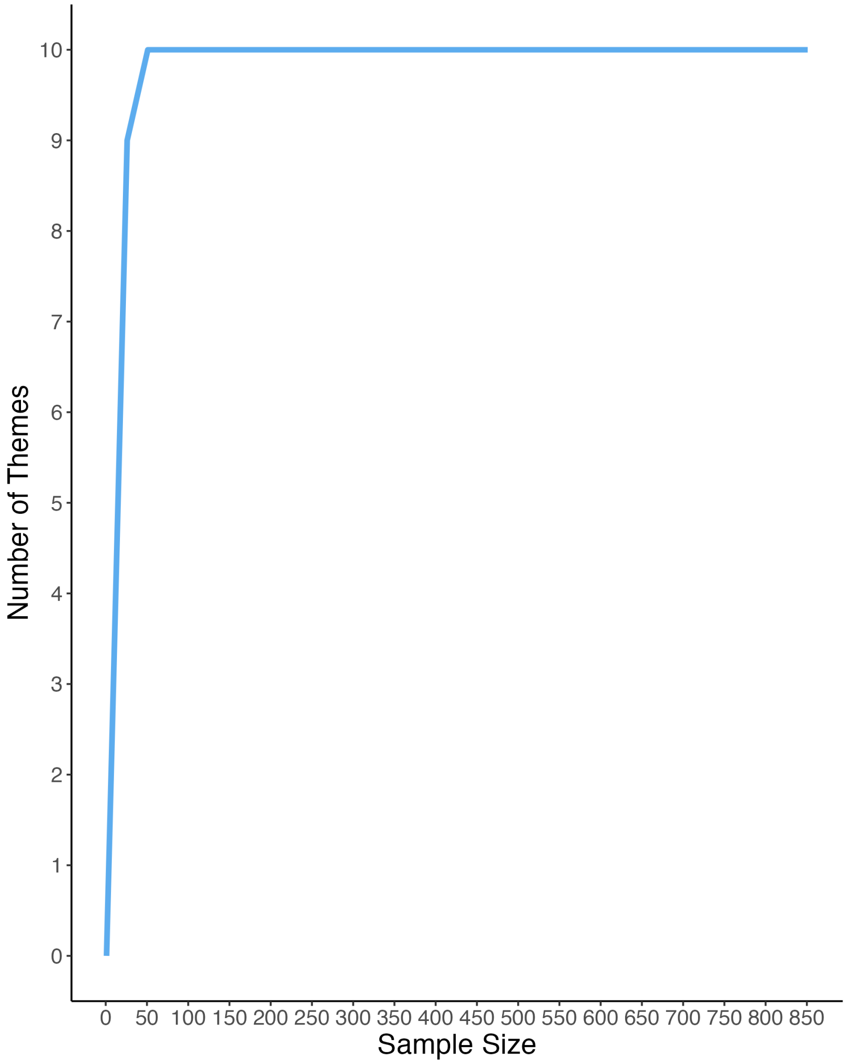

Figure 5. Data saturation across sample sizes

Bi-plot map stability. The standard deviation of brand-theme distances on the bi-plot was calculated across different sample sizes, resampling 100 times at each interval (figure 6). Results show standard deviation decreases with larger samples, indicating more stable brand-theme relationships. While stability improves, the rate slows after a sample size of 300, showing diminishing returns.

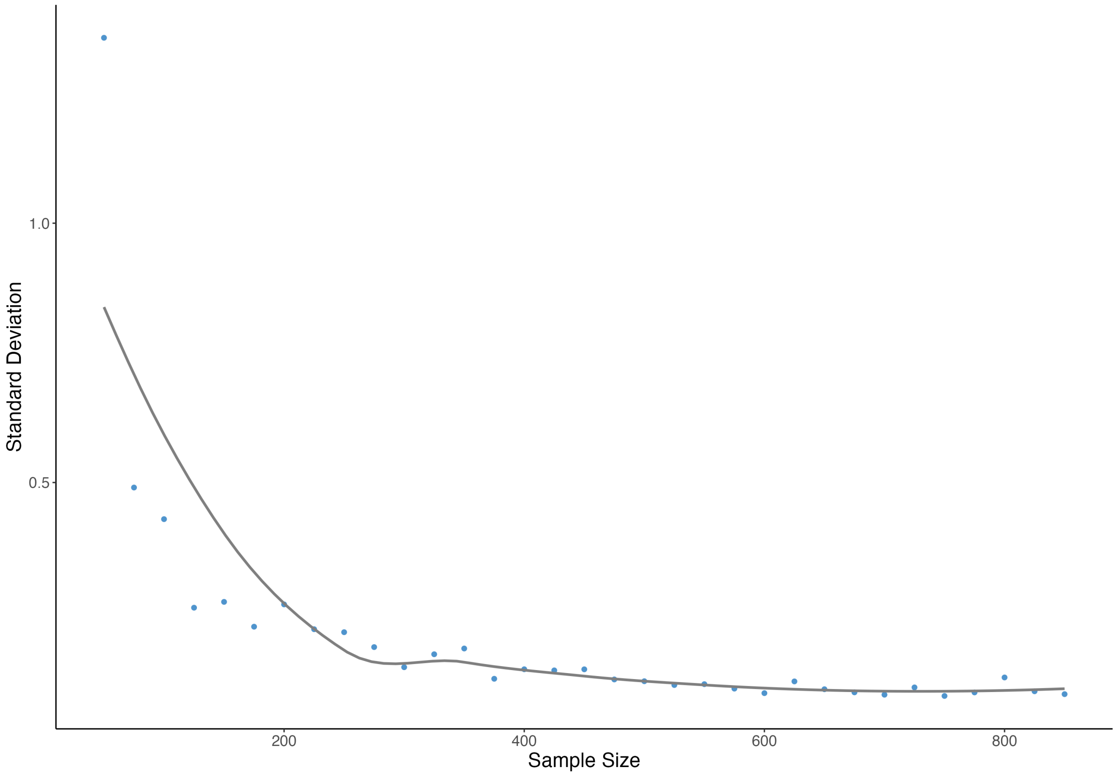

Figure 6. How sample size affects variance of the brand-theme relationship

5.4 Correspondence Analysis (CA) Results

The principal axes, or dimensions, obtained from the CA provide meaningful insights into the associations between brands and their themes. Together, these two dimensions explain 74.7% of the total variance. Figure 7 shows the final bi-plot.

5.4.1 Principal Dimension 1: Perceived Value vs. Aesthetic and Functional Appeal

This dimension (Figure 8, left side) explains 34.8% of the variance and captures the contrast between practical financial considerations and aesthetic and functional characteristics. The left side emphasizes *Investment and Value for Money*, and *Durability - negative*, with *Investment and Value for Money* contributing significantly (51.78%, left side), and *Durability* making a moderate contribution (24.01%, left side). The right side highlights *aesthetic appeal*, which contribute strongly (15.15%, right side).

5.4.2 Principal Dimension 2: Ease of Use vs. Lifestyle

Explaining 39.9% of the variance, this dimension (Figure 8, right side) contrasts *Ease of Use* with factors like *Lifestyle*, *aesthetic appeal*, *Performance*, and *Durability*. The top of this axis highlights *Ease of Use* (32.77%, top) and *Durability* (10.89%, top), while the bottom reflects *Lifestyle Integration* (25.18%, bottom), *aesthetic appeal* (15.97%, bottom), and *Durability* (3.08%, bottom).

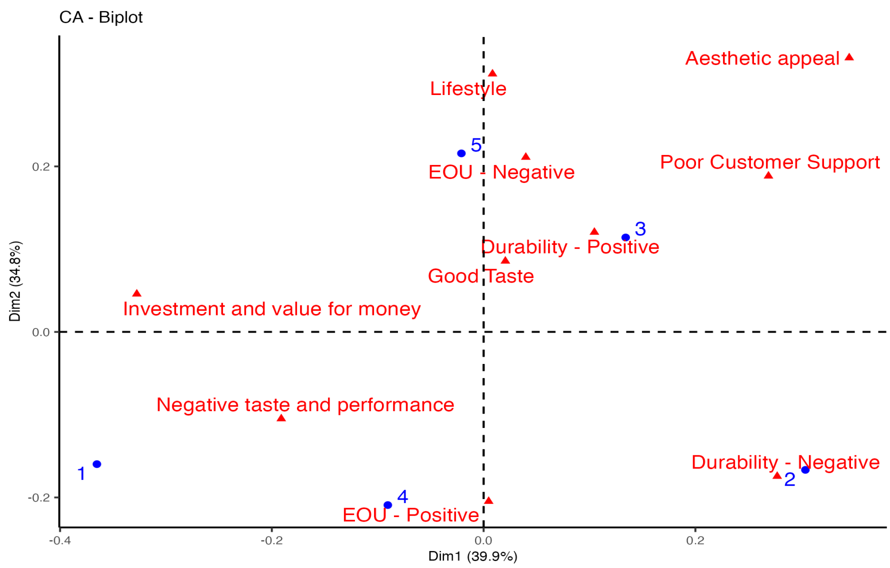

Figure 7. Correspondence analysis bi-plot showing brand positions: 1 = Casabrews 2 = De Longhi, 3 = Gaggia, 4 = Philips, 5 = Breville,

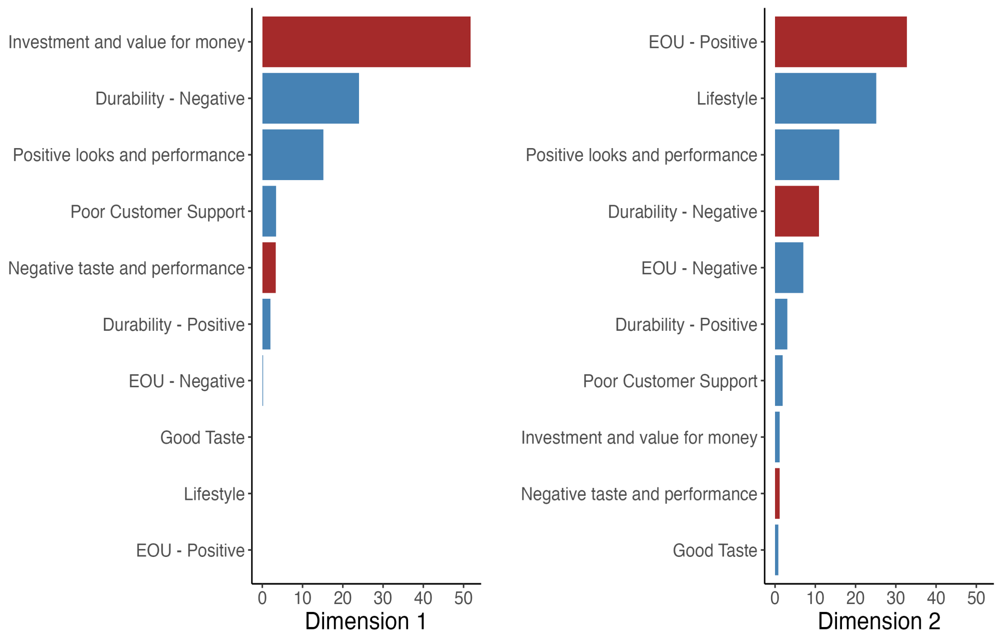

Figure 8. Contributions of themes to dimension 1 (Perceived Value vs. Aesthetic and Functional Appeal) on the left and dimension 2 (Ease of Use vs. Lifestyle Integration) on the right. The blue/red colouring indicates whether they are on the left or right of dimension 1, or top or bottom of dimension 2.

5.5 Brand-theme association

A heatmap (figure 9) was created to clarify the associations using equation 1, as it is important to avoid assuming that categories are closely related based on proximity on the plot. Breville is most closely linked to "Lifestyle" and "aesthetic appeal and performance," and least associated with "EOU - Positive" and "Durability - Negative." Casabrews is most associated with “Investment and value for money” and “Negative taste and performance,” and least associated with “aesthetic appeal and performance” and “Poor Customer Support.” De Longhi is most closely associated with “Durability - Negative” and “Poor Customer Support,” while the least association is with “Investment and value for money” and “Lifestyle.” Gaggia is most associated with “aesthetic appeal and performance” and “Poor Customer Support,” and least associated with “Investment and value for money” and “Negative taste and performance.” Finally, Philips shows the highest association with “EOU - Positive” and “Negative taste and performance,” and the least association with “aesthetic appeal and performance” and “Lifestyle.”

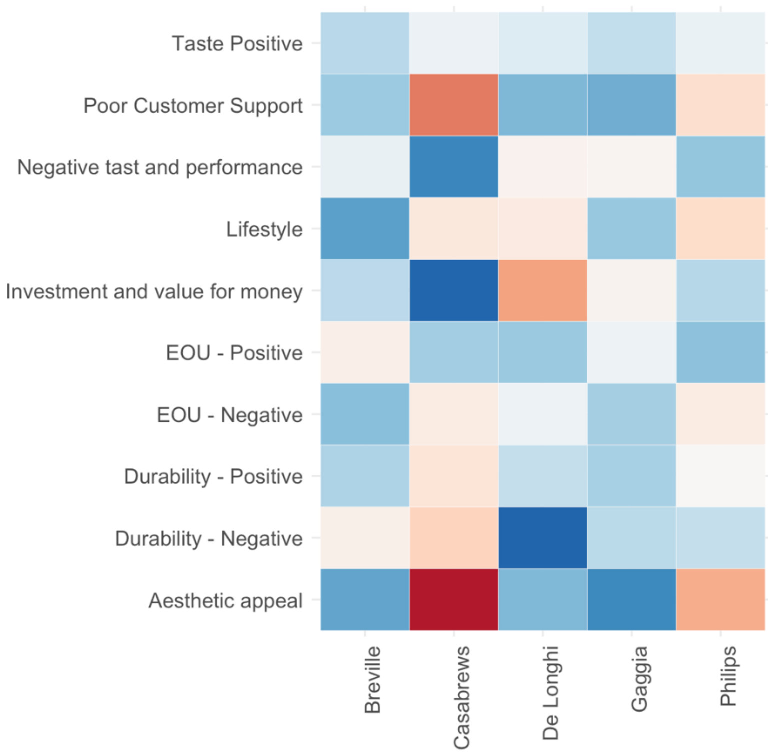

Figure 9. Heatmap results from the correspondence analysis. Darker blue colour represent higher associations, while darker red represents lower associations.

5.6 Summary of Findings

Thematic analysis found key themes like ease of use, sentiment, taste, customer support, and durability, showing varied user experiences. Positive ease of use and strong recommendations were frequently noted, while setup challenges and customer support issues were common negative aspects. Both taste and durability received mixed feedback, reflecting varied user experiences.

The bootstrapped sampling results clearly show evidence of data saturation at about 50 reviews, and diminishing returns on after 300 reviews for bi-plot stability. The bi-plot effectively visualizes the relationship between brands and themes, offering actionable insights into brand positioning by highlighting key associations such as user experience and product value. Finally, the heat map provides a clear and detailed illustration of the brand-theme associations, revealing distinct patterns in customer perceptions that can guide targeted brand strategy and improvement efforts.

6. Discussion

This study aimed to address the shortcomings of purely machine-based text mining methods by incorporating techniques that preserve nuance, context, and word order, while also prioritizing practical, accessible approaches for real-world marketing applications, rather than relying on overly complex methodologies. The research focused on three key questions by employing a mixed-method approach that combines qualitative thematic analysis with quantitative techniques.

First, how useful are themes derived from thematic analysis in extracting psychological, behavioural, and sentiment factors from online reviews? Second, how many online reviews are necessary to achieve data saturation and establish a stable brand position? Finally, can CA be used to gain actionable insights for improving brand positioning based on consumer perceptions?

6.1 Research Question 1

The thematic analysis effectively addresses research question one by identifying key psychological, behavioural, and sentiment factors such as ease of use, taste, customer support, and durability. This approach uncovers nuanced insights from online reviews that previous methods, LIWC, bag-of-words, and keyword approaches, have not achieved. While those methods often lose context and fail to capture the depth of consumer sentiment, thematic analysis allows for a more comprehensive understanding of consumer experiences, both positive and negative. This demonstrates the strength of thematic analysis in providing a richer, more detailed exploration of consumer feedback, making it a superior method for understanding complex factors that influence brand perception.

The present study builds on the existing body of research by employing mixed-methods to analyse eWOM for brand positioning. This approach integrates thematic analysis with CA, aiming to offer a more nuanced and accessible method for understanding brand positioning in the context of online reviews. To evaluate the effectiveness and novelty of this approach, a comparative analysis with existing studies is warranted, focusing on methodology, data handling, and the practical implications of findings.

6.1.1 Methodological Integration and Novelty

Compared to previous studies, the mixed-methods approach in this research offers a balanced blend of qualitative and quantitative techniques, which is relatively underexplored in eWOM analysis. For instance, studies like Taecharungroj (2022) and Sun et al. (2023) predominantly rely on quantitative text-mining methods such as LDA and sentiment analysis. While these methods are effective in processing large datasets, they often suffer from significant limitations, particularly the loss of contextual understanding and word order, as highlighted in the critique of these studies. The LDA model used in Taecharungroj (2022), for example, was effective in identifying topics but failed to preserve the deeper nuances and context within the reviews, leading to a less reliable understanding of consumer sentiments and experiences.

In contrast, the present study's use of thematic analysis addresses this gap by allowing for the identification of themes that emerge naturally from the data, preserving the richness and complexity of consumer feedback. This is particularly important when dealing with unstructured textual data, as it enables the researcher to capture subtleties that automated methods like LDA might overlook. For example, the theme of "Ease of Use" in this study was split into positive and negative aspects, a distinction that might be lost in a more automated approach.

6.1.2 Handling of Context and Nuance

The issue of context loss, particularly in the application of lexicon-based methods like LIWC as seen in Alzate et al. (2022) and Prihananto et al. (2024), is a critical shortcoming in many text-mining approaches. These studies, while effective in categorizing text into pre-defined psychological or linguistic variables, often miss the context in which words are used, potentially leading to misinterpretations. For example, Alzate et al. (2022)'s reliance on PCA to reduce LIWC themes into components for brand positioning resulted in a loss of nuanced insights that could have been crucial for a more accurate depiction of brand perception.

The present study circumvents this issue by using a more human-centric approach to data analysis, where themes are derived directly from the text without the constraints of predefined dictionaries. This allows for a more flexible and contextually aware analysis, ensuring that the themes are genuinely reflective of consumer sentiments. The thematic analysis process employed here, following the guidelines of Braun and Clarke (2006), ensures that the context and nuance of consumer feedback are preserved, leading to more meaningful and actionable insights.

6.1.3 Temporal Relevance and Review Data Handling

Two studies took their data from a wide date span. Alzate et al. (2022) collected all available reviews up to Feburary 2017, and Taecharungroj (2022) earliest review of the dataset was published in December 2004 and the latest one was published at the end of April 2021. This can dilute the temporal relevance of the findings.

This is particularly important in the context of shifts in consumer behaviour and attitudes, as with the recent moves towards a concern for greener products (White et al., 2019). Furthermore, brand positioning changes, not just from shifts in consumer beliefs but changes in product or service offering from both the focal company, and the competitors. By contrast, the present study focused on a dataset of reviews from a single year (2023), ensuring that the findings are temporally relevant and reflective of current consumer sentiments, behaviours, thus it is a more accurate representation of true brand position.

6.1.4 Practical Implications and Accessibility

This study aims to bridge the gap between academic research and practical application, particularly for brand managers lacking advanced training in statistical or text-mining methods. Unlike studies such as Dressler and Paunovic (2021) and Sun et al. (2023), where complex methods like netnographic analysis and advanced sentiment analysis limit practical usability, this study employs thematic analysis, a simple and flexible method that allows brand managers to engage directly with consumer feedback without needing extensive methodological expertise. However, relying on human-led research could increase the workload for brand managers compared to machine-led approaches, especially in studies involving large datasets like Taecharungroj (2022), which included over 51,000 reviews. To address this, a resampling approach to evaluate data saturation and the stability of the CA biplot was introduced.

6.2 Research Question 2

This study challenges the conventional wisdom in brand positioning research, particularly the assumption that analysing large volumes of online reviews is necessary for meaningful insights. For instance, studies such as Taecharungroj (2022), which analyzed over 51,000 reviews, exemplify the data-heavy approaches that dominate the field. However, my research provides compelling evidence that such extensive data collection may not only be unnecessary but could also dilute the clarity and actionable value of the findings.

The present study contrast previous research by demonstrating that thematic saturation can be achieved with as few as 50 reviews and that stable brand-theme relationships are established with around 300 reviews, highlighting a more efficient and targeted approach to data analysis. The results suggest that beyond a certain point, additional reviews contribute little to improving the depth of insight, effectively introducing diminishing returns. In fact, the excessive volume of reviews analyzed in studies like Taecharungroj (2022) may introduce noise and unnecessary complexity, rather than enhancing the understanding of brand positioning.

This research advocates for a reevaluation of methodological approaches in brand positioning studies. Instead of prioritizing the sheer quantity of data, researchers and practitioners should focus on the quality and relevance of the data they analyze. These findings indicate that a smaller dataset with a more robust qualitative analysis might yield more useful results.

6.3 Research Question 3

Two main dimensions emerged from the analysis: “Perceived Value vs. Aesthetic and Functional Appeal” and “Ease of Use vs. Lifestyle,” each revealing distinct factors that influence consumer preferences and brand positioning.

The study’s findings align closely with the principles of Davis (1989) that perceived ease of use is critical in technology adoption. Dimension 2 ease of use, both positive and negative, was a significant factor influencing customer sentiment. These findings illustrate how technology acceptance research, originally designed for IT systems, may also to consumer technologies such as coffee machines. As a consequence, readers of reviews may be influenced by the review writer’s mentions of ease of use of the machines.

However, while ease of use and durability have been identified as important factors across both dimensions, research suggests that their competitive advantage is diminishing. Commoditization of performance-based features are becoming an expectation for consumers Fujimoto and Miller (2007). This is leading to competitive shifts from functional factors like durability and performance to emotional values, driven by design and user experience (Noble and Kumar, 2008). Companies that continue to fall short on performance-based features risk losing relevance, as consumers now expect these as a given before considering emotional value or user experience. In the present study, this may explain why the themes **Durability - Negative**, and **Investment and value for money** are at polar opposites on dimension 1, customers feel a strong sense of poor investment if the basic performance needs are not met. In addition, **Durability - Positive**, and **Investment and value for money** are on the same side and, despite their alignment, **Durability - Positive** was not a significant contributor to the dimension. This suggests that failure to meet customer expectations around performance has a stronger impact on percieved value than meeting their expectations on this theme.

Aesthetic appeal and lifestyle were important contributors to dimension 2. Toufani et al. (2017) found that while aesthetic appeal positively influences purchase intention, this effect is moderated by emotional and social value, meaning that aesthetics drive purchases through the emotional satisfaction and social status they provide to consumers. This research is supported by Bell et al. (1991) who found that consumers may believe that by choosing visually appealing products they can improve their social status. In the present study, the emotional components of aesthetics and lifestyle was at polar opposites with the perception that the product was easy to use. The present finding is support by Buker et al. (2022), who state that emotional appeal are often difficult to balance, as the product requirements stemming from emotional user preferences frequently lead to conflicting goals.

Existing research often utilises Principal Component Analysis Alzate et al. (2022) or CA bi-plots (Karadağ et al., 2022) to visualize brand positioning. While effective, methods like CA require users to understand how to calculate the inner-product between column and row vectors (i.e., between brands and themes) to accurately interpret the data. This complexity can be a barrier for brand managers who may not be familiar with advanced mathematical concepts, limiting the practical utility of bi-plots. The present study advances on existing research in a simple way by introducing heatmaps as a complementary tool to CA bi-plots. Heatmaps offer a more intuitive and straightforward way to visualize the relationships between brands and themes, eliminating the need for complex calculations. This approach assumes that software or tools can handle the underlying linear algebra, presenting the results in an accessible format that is easy for brand managers to interpret. This offers clear practical benefits for brand managers since heatmaps provide a clear, actionable view of brand positioning. They can quickly identify strong or weak associations with key themes, enabling more informed decision-making.

7. Conclusion

The present study demonstrated that thematic analysis offers richer insights into consumer sentiments compared to automated methods, capturing psychological and behavioral factors. It also found that significant insights can be gained without reading large volumes of reviews. The thematic analysis revealed also that key themes, such as ease of use, taste, customer support, and durability, influence user experiences with coffee machines. Positive aspects like the machines being considered easy to use and good value for money were frequently praised, while negative aspects included complex setup, poor customer support, and issues with durability.

7.1 Implications for Practice

Marketing professionals could consider adopting mixed-methods approaches to analyze consumer sentiment in a way that captures nuance and context, while also providing actionable, data-driven insights. Popular tools such as Qualtrics (Qualtrics, 2023) have an abundance of textmining options within their interface, but are deeply lacking in rodust mixed methods approaches. Building a more robust system could help to bridge the gap between academic rigor and practical application. Marketers do not need to analyze an overwhelming volume of reviews to understand brand perception. Focusing on around 50–300 reviews can provide meaningful insights without unnecessary data overload, allowing for more efficient use of resources. One important finding from the CA and its discussion is that even successful brands often struggle to balance ease of use with emotional design. Overcoming such issues could lead to a new wave of brand adopters that are attracted by emotional design aspects, but intimidated by complexity.

7.2 Limitations

The data from the current research come from a single source, and were codeded and thematacized by a single researcher, presenting two problems. Firstly, it could be that some shopping or review platforms attract a different audience and thus different issues discussed on their reviews. Furthermore, platform design has been shown to influence how consumers share product information (Adjei et al., 2022; Chan and Yang, 2021). Triangulating across varied collection modes and sources is argued to increase the credibility of findings and interpretations (Nowell et al., 2017). Secondly, a single researcher coding the data may have limited the quality and depth of the themes while increasing the risk of some themes being missed Nowell et al. (2017).

Balancing the merging of themes is crucial, as too much merging can lead to loss of nuance, while too many themes can obscure important details. The “correct” level of merging still remains unclear and a different approach may well have produced a very different outcome. Secondly, the level at which themes are merged will result in different points of data saturation and the stability of the bi-plot found in the resampling. It is important that researchers and practitioners do not take the values in the present study as a definitive cut-off but more so as a platform on which to build further knowledge around how much qualitative data is really needed to achieve desired outcomes.

New consumer interests and attitudes change over time, such as the present rise in green issues (White et al., 2019) and how cosnumer data is gathered and used (Grewal et al., 2019). Thus, practitioners who want to study and employ such tools for brand positioning must not become complacent and be ready to find new emergent themes. Another key point is that it is extremely unlikely that these same themes carry over to other product categories, therefore it is very important that each new product category is approached from scratch.

While this practical and pragmatic approach comes from the fact that CA makes the results of a qualitative analysis much more easily interpretable to brand managers, it is unlikely to be adopted in its current form. Firstly, CA is still a relatively unheard of statistical approach outside of academic social science research. Mixed methods remains a challenge for many researchers, more so professional marketers. Any future attempts at achieving adoption of such methods must take in to account key aspects from technology acceptance research.

7.3 Future Recommendations

The TAM model (Davis, 1989) states clearly that intention to adopt a new technology are affected by its perceived ease of use and its perceived usefulness. There still exists a chasm between what academic marketing researchers find and what is practiced in businesses, perhaps because of the previous 2 factors not being met. Thus, researchers in marketing science must continue to devise methods that make their work accessible to practitioners. This includes not just accessible methods, but systems (e.g., software) that allow for easier and speedier application of more complex methods and increased academic-industry knowledge transfer.

One major future area of research is in the developments around supervised machine learning for text classification. Firstly, such methods could be used to create larger thematisized datasets, or classify new datasets from existing codes and themes. For example, once a sufficiently large dataset is coded by researchers, this then forms the basis to train a model from the language of these texts to classify new reviews that come through into existing themes.

Methods such as naive bayes (Kim et al., 2006; Kibriya et al., 2005; Frank and Bouckaert, 2006) and deep learning (Kim, 2014; Sutskever et al., 2014) have been shown to be extremely accurate when classifying text into categories from pre-trained data. This would retain the benefit of a human researcher first developing the themes, and thus the nuance of human understanding, and the machine capabilities of taking this understading to a wider set of texts than is reasonable for a human to be read.

Another promising area of future research could look into the application of the current method in marketing practice. This should be UX-led and include the design of a system that is accessible from both qualitative and quantitative perspectives, ensuring that marketing professionals can easily integrate these insights into their strategies without needing advanced methodological expertise.

This paper sought to overcome some of the limitations of either using a data science approach to understanding brand positioning or the challenges faced from solely using a qualitative methods. In doing so, it has laid the potential pathway for methods of discovering emergent themes from reviews of products to positioning them on a bi-plot with the use of CA. It presented clear and easy methods to enhance marketing strategy by harnessing an understanding how brands are benchmarked against their competitors and on what thematic dimensions.

8. Reflections

This dissertation journey has been a great learning experience, expanding my research capabilities, particularly, the potential now to integrate mixed-methods into my professional life. Entering this project, I possessed a solid foundation in quantitative marketing methodologies with limited experience in qualitative research, synthesizing these two approaches within a single research framework was hugely beneficial. What follows is a brief reflection using Gibbs (1988) on how I overcame a key challenge.

8.1 Key Challenge

The study aimed to explore brand positioning using a mixed-methods approach, combining thematic and CA. However, as I was reading and coding the reviews, the large volume of data quickly revealed the method's impracticality - one of the goals was to give me better tooling for my professional life. What initially seemed a thorough and realistic approach became time-consuming and cumbersome. This was only the qualitative portion of the mixed-methods; the quantitative analysis was still pending, further complicating the process.

8.2 Feelings

Feelings At this stage, I was concerned about the feasibility of the study within the given time frame and resources. The task of manually coding and analysing such a large dataset was overwhelming, and a main purpose, applicability, was being comprimised. This led to a sense of frustration, as I had high hopes for the study to not only be academically rigorous but also practically valuable for industry professionals. While a null result is not a failure, I still wanted to find a solution for this issue.

8.3 Evaluating

I recognized that the original methodology, while theoretically sound, needed a practical adjustment to maintain the study's relevance and feasibility. The extensive time required for thematic analysis raised questions about the volume of data needed for data saturation and whether the themes identified would genuinely reflect the overall dataset. I took some time to reflect on this and come up with a solution. Remembering what I had read about focus and diffuse mode of thinking in tackling mathematics (Oakley, 2014) I stepped away to focus and distract myself. It was only in taking some time away that I was able to discover a new pragmatic solution.

8.4 Analysis

I believe that it was only because I allowed this reflection time that I considered the inclusion of a resampling method to address these concerns. Resampling, specifically a bootstrapping approach, provided a statistical basis for determining of how much data is needed on future occasions. I figured that, if data saturation only occured at high volumes of data, applicability (not the usefulness of the primary results) would have been at risk. A null result would not have been a bad thing, but I needed to test this to see.

8.5 Final Thoughts

The reflexive decision to include a resampling approach was a challenge as this added a further analytical portion. This took a long time to conduct, a large volume of code, and increased the pressure on the word count. A word count that was already under stress because of the mixed-methods nature. In the end it worked, but more importantly, I only came up with this solution because I was willing to take time away from the problem and even go back through some old machine learning books on my book shelf.

8.6 Action Plan

I will continue to employ reflective practices, such as Gibbs' Cycle, in my research endeavors and time away from challanges for creativity in my solution to issues. This experience has underscored the value of being adaptable and responsive to the challenges that arise during a research project. In future studies, I plan to integrate reflective cycles more explicitly from the outset, ensuring that any necessary methodological adjustments are guided by a structured and critical reflection process.

References

Adjei, M.T., Zhang, N., Bagherzadeh, R., Farhang, M. and Bhattarai, A., 2023. Enhancing consumer online reviews: the role of moral identity. *Journal of Research in Interactive Marketing*, *17*(1), pp.110-125.

Alpert, L. and Gatty, R., 1969. Product positioning by behavioral life-styles. *Journal of marketing*, *33*(2), pp.65-69.

Alzate, M., Arce-Urriza, M. and Cebollada, J., 2022. Mining the text of online consumer reviews to analyze brand image and brand positioning. *Journal of Retailing and Consumer Services*, *67*, p.102989.

American Psychological Association (2017) Ethical Principles of Psychologists and Code of Conduct.

Ankers, P. and Brennan, R., 2002. Managerial relevance in academic research: an exploratory study. *Marketing Intelligence & Planning*, *20*(1), pp.15-21.

Aurier, P. and Guintcheva, G., 2014. Using affect–expectations theory to explain the direction of the impacts of experiential emotions on satisfaction. *Psychology & Marketing*, *31*(10), pp.900-913.

Bell, S.S., Holbrook, M.B. and Solomon, M.R., 1991. Combining esthetic and social value to explain preferences for product styles with the incorporation of personality and ensemble effects. *Journal of social behavior and personality*, *6*(6), p.243.

Bhat, S. and Reddy, S.K., 1998. Symbolic and functional positioning of brands. *Journal of consumer marketing*, *15*(1), pp.32-43.

Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), pp.77-101.

Buker, T., Schmitt, T., Miehling, J. and Wartzack, S., 2022. What's more important for product design–usability or emotionality? An examination of influencing factors. *Journal of Engineering Design*, *33*(8-9), pp.635-669.

Chan, H. and Yang, M.X., 2021. Culture and electronic word of mouth: a synthesis of findings and an agenda for research. *Journal of Global Marketing*, *34*(3), pp.165-169.

Charmaz, K., 2008. Grounded theory as an emergent method. Handbook of Emergent Methods, 155-172.

Côté, L. and Turgeon, J., 2005. Appraising qualitative research articles in medicine and medical education. *Medical teacher*, *27*(1), pp.71-75.

Crawford, C.M., 1985. A new positioning typology. *Journal of Product Innovation Management: An International Publication of the Product Development & Management Association*, *2*(4), pp.243-253.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pp.319-340.

Dressler, M. and Paunovic, I., 2021. Reaching for customer centricity—wine brand positioning configurations. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(2), p.139.

Dwivedi, Y.K., Slade, E. and Williams, M.D., 2017. Electronic Word of Mouth in the Marketing Context: A State Of The Art Analysis And Future Directions. *Switzerland: Springer Nature*

Frank, E. and Bouckaert, R.R., 2006. Naive bayes for text classification with unbalanced classes. In *Knowledge Discovery in Databases: PKDD 2006: 10th European Conference on Principles and Practice of Knowledge Discovery in Databases Berlin, Germany, September 18-22, 2006 Proceedings 10* (pp. 503-510). Springer Berlin Heidelberg.

Fujimoto, T., 2007. Architecture-based comparative advantage—a design information view of manufacturing. *Evolutionary and Institutional Economics Review*, *4*, pp.55-112.

Fujimoto, T. and Miller, B. (2007) *Competing to be really, really good: The behind-the-scenes drama of capability-building competition in the automobile industry*. Vol. 22. Tokyo: International House of Japan.

Gibbs, G. (1988) *Learning by doing: A guide to teaching and learning methods*. Oxford Polytechnic, Oxford: Further Education Unit.

Greenacre, M. (2017) *Correspondence analysis in practice*. chapman; hall/crc.

Greene, J.C., 2008. Is mixed methods social inquiry a distinctive methodology?. *Journal of mixed methods research*, *2*(1), pp.7-22.

Grewal, D., Hulland, J., Kopalle, P.K. and Karahanna, E., 2020. The future of technology and marketing: A multidisciplinary perspective. *Journal of the Academy of Marketing Science*, *48*, pp.1-8.

Hacking, C., Verbeek, H., Hamers, J.P.H. and Aarts, S. (2023) Comparing text mining and manual coding methods: Analysing interview data on quality of care in long-term care for older adults B. Gobin (ed.). *PLOS ONE*, 18(11) e0292578.

Harrison, R.L. and Reilly, T.M. (2011) Mixed methods designs in marketing research. *Qualitative market research: an international journal*, 14(1) 7–26.

Hjellbrekke, J. (2018) *Multiple correspondence analysis for the social sciences*. Routledge.

Hu, X. and Ha, L. (2015) Which form of word-of-mouth is more important to online shoppers? A comparative study of WOM use between general population and college students. *Journal of Communication and Media Research*, 7(2) 15–35.

Ismagilova, E., Slade, E.L., Rana, N.P. and Dwivedi, Y.K. (2020) The effect of electronic word of mouth communications on intention to buy: A meta-analysis. *Information Systems Frontiers*, 22 1203–1226.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2021) An introduction to statistical learning: With applications in r. Springer texts in statistics. 2nd edition. Springer New York, NY.

Jiang, G., Liu, F., Liu, W., Liu, S., Chen, Y. and Xu, D. (2021) Effects of information quality on information adoption on social media review platforms: Moderating role of perceived risk. *Data Science and Management*, 1(1) 13–22.

Kahneman, D. (2017) *Thinking, fast and slow*.

Kanouse, D.E., 1984. Explaining Negativity Biases in Evaluation and Choice Behavior: Theory and Research. *Advances in consumer research*, *11*(1).

Karadağ, H., Tosun, P. and Ayan, B. (2022) User-generated and brand-generated content as indicators of university brand personality and business strategy. *Journal of Marketing for Higher Education*, 1–23.

Kibriya, A.M., Frank, E., Pfahringer, B. and Holmes, G. (2005) Multinomial naive bayes for text categorization revisited. In: *AI 2004: Advances in artificial intelligence: 17th australian joint conference on artificial intelligence, cairns, australia, december 4-6, 2004. Proceedings 17*. 2005 Springer, 488–499.

Kim, S.-B., Han, K.-S., Rim, H.-C. and Myaeng, S.H. (2006) Some effective techniques for naive bayes text classification. *IEEE Transactions on Knowledge and Data Engineering*, 18(11) 1457–1466.

Kim, Y. (2014) *Convolutional neural networks for sentence classification*. In: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 2014 Association for Computational Linguistics,.

King, W.R. and He, J. (2006) A meta-analysis of the technology acceptance model. *Information & management*, 43(6) 740–755.

Kotler, P., Brady, M., Goodman, M. and Hansen, T. (2019) *Marketing management: European edition*. Pearson Higher Ed.

Kutabish, S., Soares, A.M. and Casais, B. (2023) *The influence of online ratings and reviews in consumer buying behavior: A systematic literature review*. In: *ICDEc 2023, LNBIP 485*. 2023 Braga, Portugal: School of Economics; Management, University of Minho; Springer Nature Switzerland AG,.

Le, Q. and Mikolov, T. (2014) Distributed representations of sentences and documents. In: *Proceedings of the 31st international conference on machine learning (ICML)*. JMLR: Workshop and conference proceedings. 2014 Beijing, China: JMLR.org; Journal of Machine Learning Research, 1188–1196.

Lee, J., Park, D.-H. and Han, I. (2008) The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic commerce research and applications*, 7(3) 341–352.

Leeflang, P.S.H., Verhoef, P.C., Dahlström, P. and Freundt, T. (2014) Challenges and solutions for marketing in a digital era. *European Management Journal*, 32 1–12.

Lilien, G.L. (2011) Bridging the academic-practitioner divide in marketing decision models. *Journal of Marketing*, 75(4) 196–210.

Liu, X., Lin, J., Jiang, X., Chang, T. and Lin, H. (2024) eWOM information richness and online user review behavior: Evidence from TripAdvisor. *Journal of Theoretical and Applied Electronic Commerce Research*, 19(2) 880–898.

Machová, K., Szabóova, M., Paralič, J. and Mičko, J. (2023) Detection of emotion by text analysis using machine learning. *Frontiers in Psychology*, 14.

Malter, M.S., Holbrook, M.B., Kahn, B.E., Parker, J.R. and Lehmann, D.R. (2020) *The past, present, and future of consumer research*. *Springer Science Business Media, LLC, part of Springer Nature*,.

Market Research Society (2023) MRS Code of Conduct 2023.

Maslowska, E., Malthouse, E.C. and Viswanathan, V. (2017) Do customer reviews drive purchase decisions? The moderating roles of review exposure and price. *Decision Support Systems*, 98 1–9.

McAuley Lab (2023) Amazon reviews 2023. Hugging Face.

Miles, M.B. and Huberman, A.M. (1994) *Qualitative data analysis: An expanded sourcebook*. sage.

Molina-Azorin, J.F. (2016) *Mixed methods research: An opportunity to improve our studies and our research skills*.

Morgan, D.L. (2007) Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1) 48–76.

Nenadic, O. and Greenacre, M. (2007) Correspondence analysis in r, with two- and three-dimensional graphics: The ca package. *Journal of Statistical Software*, 20(3) 1–13. Available from *http://www.jstatsoft.org*.

Noble, C.H. and Kumar, M. (2008) Using product design strategically to create deeper consumer connections. *Business Horizons*, 51(5) 441–450.

Nowell, L.S., Norris, J.M., White, D.E. and Moules, N.J. (2017) Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1) 160940691773384.

Oakley, B. (2014) *A mind for numbers: How to excel at math and science (even if you flunked algebra)*. Penguin.

Palmer, A. (2013) *EBOOK: Principles of services marketing*. McGraw Hill.

Park, C.W., MacInnis, D.J. and Eisingerich, A.B. (2016) *Brand admiration: Building a business people love*. John Wiley & Sons.

Pieters, R.G.M. and Bijmolt, T.H.A. (1997) Consumer memory for television advertising: A field study of duration, serial position, and competition effects. *Journal of Consumer Research*, 23(4) 362–372.

Prihananto, P., Yusvianty, R.S.A., Hakim, N.S., Bhawika, G.W. and Agustin, H. (2024) Analyzing brand positioning and brand image of smartphone brands in indonesia by mining online review. *Procedia Computer Science*, 234 318–332.

Qualtrics (2023) Qualtrics XM/os2.

R Core Team (2024) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Riebe, E. and Dawes, J. (2006) Recall of radio advertising in low and high advertising clutter formats. *International Journal of Advertising*, 25(1) 71–86.

Rutter, R., Lettice, F. and Nadeau, J. (2016) Brand personality in higher education: Anthropomorphized university marketing communications. *Journal of Marketing for Higher Education*, 27(1) 19–39.

Saldaña, J. (2021) *The coding manual for qualitative researchers*.

Saqib, N. (2020) Positioning–a literature review. *PSU Research Review*, 5(2) 141–169.

Simms, C. and Trott, P. (2007) An analysis of the repositioning of the 'BMW mini' brand. *Journal of Product & Brand Management*, 16(5) 297–309.

Sun, X., Cao, X. and Lau, Y. (2023) Exploring cruise tourists' sentiment expression pattern from online reviews: An insight into market positioning. *Tourism Management Perspectives*, 49 101195.

Sutskever, I., Vinyals, O. and Le, Q.V. (2014) Sequence to sequence learning with neural networks. In: *Proceedings of the 27th international conference on neural information processing systems (NIPS)*. 2014 Curran Associates Inc., 3104–3112.

Taecharungroj, V. (2022) Experiential brand positioning: Developing positioning strategies for beach destinations using online reviews. *Journal of Vacation Marketing*, 29(3) 313–330.

Tao, H., Sun, X., Liu, X., Tian, J. and Zhang, D. (2022) The impact of consumer purchase behavior changes on the business model design of consumer services companies over the course of COVID-19. *Frontiers in Psychology*, 13 818845.

The British Psychological Society (2021) Code of ethics and conduct. The British Psychological Society.

Toufani, S., Stanton, J.P. and Chikweche, T. (2017) The importance of aesthetics on customers' intentions to purchase smartphones. *Marketing Intelligence & Planning*, 35(3) 316–338.

Trout, J. and Ries, A. (1986) *Positioning: The battle for your mind*. McGraw-Hill New York, NY.

Turner, M., Kitchenham, B., Brereton, P., Charters, S. and Budgen, D. (2010) Does the technology acceptance model predict actual use? A systematic literature review. *Information and software technology*, 52(5) 463–479.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003) User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425–478.

Webb, P.H. and Ray, M. (1979) Effects of television clutter. *Journal of Advertising Research*, 19 7–12.

White, K., Hardisty, D.J. and Habib, R. (2019) The elusive green consumer: People say they want sustainable products, but they don't tend to buy them. Here's how to change that. *Harvard Business Review*, 97(4) 124–133.

Wright, M. (1997) When strength means death. *Brand Strategy*, 107 3–5.