Using Structural Topic Modelling to predict users' sentiment towards Intelligent Personal Agents. An application for Amazon's Echo and Google Home.

Authors:

Manuel J. Sánchez-Franco (*), Department of Business Administration and Marketing, University of Sevilla, Spain.
Francisco J. Arenas-Márquez, Department of Finance and Operations Management, University of Sevilla, Spain.
Manuel Alonso-Dos-Santos, Department of Administration, Universidad Católica de la Santísima Concepción, Chile / Department of Marketing, University of Granada, Spain

(*) Corresponding author.

Abstract:

Despite growing levels of usage of Intelligent Personal Assistants (hereinafter, IPA), their in-home usage has not been studied in depth by scholars. To increase our understanding of user interactions with IPA, our research created a theoretical framework rooted in technology acceptance models and Uses and Gratification Theory. Our empirical method designs an ambitious analysis of natural and non-structured narratives (user-generated content) on Amazon's Echo and Google Home. And to identify key aspects that differentially influence the evaluation of IPA our method employs machine-learning algorithms based on text summarisation, structural topic modelling and cluster analysis, sentiment analysis, and XGBoost regression, among other approaches. Our results reveal that (hedonic and utilitarian) benefits gratification, social influence and facilitating conditions have a direct impact on the users' sentiment for IPA. To sum up, designers and managers should recognise the challenge of increasing the customer satisfaction of current and potential users by adjusting doubtful users' technical skills and the (hedonic, cognitive, and social) benefits and functionalities of IPA to avoid boredom after a short lapse of time. Finally, the discussion section outlines future lines of research and theoretical and managerial implications.

Type: Research article

Keywords: *Intelligent Personal Assistants; technology acceptance models; Uses and Gratification Theory; text analytics; sentiment analysis; Structural Topic Model, XGBoost regression.*

1. Introduction:

Human-computer interaction enables making domestic life easier and adapted to the preferences of users who are seeking not only to multitask through voice-control platforms (Edwards *et al.*, 2019) but also to stimulate companionship (Lopatovska &Williams, 2018). "Conversation is [thus] becoming an essential mode of human-computer interaction" (Luger & Sellen, 2016, p.5286). In this regard, ambient intelligence research (hereinafter, AmI; *cf.* Vasilakos & Pedrycz, 2006) is increasingly focusing on natural and non-structured (and more efficient) *lifelike conversational* interactions between humans (*e.g.*, families) and machines (Hoy, 2018; McLean & Osei-Frimpong, 2019a; *cf.* also Alepis & Patsakis, 2017).

AmI research on smart homes (Marikyan *et al.*, 2019) and their related technologies (Shin *et al.*, 2018) such as Intelligent Personal Assistants (hereinafter, IPA), has recently been receiving recent and increasing scholarly interest. However, research on the voluntary usage of in-home devices is still necessary to assess how to promote increased levels of non-economic satisfaction based on cumulative gratifying and easy interactions (Chiang *et al.*, 2020; Eyssel *et al.*, 2012). "The interaction experience is even more important to the users than its outcome" (Lopatovska & Williams, 2018, p.267). Our study, therefore, stems from the growing interest in knowledge representation that allows the systems (1) to convert and store what they know and hear, and consequently, (2) to increase personification and facilitate interactions (*cf.* self-service technology, *e.g.*, Belanche *et al.*, 2020; Nijssen *et al.*, 2016; Pantano & Pizzi, 2020).

To understand the adoption and voluntary usage of IPA, our research is based on technology acceptance models (*cf.*, Davis, 1989, Davis et al, 1992, Venkatesh *et al.*, 2003, 2012, among others), and comprehensively integrates the Uses and Gratification Theory (hereinafter, U&GT, *cf.* Katz *et al.*, 1974). On the one hand, there is a growing body of theoretical and empirical evidence related to the relevance of extrinsic factors (such as perceived usefulness, ease-of-use or facilitating conditions) for adopting and using different technologies in many contexts. For instance, TAM is a parsimonious, valid and robust model that has been widely used. On the other hand, individuals are active in integrating media into their own lives (Gan & Li, 2018), and select an innovation (*e.g.*, in-home voice assistants) because it satisfies a range of needs (*e.g.*, enjoyment or pleasure, among others). In this regard, previous research became aware of the relevance of additional motivations of use, such as intrinsically enjoyable experiences or social motivations, in understanding technology adoption and usage behaviours (*cf.* Holsapple & Wu, 2007; Gaisser & Utz, 2020; Sánchez-Franco & Roldán, 2005; McLean & Osei-Frimpong, 2019a; Ryan & Deci, 2000; Venkatesh *et al.*, 2012). And it is precisely U&GT (Katz *et al.*, 1974, among others) to allow a proper understanding of individuals' additional motivations to fulfil unsatisfied needs, as well as IPA adoption and use. As Joo and Sang (2013, p.2513) propose, "the combination overcomes the limitations of TAM while encompassing users' motivations yet retaining the parsimonious feature of TAM".

Moreover, although previous studies analyse the adoption and use of IPA through self-reported surveys (Lovato & Piper, 2015; Pradhan *et al.*, 2018), our research here develops a feature-oriented product approach based on a bulk of customers' natural and non-structured user-generated content (hereinafter, UGC) extracted from Amazon.com, Amazon.co.uk and Bestbuy.com sites (hereinafter, Amazon and Bestbuy sites) and centred on Amazon's Echo and Google Home devices. These are conceptualised as "a software agent that provides professional/administrative, technical and social assistance to human users by automating and easing many day-to-day activities" (*cf.* Han & Yang, 2018). To understand the richness of latent content related to motives towards using Amazon's Echo and Google Home devices, our research applies a novel application of topic modelling (here STM) and hierarchical clustering in the in-home assistant domain. Additionally, to assess how a topic predicts the acceptance and user experience (extracted from a sentiment analysis) when using IPA, our analysis employs XGBoost regressors, SHAP values and a one-sided hypothesis test based on Bayesian estimation. In our opinion, our results -on how the IPA improves their home automation- are more reliable and accurate than previous statistical results based on limited sample data, and provide novel clues not explicitly indicated by customers' ratings.

To sum up, our research extends previous research questions focused on predicting topics' influences on the adoption and usage of in-home devices by users. To our knowledge there is a paucity of research on (1) why users are adopting and accepting in-home voice assistants and (2) how they can be better designed to fit, speed up, or improve habits and routines while avoiding unnecessary tasks (*e.g.*, Bentley *et al.*, 2018; de Barcelos Silva *et al.*, 2020; Livaditi *et al.*, 2003; McLean & Osei-Frimpong, 2019a; Mennincken & Huang, 2012; Ramírez-Correa *et al.*, 2019). The following research questions are then important:

> RQ1. What are the motivations (here, topics) of users related to their encounters with IPA and highlighted in natural and non-structured user-generated content?
>
> RQ2. How valid are the extracted topics by analysing an additional underlying semantic structure identified through a meaningful hierarchical clusterisation?
>
> RQ3. What are the different levels of sentiments (associated with cumulative gratification and easy interactions) when users choose Amazon's Echo or Google Home?
>
> RQ4. What are the main topics that predict users' sentiments and the extent to which the use of IPA could become a routinised activity or habit?

This paper theoretically addresses these questions by focusing on the audience's usage motives associated with users' expectations, predictions, goals and desires. The research method section provides details of the data gathering, data mining, and findings on Amazon's Echo and Google Home. Figure 1 outlines the steps for transforming free-form text into a structured form, and main research questions. The discussion section presents future research lines and theoretical and managerial implications.

::Figure 1 near here::

2. Theoretical framework

Lopatovska and Williams (2018, p.267) note that "the most frequent types of [Alexa] interactions included quick weather checks and music-playing requests" (see also Lopatovska *et al.*, 2018). Bentley *et al.* (2018) conclude that the median household performs 4.1 commands per day, and the most frequent requests are, for instance, for music (40 %) and information (17 %) -see also Hoy (2018), and McLean and Osei-Frimpong (2019b). Our research, therefore, focuses on extrinsic motives (related to IPA, such as perceived usefulness, ease of use or facilitating conditions) that influence individuals to act in a certain way to properly fulfil a need or desire (*e.g.*, education assistance or medical assistance, among others) or enhancing daily uses (*e.g.*, controlling lights, air-conditioning systems, or setting timers, security access control of users' home and family). On the other hand, IPA foster intrinsic (or ritualised) motivations such as enjoyment through listening to music or symbolic or social motivations to create or maintain social relationships through an emotional connection or companionship from nonhuman agents.

Accordingly, our research assumes a user-level insight in understanding the voluntary usage of in-home devices. And it is grounded on technology acceptance models, such as TAM -widely used to identify the drivers of technology acceptance (Davis, 1989; Davis *et al.*, 1992)-, TAM2 (Venkatesh & Davis, 2000), TAM3 (Venkatesh & Bala, 2008), UTAUT (Venkatesh *et al.*, 2003) or UTAUT2 (Venkatesh *et al.*, 2012), among others, and additionally, on U&GT. "The U&G approach allows us to understand intrinsic factors of individual users, a weak point in TAM" (Yoo & Sang, 2013, p.2517). In other words, by incorporating the works of Katz *et al.* (1974), Weibull (1985) and other research such as that of Rauschnabel *et al.* (2018) and McLean and Osei-Frimpong (2019a), our extending theoretical background enables jointly identifying the main drivers that motivate the users' psychological assessment of IPA such as utilitarian-, hedonic- or symbolic-benefits gratifications.

Our study, therefore, outlines several types of motivations –referring to the intended or expected gains from use:

- Utilitarian (or instrumental) benefits gratifications are related to cognitive needs –also called performance expectancy by Venkatesh *et al.* (2012). They are (1) defined as "the degree to which a person believes that using a particular system would enhance his or her performance" (Davis, 1989, p.320), and (2) associated with the subjectively assessed practicality of IPA based on goal-oriented drivers. IPA provide information and services to reduce the effort and complexity of task accomplishments (*e.g.*, checking weather) and enable users to be hands-free during activities like:
    a. Searching for real-time information to complete a task.
    b. Learning based on the need to know (*e.g.*, looking up recipes or creating notes related to content gratification).
    c. Looking up customer services.
- Perceived ease of use is associated here with IPA usability. In particular, the IPA's malleability or customisability lead to personification and facilitate interactions to make domestic life less effortful and, consequently, more enjoyable (Sánchez-Franco & Roldán, 2005). Devices that are easier to use (1) provide better feedback to users' stimuli (*e.g.*., through their impact on the user's perception of self-efficacy) and, therefore, (2) lead to increased emotional responses of pleasure.
- Facilitating conditions (*e.g.*, Bluetooth capabilities or easy subscriptions) are an additional main driver. Facilitating conditions here are defined as "perceived enablers or barriers in the environment that influence a person's perception of ease or difficulty of performing a task" (Teo, 2010, p.255).
- Hedonic benefits gratification is precisely related to tension-release needs (Rubin, 2002), and could be defined here as "the extent to which the activity of using the computer is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated" (Davis *et al.*, 1992, p.1113; *cf.* also Venkatesh & Davis, 2000). Precisely, "hedonic gratifications are associated with various positive outcomes such as the reduction of boredom or pleasure" (Rauschnabel, 2018, p.218). For instance, individuals may use IPA to seek entertainment associated with a particular task (or also fun, fantasy or escapism), social gaming or enjoyment and, consequently, a reduction of boredom through listening to music, playing games, listening to audiobooks, telling jokes, among others.
- Social gratification is related to individuals who use key media for social needs, such as creating or maintaining social relationships (*e.g.*, McLean & Osei-Frimpong, 2019a). This is conceptualised as "an investment whereby the parties commit resources to each other to receive benefits in the future" (Camén *et al.*, 2011, p.378).
- Symbolic benefits gratification is related to the desire for the service to satisfy internal needs, such as self-enhancement or group membership (Park *et al.*, 1986). Individuals may use specific innovations (*e.g.*, IPA) to reaffirm their social status (and to gain a symbolic reward such as making a favourable impression on others, see Goodin, 1977). Symbolic benefits gratification is indeed derived from socialisation processes and social interactions (Wee & Ming, 2003).

To sum up, by combining TAM and U&GT frameworks, our research comprehensively explains and predicts IPA adoption by users and details motivations that can affect the user acceptance of own-brand devices such as Amazon's Echo or Google Home. Both smart speakers allow the intelligence of the agent to be accessible (Alexa or Google Assistant) and this leads to redesigning the daily self-service assistance to enhance their adoption and acceptance as well as the short- and long-term emotional state of users.

3. Materials and method:

To identify key aspects that influence the evaluation of IPA, our research focuses on machine-learning algorithms based on text summarisation, structural topic models, hierarchical clustering, sentiment analysis, and XGBoost regression, among other approaches. Our study aim lies in improving the performance of mining tasks associated with customer reviews through:

- Examining the underlying semantic structure by applying topic modelling.

- Combining the number of identified topics into *handy* clusters to interpret them semantically more easily.
- Estimating a topics' relationship to individuals' sentiment towards IPA.
- Predicting the development of mutually beneficial relationships with users.

3.1. Data collection:

Our research gathers data on Amazon's Echo and Google Home stored at the Amazon and Bestbuy sites. In 2019, the top two IPA were Amazon's Echo and Google Home. According to eMarketer (2018), in the US 66.6 % of individuals -who use at least one smart speaker at least once per month- mention Amazon, followed by Google (29.5 %). According to the statistics corporation Statista (2020), Alexa and Google alone gather 63 % of the global market share.

Our dataset contains 31,026 reviews of different configurations of (1) the Amazon Echo, Echo Dot and Echo Plus (20,378) devices and (2) Google Home and Google Home Mini (10,648), spanning the period between 2017 and 2019. See Figure 2.

::Figure 2 near here::

3.2. Data cleansing process:

Our research: (1) discards punctuation, capitalisation, digits, and extra whitespaces, (2) checks correct spelling, (3) tokenises and lemmatises the terms, and (4) removes a list of common stop words to filter out overly common terms. To keep the language variable consistent across texts, it eliminates non-English reviews (and characters). It also omits terms shorter than a minimum of three characters (Sánchez-Franco *et al.*, 2020).

3.3. Extracting terms

Our research selects a dictionary based on Term Frequency–Inverse Document Frequency (TF-IDF) values above the median. TF reflects how frequently a term *t* occurs in document *d*. IDF measures if the term *t* is rare or common across the corpus. IDF thus allows to filter out common terms, *g*, a term *t* reaches a high TF-IDF value through a high term frequency (in document *d*) and a low document frequency of the term in the corpus. This selecting process finally yields a handy dictionary of 1,256 terms.

4. Data mining:

4.1. Research question 1. Structural topic model:

Our research estimates the relationships between terms and documents through a text-mining algorithm to discover hidden semantic structures in the IPA users' narratives (*e.g.*, facilitating conditions, performance expectancy, effort expectancy, portability, among others). Our approach focuses on the Structural Topic Model as a generative model of term counts (hereinafter, STM; *cf.* Roberts *et al.*, 2018, 2020). The STM algorithm enables the researchers to discover topics that can be correlated and estimate their relationships to document metadata. "Researchers can [thus] change the topic prevalence parameters by adding covariates [here, IPA models] to affect the topic proportion of the document" (He *et al.*, 2020, p.7).

Assuming that there is no single correct path to select an optimal number of topics, our research estimates different STM models among 10 and 60 topics and assesses the trade-off between exclusivity and semantic coherence (Blei *et al.*, 2003; Gerring, 2001, Mimno *et al.*, 2011; Roberts *et al.*, 2014). Although terms that are most likely under a topic should co-occur within the same document (semantic coherence metric), they should additionally be exclusive to the topic *t* (exclusivity metric) for communicating content (Bischof & Airoldi, 2012). Here 24 topics are an acceptable (not optimal) number of topics for uncovering useful features about the natural content of the users' reviews concerning the IPA experiences.

Figure 3a displays estimated topic proportions (*i.e.*, expected proportion of the corpus that belongs to each topic), as well as Gini coefficients. In this regard, the highest prevalent topics correspond closely to general topics (topics 3, 13 and 12, in this order), and are associated with wearable comfort and physical features to create an enjoyable (and ubiquitous) environment, *i.e.*, to ensure fun and pleasure deriving from the conveniences (see Figure 4). Moreover, the Gini coefficient is used to measure the inequality of each topic across reviews and varies between 0 (complete equality) and 1 (complete inequality). Topics 3, 17 and 14 are extremely concentrated; on the other hand, topics 9, 22 and 24 are discussed with analogue attention by our sample (that is to say, extremely spread). Figure 3b shows the comparison of topic composition between Amazon's Echo and Google Home based on estimated topic proportions.

::Figures 3ab near here::

According to usefulness for interpreting topics, in Figure 4 topics are conceptualised with a list of the most representative terms ordered from highest to lowest relevance (here, top 10 terms, and $\lambda = 0.6$; *cf.* Sievert & Shirley, 2014). Figure 4 also highlights the terms whose saliency coefficients are higher than their median. Highlighting terms by saliency (and distinctiveness) helps the classification and disambiguation of topics. Moreover, the distinctiveness of *w1* measures "how informative the specific term *w1* is for determining the generating topic, versus a randomly-selected term *w2*" (Chuang *et al.*, 2012), *i.e.*, how much less the term *w1* is shared across topics. In turn, the saliency of a terms *w1* is the result of multiplying the probability of term *w1* by its distinctiveness (*cf.* also Sievert & Shirley, 2014), *i.e.*, saliency penalises the terms shared across several topics and fosters terms that are acceptable predictors of topics. Relating the values of saliency and distinctiveness of most relevant terms, our results concur with the findings of Lopatovska and Williams (2018, p.267). See Figure 5. Overall, users are concerned either with the core functions of the IPA (*e.g.*, music or questions) or customer service (*e.g.*, setup or purchase). Users mention utilising IPA to make calendar appointments and create reminders, to find local restaurants, or to answer trivia questions, among others. Users also employ IPA as an alarm or for general internet searches, while playing music was the primary reason for using IPA.

::Figure 4 near here::

::Figure 5 near here::

4.2. Research question 2. Hierarchical clustering:

To validate how valuable scholars and managers find the extracted topics as subject matter, our research here focuses on the model's semantic validity, *i.e.*, "the extent to which each category or document has a coherent meaning and the extent to which the categories are related to one another in a meaningful way" (Quinn *et al.*, 2010, p.216). In this regard, one way to define topics here is to categorise them by meaningful groupings of topics revealing their organisation by applying a clustering analysis. Representative topics in each grouping should form a meaningful and consistent cluster.

Hierarchical clustering is applied for its simplicity and ease of understanding. The main steps are outlined below:

- Our research compares an agglomeration hierarchical clustering (Ward's method) with a divisive hierarchical method. The analysis is implemented by using the logarithmised theta matrix (representing the distribution of topics over documents) as a distance matrix input.
- Although there is not an optimal number of groupings, both dendrograms intuitively suggest an acceptable (and handy) number of 5 clusters. Regarding this, Figure 6 provides an easy-to-interpret view of its agglomeration clustering structure based on a dendrogram that illustrates the relationships between clusters.

- k being= 5, the agglomerative coefficient of the agglomeration hierarchical method is also closer to one ($ac_{ward} = 0.671 > ac_{complete} > ac_{average} > ac_{single}$) than the coefficient associated with the divisive hierarchical method (dc = 0.596) [1].
- Our study selects k = 5 and the agglomeration hierarchical method.

::Figure 6 near here::

Next, our research proceeds to interpret each of the groupings, focusing our interpretation on the dendrogram's bottom level.

#1a: Facilitating conditions.

- Topics 10 and 16 describe (poor) facilitating conditions (even in the presence of overall effort expectancy) related to, for instance, setting up the IPA (*e.g.*, topic 10: "*I liked it but had a little trouble getting it setup. Once I found out how to install the application, I had no problems. Cannot wait to learn more about how to use it* (…)"). Negative interactions associated with the voice-activation could become common barriers to impede the continued use of IPA.

#1b: Integration between devices, and language compatibilities.

- Topic 5 refers to (poor) integration between smart devices (*e.g.*, control home lights, door locks or temperatures) or services (*e.g.*, subscriptions), as an incentive to purchase (affecting limited performance expectancy). Indeed, a main source of frustration lies in the lack of integration with other devices or service subscription plans as suggested in topic 5: "*It was not made obvious that you have to sign-up to Amazon Prime to get access to most of the music you want. You get signed up to the trial of Amazon Prime when you buy the Amazon Echo by default. I cancelled and lost all prime access which I did not realise I needed. When I renewed, I could not continue with the trial*".
- Topics 4 and 20 refer to (poor) feeling understood based on experienced conditions related to, for instance, problems understanding language accents. Language incompatibilities or poor integration services provided by IPA indeed promote scant value for their personal needs. For instance, topic 4 contains: "*A little sad Alexa is only programmed in English. I took her home to Holland and she just is not the way she was when I got to know her in the USA. Besides, because of my accent, she often does not understand what I am asking. I am pretty much disappointed with the Amazon device*"). Or topic 20: "*It's a cool device. The problem comes when you are playing something either on the device or another source such as a laptop. The voice recognition is severely affected*".

#2a: In-home control, affordability and simplicity.

- Topics 19 and 6 (home automation and compatibility related to switching costs), and 15 and 22 (hands-free human-device interaction) are related to in-home control that enhances affordability and simplicity (Yang & Lee, 2019). Following Shin *et al.* (2018, p.247), "compatibility is an important factor when choosing smart home services […] since smart home services require connections and communications among various home appliances". IPA help users thus become freer. For instance, topic 6 contains the following high topical content: "*I am a huge Google fan, but Google does not even come close to Alexa on home-automation, setup and compatibility. Google Home is better at answering questions*"). Or topic 15: "*Much better than Amazon Echo device. Sleek and aesthetic design, good audio, integrates well with a host of the smart home device. The only gripe is no aux-input and Bluetooth to speakers (unless you have bought a Chromecast device)*".

[1] The Adjusted Rand Index (hereinafter, ARI; *cf.* Hubert & Arabie, 1985), *i.e.*, a well-known measure of similarity between two partitions of a dataset, is here equal to 0.348 > 0 (RI = 0.779).

#2b: Price/value for money.

- Topics 2 and 21 are associated with price/value for money –defined as "consumers' trade-off between the perceived benefits of the applications and the monetary cost for using them" (Venkatesh *et al.*, 2012, p.161). For instance, "*This is well worth the money! I have never owned an Amazon device, but this is a very cost-effective solution to help automate your home. In my opinion, this is a better option than the normal Google Home as it has the same functionality at a fraction of the price!*" (Topic 21).

#3: Quality performance.

- Topics 11 and 24 are associated with quality performance (*e.g.*, topic 24: "*Only had this a couple of days so yet to find out how much it can do, or its shortcomings. But so far, I am impressed. It was very easy to set up using the application*").

#4a: Social influences, and hedonic motivations.

- Topic 12 is related to social influence. For instance, "*The device ease of use and convenience brought joy and excitement to the family. When family and friends visit our home, they enjoy the features and benefits. Excellent purchase*.". In this regard, Tsai *et al.* (2015) conclude that "older adults are more likely to use technology (…) if they are gifted the product by a family member" (*cf.* also Koon *et al.*, 2020).
- Topics 7 and 23 are (enjoyable and entertaining) benefits, relating to, on the one hand, motivating their adoption as a function of family, friends or society, and, on the other hand, the fun and pleasure derived from the IPA conveniences (*e.g.*, topic 7: "*The device ease of use and convenience brought joy and excitement to the family. When family and friends visit our home, they enjoy the features and benefits. Excellent purchase …*"). Or topic 23: "*My wife likes using it around the house it is like having an encyclopedia of knowledge when you ask a question it gives you the answers kind of like Siri does, we have fun using it*".

#4b: Effort expectancy.

- Topics 14 and 18, and topics 8 and 9 are related to (rich) effort expectancy (relating to performance expectancy or cognitive benefits gratification, such as seeking news and playing music, checking the weather and traffic, or learning) that enhances daily routines (*e.g.*, making life easier, even improving it by calendar management), *i.e.*, "the extent to which people tend to perform behaviours automatically because of learning" (Limayem *et al.*, 2007, p.709). For instance, topic 8 contains the following high topical content: "*Love the convenience of adding things to shopping list. Would appreciate the ability to put list items into 'groups' for easier reference when shopping*". Or topic 9: "*This device has delivered all that was promised. I look forward to exploring and expanding the capabilities of this useful purchase and hope that it continues to deliver. Sound is great, the ease in operation is also very satisfactory*". To sum up, topics 8 and 9 emphasise "the extent to which people tend to perform behaviours automatically because of learning" (Limayem *et al.* 2007, p.709)

#5: Novel design features.

- Topics 3 and 13 are associated with novel design features (wearable comfort influenced by IPA's physical features, *e.g.*, *smaller devices in each room*, and aesthetic properties such as the sound quality) that allow users to foster symbolic and aesthetic properties, to create an enjoyable environment, and to transport novel technologies (portability) which multiply the service access points of users (Yang *et al.* 2017). For instance, topic 3 contains the following high topical content: "*New design looks much better than the second generation. sound has improved as well as some bass now. New style visually appears to be of higher quality than previous versions*". Or topic 13: "*I put the charcoal mini in a bedroom with darker furniture. These are great additions, especially in large homes, to the larger portable unit. For very small rooms, where you do not need portability, these are perfect!*". To sum up, "in the same way individuals may furnish their home with designer hard and soft furnishings to elicit symbolic benefits, the in-home voice assistant may become part of this social enhancing activity" (McLean & Osei-Frimpong, 2019a, p.34).

4.3. Research questions 3. Sentiment analysis:

Our research analyses users' emotions as well as their positive or negative valence of a single (and robust) sentiment with IPA. IPA ratings (here, score or punctuation published by the site) tend to be extremely high and promote loss of informative value. The score distribution is indeed strongly skewed towards the 4- and 5-star ratings. See Table 1 and Figure 7. For this reason, our research applies an additional machine learning approach to extract the best features of the narratives according to users' sentiments regarding the IPA. Our study employs the NRC Word-Emotion Association Lexicon and its associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive) (*cf.* Mohammad & Turney, 2010, 2013). As Han *et al.* (2016, p.6) recommend, "these [dictionaries] are created to summarise the opinions within online customer reviews and to perform tonal analysis of social media blogs".

On the one hand, to find out whether reviewers (by IPA models) significantly employ more terms of emotions, our research applies a configural frequency analysis (hereinafter, CFA, *cf.* Bortz *et al.* 1990). CFA is a method for the analysis of bivariate or multivariate cross-classifications of categorical variables that detects patterns in our dataset that occur significantly more (types) or significantly less often (antitypes) than expected by chance. In this sense, Table 2 contains the Q coefficients and several additional local tests which reveal that all the configurations are significantly more frequent than expected by chance. In particular, in Figure 8 the Q coefficient (coefficient of pronouncedness) displays the size of the effect of each configuration and reveals if the reviewers of a model more frequently use positive (vs negative) terms compared to its competitor (*i.e.*, the larger the Q, the stronger the configuration):

- Compared to its opponent's IPA model, negative categories such as disgust, anger, fear and sadness, and a negative sentiment valence are significantly more employed in Amazon's Echo reviews.
- Concerning positive emotions, categories such as joy, anticipation, trust and surprise, there is no precise preference according to the IPA model analysed.
- In Google Home reviews positive emotions, such as joy and anticipation, and a positive sentiment valence are significantly more employed in comparison to trust and surprise. On the other hand, surprise and trust are preferred by users of Amazon's Echo.

On the other hand, the quality of the prediction (or research question 4, or epigraph 4.4) here is based on the levels of different sentiments associated with cumulative gratification and easy interactions (as the target variable of the regressor) when users choose Amazon's Echo or Google Home and the performance of an XGBoost regressor trained with the topics extracted. Related to our main research goals centred on a polarity sentiment for each review as a response variable, our study focuses on review-level sentiment based on the difference between the positive and negative valences for each review through the terms contained in them. For illustrative purposes only, the most frequent (lemmatised) terms related to positive valence are 'love', 'music', 'question', 'fun', 'recommend', 'happy' or 'gift', among others. The most negative influences are from 'problem', 'bite', 'alarm', 'disappoint', 'wait', 'frustrate', or 'annoy'. See Figure 9.

::Table 1 near here::

::Figure 7 near here::

::Figure 8 near here::

::Table 2 near here::

::Figure 9 near here::

4.4 Research question 4. Predictive analysis:

Our research derives an XGBoost regressor (Climent *et al.*, 2019; Chen & Guestrin, 2016; Sánchez-Franco *et al.*, 2020) trained with the topics extracted and uses the coefficients root mean squared error (RMSE) and r-squared (*r2*) scores to validate the research model. To find the best-fitted model based on the *r2* metric, our research splits the dataset into two subsets: 75% of the observations (train dataset) allow us to train the model. A 10-fold cross-validation helps with finding the optimal combination of parameters (Climent *et al.*, 2019). The remaining 25% (validation dataset) enable us to validate the performance of the *best-fitted model* and to ensure that the model does not fail to reliably fit additional data or predict future observations (over-fitting). Following the tuning process, the *r2* value *obtained* with the validation dataset (24 topics) is 0.96 (and has an RMSE of 0.403) with the use of early stopping (30) to avoid over-fitting, 100 rounds, a learning rate (η) of 0.20, a subsample equal to 1 and a maximum depth of each tree of 6. The validated model presents a high predictive power (*r2* = 0.96), with 96 % of variance explained.

Secondly, our study identifies the key features related to interactions with their IPA by using SHAP values of each topic for every review (*SHapley Additive exPlanations*, hereinafter, SHAP). SHAP values are conceptualised as “a unified measure of feature importance” (Lundberg & Lee, 2017, 4), and help us to interpret predictions not across the entire population but at an individual level. In this regard, SHAP values describe how much each topic in our model contributes, either positively or negatively, to increasing (or decreasing) users’ sentiment levels. Figure 10 shows the most influential topics in our model output. Social influence (topic 12), facilitating conditions to perform them (topic 17) and the integration between devices and services (topic 5) are, in this order, the largest key features to adopt and accept an IPA:

- Topic 12 is related to social influence, *i.e.*, users preferentially (and unsurprisingly) assess more (less) devices highly (lowly) rated features and benefits related to the pleasure obtained from interacting with family and friends, *e.g.*, “*The device ease of use and convenience brought joy and excitement to the family. When family and friends visit our home, they enjoy the features and benefits. Excellent purchase*”. If the family, as a relevant influencing factor in an in-home device, is satisfied with an IPA, individuals will buy it.
- An increasing value of topic 17 (related to negative facilitating conditions) is highly associated with negative effects on the model output, *i.e.*, “(…) *device refuses to connect to Wi-Fi after multiple attempts. Have reset the device and router - all other devices have no issues connecting to Wi-Fi (which is super-fast 1gbps fibre optic - so no problems with speed either)*”. SHAP dependence plots (Figure 11) coordinate with the feature values on the x-axis (topic i) and the corresponding SHAP values (log odds ratio) on the y-axis and help to identify negatively influential topics on our research model output.
- On the contrary, topic 5 (related to, for instance, *premium member of Spotify*) shows the opposite effect — higher values lead to a higher prediction of users’ sentiments.
- Like topic 17, growing values of topic 16 (related to negative facilitating conditions) are also associated with decreasing values in the reviewers’ sentiments.
- …

Finally, it is reasonable to propose the following additional question: How strong is the evidence that, for instance, the average SHAP values of topic *t* differ in their association with the users’ sentiments, *i.e.*, Amazon’s Echo and Google Home? Our research precisely reveals the differential model-based influence of the topics identified on users’ sentiments by the comparison between two means based on a Bayesian estimation that is analogous to an unpaired t-test using the current standard in the field of psychology, *i.e.*, a zero-centred Cauchy prior with scale $1/\sqrt{2}$ for effect size (see Morey & Rouder, 2018). Our study here applies a one-sided hypothesis test. The numerator is restricted to H1: $\delta_{\text{Amazon}} > 0$, and the denominator is restricted to H1: $\delta_{\text{Google}} > 0$. The main results (related to topics previously analysed) are set out below:

A) The corresponding Bayes factor ratio (see BF $_{Google}$ > BF $_{Amazon}$) provides greater evidence for the Google model compared to the Amazon model in the following cases:
- Cluster 1a: Topics 10, 16 and 17 related to the perception of adequacy in facilities and environment for IPA actions.
- Cluster 1b: Topics 4 and 20 referred to feeling understood based on conditions experienced.
- Cluster 2a: Topics 15 and 22 related to hands-free human-device interaction.
- Cluster 3: Topics 11 and 24 referring to quality performance.
- Cluster 4a: Topics 7 and 23 referring to (fun and entertaining) benefits; for instance, playing music and questions on Encyclopedia of knowledge deriving from the IPA conveniences that users have fun using.

B) The corresponding Bayes factor ratio (see BF $_{Google}$ > BF $_{Amazon}$) provides greater evidence for the Amazon model relative to the Google model in the following cases:
- Cluster 2a: Topics 6 and 19 (related to home automation and compatibility linked to switching costs).
- Cluster 2b: Topics 2 and 21 (referred to a cost-effective solution to help automate home)

Figures 12ab display density plots based on SHAP values grouped by devices and different evidence for Google Home (in Figure 12.a) or Amazon's Echo (in Figure 12.b).

::Figure 10 near here::

::Figure 11 near here::

::Figures 12ab near here::

5. Discussion:

Summary of results:

Based on TAM and U&GT, our study: (1) contributes to the literature on smart home technologies from the perspective of users' adoption and acceptance motives, (2) offers insights for theoretical and managerial implications concerning the effect of customers' service encounters (narrated in reviews) on their psychological sentiment, and (3) derives future recommendations. In this regard, despite analysts predicting that the worldwide spending on the Internet of Things is forecast to reach $742 billion in 2020 (IDC, 2020), and the growth and availability of IPA (Sharma, 2018), there is still scepticism in society regarding these innovations, and a need to examine their success factors to increase the popularity expected in (virtual) environments that are familiar to the users (*e.g.*, in-home).

The main results are, in this order:

1. The social entertainment associated with a particular task (preferentially related to playing music), social gaming or enjoyment, and a reduction of boredom from interacting or using in-home voice assistants with family or friends (social influence),
2. Their (doubtful) facilitating conditions or resources or perceived support available to perform a behaviour, *e.g.*, setting up the IPA (even in the presence of effort expectancy among *a bit lazy users*), and their integration with other platforms, other brands, mobile applications, or services (*e.g.*, subscriptions) that could create a barrier to entry and become incentives for utilisation to be convenient, ubiquitous and timesaving (enhancing limited performance expectancy),
3. Home automation and compatibility related to switching costs, performing behaviours automatically because of learning.
4. Effort expectancy (relating to hands-free human-device interaction while users do other things at the same time).

Theoretical implications:

Our research has a double complementary research purpose, exploratory and predictive. Moreover, our study faces a scenario that is "data-rich and theory-skeletal", which is characteristic of big data analysis. Consequently, our research examines data and evaluates different configurations using different techniques, being STM, clustering analysis, and XGBoost being especially appropriate for this task. Following such an exploratory approach, our analysis tries to locate topics of interest and to avoid any confirmatory purpose on potential associations between topics and IPA acceptance. Following Sánchez-Franco and Alonso-Dos-Santos (2020, p.16) "our research could be understood as a heuristic for theory building, applying an inductive perspective of reasoning to obtain clues that may point researchers and practitioners in a promising direction".

Accordingly, our research examines how the presence or absence of features, as well as the consequences of using the service, affect users' overall assessments. On the other hand, our research is therefore consistent with the proposal that the users' narratives offer valuable information. On the other hand, our research guides further research on the design of IPA to reflect the willingness to use on the users' side. Although our research model produces a relevant summary of customers' reviews and overall confirms and addresses the results of previous research, our results additionally provide insights not indicated in customers' ratings into how IPA improves their home automation, and additionally, they are more reliable and accurate than previous statistical results based on limited sample data.

Managerial implications:

Unlike the results of McLean and Osei-Frimpong (2019a), hedonic benefits from IPA are key to success. IPA should thus be capable of offering novel new ways to have fun and entertain which have not been there before and are beneficial for everyday life (*cf.* also Brown & Venkatesh, 2005). In particular, assessing the distinction between Amazon's Echo and Google Home, and according to topic 2 (referring to a cost-effective solution to help automate the home), there exists a higher consistency between the actual cost of IPA, especially in the case of Amazon's Echo, and how users compare the perceived benefits of the IPA to the monetary cost of using it. In the case of Google Home, designers should thus emphasise the beneficial trade-off between price and gratification for daily life by IPA connection to as many devices as possible, and foster network externalities, *i.e.*, increasing their adoption as a function of family, friends, or society.

Furthermore, when users (without specialist technical skills) perceive, for instance, that "an organizational and technical infrastructure exists to support the use of the system" (Venkatesh *et al.*, 2003, p.453), they may be strongly motivated to adopt it. 'Home automation and compatibility related to switching costs' are the features most highly rated in the comments of the Echo's reviewers. On the contrary, the frustration based on the perception of (poor) suitability in facilitating conditions is more present in Amazon's Echo reviewers in comparison with Google Home reviewers.

Accordingly, benefits gratification and facilitating conditions play a critical role and have a direct impact on the users' sentiment for IPA. Designers and managers recognise the challenge of increasing the customer satisfaction of current and potential users by adjusting doubtful users' technical skills and the benefits and functionalities of IPA to avoid getting bored after a short lapse of time. To maintain a truly loyal relationship, future human-centric computer interaction research must consider cognitive, hedonic, and social drivers related to multiple benefits gratifications when allocating their IPA efforts:

1. To increase customers' assessment by training instructions (where users find detailed instructions on how to set up and use the in-home voice device, and their skillset; *cf.* instructional protocol proposed by Koon *et al.*, 2020) to reduce the frustration of not achieving a setting up goal among users with a lower sense of self-efficacy.
2. To overcome relevant weaknesses, such as a lack of understanding of how the IPA works, the interoperability of technologies, and secure data storage, *i.e.*, what would users need to use IPA fully?

Finally, users are concerned about listening to their music playlist on their IPA or connecting it to different devices or apps (network effect). Otherwise, IPA could seem gimmicky by lacking the ability to improve users' everyday lives (life efficiency). In this sense, individuals are also highly motivated by social (family) benefits. If an IPA seeks to foster social spaces where family members interact and, consequently, maintain a satisfying loyal relationship with its users, managers must consider the main features related to both relationship quality components. In other words, IPA that provide the most appropriate combination of features and benefits and are aware of the limited skills of their users are more likely to achieve a competitive advantage and enable IPA diffusion in homes.

Limitations

Reviews are not written independently of each other. A topic-network structure may also change over time. Additionally, an improvement of algorithms is necessary to more easily develop topic models and select the optimal number of topics and metatopics. In this vein, a limitation lies in the self-selection bias when customers publish their reviews. The sample chosen is also made up of IPA users. Future research should select non-users to draw generalised conclusions about potential customers and assess the moderating effect of the self-efficacy of users, and their cognitive innovativeness (Huang & Liao, 2015).

Secondly, IPA adoption, acceptation and post-adoption need to be further analysed in a larger dataset, and a cross-demographic approach should be applied. In this context, future studies should analyse age-based differences. Digital technologies have been launched on the market to obtain, for instance, health benefits (Rondán-Cataluña *et al.*, 2020). The elderly population seeks to live autonomously and with an enhanced sense of overall well-being, and IPA to support ageing in place has received little attention (Corbett *et al.*, 2021). "With ageing, there can be problems with vision; hence voice assistance can be integrated with the smart-homes meant for elderly people" to promote cost-effective benefits for their health and their well-being (Pal *et al.*, 2018, p.51248). To sum up, the elderly population assesses in-home device experiences using different cognitive criteria -centred, for example, on needs for daily usefulness (*e.g.*, caregiving support) or a perceived sense of safety- and consequently, facing different facilitating conditions (*e.g.*, hands-free feature) and barriers to use.

References:

Alepis, E., and Patsakis, C. 2017. "Monkey says, monkey does: Security and privacy on voice assistants". *IEEE Transactions on Affective Computing*, 5,17841–17851.

Belanche, D., Casaló, L.V., Flavián, C., and Schepers, J. 2020. "Robots or frontline employees? Exploring customers' attributions of responsibility and stability after service failure or success". *Journal of Service Management*, 31 (2), 267-289

Bentley, F., Luvogt, C., Silverman, M., Wirasinghe, R., White, B., and Lottridge, D. 2018. "Understanding the long-term use of smart speaker assistants". In G. D. Abowd, J. Kay, J. A. Kientz, V. Kostakos, S. Santini, J. Scott, & K. Yatani, (Eds.), "Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies" (pp. 91:1–91:24). New York: ACM.

Bischof, J., and Airoldi. E. 2012. "Summarizing topical content with word frequency and exclusivity". In John Langford and Joelle Pineau (eds.), Proceedings of the 29th International Conference on Machine Learning (ICML-12). New York, NY: Omnipress, 201–208.

Blei, D.M., Ng, A.Y., and Jordan, M.I. 2003. "Latent Dirichlet allocation". *Journal of Machine Learning Research*, 3: 9931022.

Bortz, J., Lienert, G. A., and Klaus Boehnke. 1990. Verteilungsfreie Methoden in der Biostatistik. Berlin et al: Springer.

Brown, S.A., and Venkatesh, V. 2005. "A Model of Adoption of Technology in the Household: A Baseline model test and extension incorporating household life cycle". *MIS Quarterly*, 29 (3), 399-426.

Camén, C., P. Gottfridsson, and B. Rundh. 2011. "To Trust or Not to Trust? Formal Contracts and the Building of Long-term Relationships". *Management Decision*, 49 (3), 365–38

Chen, T., and Guestrin, C. 2016. Xgboost: A scalable tree boosting system. In "KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining" (pp. 785-794), New York, NY, USA.

Chiang, Y-S., Chang, R-Ch., Chuang, Y-L., Chou, Sh-Y., Lee, H-P., Lin, I-J., Chen, J., and Chang, Y-J. 2020. Exploring the design space of user-system communication for smart-home routine assistants. In "CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems" (pp. 1–14), Honolulu, USA.

Chuang, J., Manning, C.D., and Heer, J. 2012. "Termite: Visualization Techniques for Assessing Textual Topic Models". Proceedings of the International Working Conference on Advanced Visual Interfaces, May 21-25, 2012, Capri Island, Italy.

Climent, F., Momparler, A., and Carmona, P. 2019. "Anticipating bank distress in the Eurozone: An Extreme Gradient Boosting approach". *Journal of Business Research*, 101, 885-896.

Corbett, C.F., Combs, E.M., Wright, P.J. *et al.* 2021. Virtual Home Assistant Use and Perceptions of Usefulness by Older Adults and Support Person Dyads. *International Journal of Environmental Research and Public Health*, 18(3), 1113.

Davis, F., Bagozzi, R., and Warshaw, P. 1992. "Extrinsic and Intrinsic Motivation to Use Computers in the Workplace". *Journal of Applied Social Psychology*, 22(14), 1111–1132

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

de Barcelos Silva, A., Gomes, M.M., da Costa, C.A., da Rosa Righi, R. *et al.* 2020. "Intelligent Personal Assistants: A systematic literature review". *Expert Systems with Applications*, 147, 113193.

Edwards, J., Liu, H., Zhou, T., Gould, S., Michael, L., Clark, H., Doyle, P.R., and Cowan, B.R. 2019. Multitasking with Alexa: How using Intelligent Personal Assistants impacts language-based primary task performance. In “CUI '19: Proceedings of the 1st International Conference on Conversational User Interfaces” (art.4, pp. 1-7), Dublin, Ireland.

Eyssel, F., Kuchenbrandt, D., Bobinger, S., de Ruiter, L., and Hegel, F. 2012. ‘If you sound like me, you must be more human': on the interplay of robot and user features on human- robot acceptance and anthropomorphism. In “HRI '12: Proceedings of the seventh annual ACM/IEEE International Conference on Human-Robot Interaction”, (pp.125-126), Boston, USA.

Gaisser, and Utz. 2020. “Hello Alexa!” Smart Speakers as new digital roommates – An analysis of online posts and reviews. In “ICA 2020, International Communication Association”, Gold Coast, Australia.

Gan, C., and Li, H. 2018. “Understanding the effects of gratifications on the continuance intention to use WeChat in China: A perspective on uses and gratifications”. *Computers in Human Behavior,* 78, 306–315.

Gerring, J. 2001. *Social science methodology: A unified framework*. Cambridge: Cambridge University Press.

Goodin, R. E. 1977. “Symbolic rewards: Being bought off cheaply”. *Political Studies*, 25(3), 383-396.

Han, H.J., Mankad, S., Gavirneni, N., and Verma, R. 2016. “What guests really think of your hotel: Text analytics of online customer reviews”. *Cornell Hospitality Report,* 16 (2), 3-17.

Han, S., and Yang, H. 2018, "Understanding adoption of intelligent personal assistants: A parasocial relationship perspective". *Industrial Management & Data Systems*, 118 (3), 618-636.

He, L., Han, D., Zhou, X., and Qu, Z. 2020. The voice of drug consumers: Online textual review analysis using structural topic model. *International Journal of Environmental Research and Public Health*, 17*, 3648*.

Holsapple, C.W., and Wu, J. 2015. “User acceptance of virtual worlds: The hedonic framework”. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 38 (4).

Hoy, M.B. 2018. “Alexa, Siri, Cortana, and more: An introduction to voice assistants”. *Medical Reference Services Quarterly*, 37 (1), 81-88.

Huang, T., and Liao, S. 2015. “A model of acceptance of augmented-reality interactive technology: the moderating role of cognitive innovativeness”. *Electronic Commerce Research,* 15, 269–295.

Hubert, L., and Arabie, P. 1985. “Comparing partitions”. *Journal of Classification,* 2 (1), 193–218.

International Data Corporation (IDC) 2020. “Worldwide Spending on the Internet of Things will slow in 2020 then return to double-digit growth”. IDC's Spending Guides, June 18. Available at: https://www.idc.com/getdoc.jsp?containerId=prUS46609320

Jockers, M. 2017. “Extracts Sentiment and Sentiment-Derived Plot Arcs from Text”, available at: https://cran.r-project.org/web/packages/syuzhet/syuzhet.pdf (R package version 1.0.4).

Joo, J., and Sang, Y. 2013. “Exploring Koreans’ smartphone usage: An integrated model of the technology acceptance model and uses and gratifications theory”. *Computers in Human Behavior,* 29(6), 2013, 2512-2518.

Katz, E., Blumler, J. G., and Gurevitch, M. 1974. Utilization of mass communication by the individual. In J. G. Blumler, & E. Katz (Eds.), “The Uses of Mass Communications: Current Perspectives on Gratifications Research” (pp. 19-32). Beverly Hills: Sage.

Koon L.M., McGlynn S.A., Blocker K.A., and Rogers W.A. 2020. “Perceptions of digital assistants from early adopters aged 55+”. *Ergonomics in Design: The Quarterly of Human Factors Applications,* 28 (1), 16-23.

Kuhn, M. 2020. “Caret: classification and regression training”, available at: https://cran.r-project.org/web/packages/caret/caret.pdf (R package version 6.0–86).

Limayem, M., Hirt, S., and Cheung, C. 2007. “How habit limits the predictive power of intention: The case of information systems continuance”. *Management Information Systems Quarterly*, 31 (4), 705-737

Livaditi, J., Vassilopoulou, K., Lougos, C., and Chorianopoulos, K. 2003. “Needs and gratifications for interactive TV applications: Implications for designers”. HICSS '03: Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS'03) - Track 4 - Volume 4 January 2003.

Lopatovska, I., and Williams, H. 2018. Personification of the Amazon Alexa: BFF or a Mindless Companion. In “CHIIR '18: Proceedings of the 2018 Conference on Human Information Interaction & Retrieval” (pp.265–268), New Brunswick, NJ, USA.

Lopatovska, I., Rink, K., Knight, I., Raines, K., Cosenza, K., Williams, H., and Martinez, A. 2018. “Talk to me: Exploring user interactions with the Amazon Alexa”. *Journal of Librarianship and Information Science*, 51(4), 984–997

Lovato, S., and Piper, A. M. 2015. “*“Siri, is This You?”* Understanding young children's interactions with voice input systems”. In “IDC '15: Proceedings of the 14th International Conference on Interaction Design and Children” (pp. 335–338), Boston, Massachusetts, USA.

Luger, E., and Sellen, A. 2016. “‘Like having a really bad PA’”: The gulf between user expectation and experience of conversational agents”. Proceedings of the 2016 CHI-Conference on human factors in computing systems, ACM, New York, USA, 5286-5297.

Lundberg, S.M., and Lee, S-I. 2017. A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), “Advances in Neural Information Processing Systems” 30 (pp.4765–4774). Curran Associates, Inc.

Maechler M., Rousseeuw P., Struyf A. *et al.* 2021. “cluster: Cluster Analysis Basics and Extensions”, available at: https://CRAN.R-project.org/package=cluster (R package version 2.1.2).

Mair, P., Funke, S., Harloff, J., and von Eye, A. 2017. “cfa: Configural Frequency Analysis (CFA)”, available at: https://cran.r-project.org/web/packages/cfa/index.html (R package version 0.10-0).

Marikyan, D., Papagiannidis, S., and Alamanos, E. 2019. “A systematic review of the smart home literature: A user perspective”. *Technological Forecasting and Social Change*, 138, 139-154.

McLean, G., and Osei-Frimpong, K. 2019a. “Hey Alexa ... examine the variables influencing the use of artificial intelligent in-home voice assistants”. *Computers in Human Behavior*, 99 (October), 28-37.

McLean, G., and Osei-Frimpong, K. 2019b. “Chat now… Examining the variables influencing the use of online live chat”. *Technological Forecasting and Social Change*, 146, 55-67.

Mennicken S., and Huang E.M. 2012. “Hacking the natural habitat: An in-the-wild study of smart homes, their development, and the people who live in them”. In Kay J., Lukowicz P., Tokuda H., Olivier P., & Krüger A. (Eds), “Pervasive Computing. Pervasive 2012. Lecture Notes in Computer Science”, 7319. Berlin, Heidelberg: Springer.

Mimno, D., Wallach, H.M., Talley, E., Leenders, M., and McCallum, A. 2011. “Optimizing semantic coherence in topic models”. In “Proceedings of the Conference on Empirical Methods in Natural Language Processing EMNLP ’11” (pp.262–72). Edinburgh, UK: Association for Computational Linguistics.

Mohammad, S., and Turney, P. 2010. “Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon”. In “Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*”*, LA, California.

Mohammad, S.M., and Turney, P.D. 2013. “Crowd sourcing a word-emotion association lexicon”. *Computational Intelligence*, 29(3), 436–465

Morey, R.D., and Rouder, J.N. 2018. “BayesFactor: Computation of Bayes Factors for common designs”, available at: https://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf (R package version 0.9.12.4-2).

Nijssen, E.J., Schepers, J.J.L., and Belanche, D. 2016. "Why did they do it? How customers’ self-service technology introduction attributions affect the customer-provider relationship". *Journal of Service Management*, 27 (3), 276 – 298.

Pal, D., Funilkul S., Vanijja, V, and Papasratorn, B. 2018. “Analyzing the Elderly Users’ Adoption of Smart-Home Services”. IEEE Access, 6, pp. 51238-51252.

Pantano, E., and Pizzi, G. 2020. “Forecasting artificial intelligence on online customer assistance: Evidence from chatbot patents analysis”. *Journal of Retailing and Consumer Services*, 55, 102096.

Park, C.W., B.J. Jaworski, and D.J. MacInnis. 1986. “Strategic brand concept-image management”. *Journal of Marketing*, 50 (3), 135–145.

Pradhan, A., Mehta, K., and Findlater, L. 2018. “Accessibility came by accident: Use of voice-controlled intelligent personal assistants by people with disabilities”. In “CHI '18: CHI Conference on Human Factors in Computing Systems”, 459, Montreal, Canada.

Quinn, K.M., Monroe, B.L., Colaresi, M., Crespin, M.H., and Radev, D.R. 2010. “How to analyse political attention with minimal assumptions and costs”. *American Journal of Political Science*, 54 (1), 209-228.

Ramírez-Correa, P., Rondán-Cataluña, F.J., and Arenas-Gaitán, J. 2019. “Analysing the acceptation of online games in mobile devices: An application of UTAUT2”. *Journal of Retailing and Consumer Services*, 50, 85-93.

Rauschnabel, P. 2018. "A Conceptual Uses & Gratification Framework on the Use of Augmented Reality Smart Glasses". In T. Jung and M.C. tom Dieck (eds.), *Augmented Reality and Virtual Reality*, 211-227.

Rauschnabel, P. A., He, J., and Ro, Y. K. 2018. “Antecedents to the adoption of augmented reality smart glasses: A closer look at privacy risks.” *Journal of Business Research*, 92, 374–384.

Roberts, M., Stewart, B., and Tingley, D. 2018. “stm: R Package for structural topic models”. *Journal of Statistical Software*, 91 (2).

Roberts, M., Stewart, B., Tingley, D. and Benoit, K. 2020. “stm: Estimation of the Structural Topic Model”, available at: https://cran.r-project.org/web/packages/stm/index.html (R package version 1.3.6).

Roberts, M.E., Stewart, B.M, Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S.K., Albertson, B., and Rand, D.G. 2014. “Structural topic models for open-ended survey responses”. *American Journal of Political Science*, 58 (4), 1064-1082.

Rondán-Cataluña F.J., Ramírez-Correa P.E., Arenas-Gaitán J. *et al.* 2020. Social Network Communications in Chilean Older Adults. *International Journal of Environmental Research and Public Health,* 17(17), 6078.

Rubin, A.M. 2002. "The uses-and-gratifications perspective of media effects". Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Ryan, R., and Deci, E. 2000. "Intrinsic and extrinsic motivations: Classic definitions and new directions". *Contemporary Educational Psychology*, 25 (1), 54–67.

Sánchez-Franco, M. J., and Alonso-Dos-Santos, M. 2020. "Exploring gender-based influences on key features of Airbnb accommodations". Economic Research - Ekonomska Istraživanja, in press.

Sánchez-Franco, M.J., and Roldán, J.L. 2005. "Web acceptance and usage model: A comparison between goal-directed and experiential web users". *Internet Research*, 15 (1), 21-48.

Sánchez-Franco, M.J., Troyano, J.A., and Alonso-Dos-Santos, M. 2020. "Fuzzy metatopics predicting prices of Airbnb accommodations". *Journal of Intelligent & Fuzzy Systems*, forthcoming.

Sharma, A. 2018. "3 Ways Voice AI Will Change the Way We Shop in 2018". Adweek, January. Available at: https://www.adweek.com/digital/3-ways-voice-ai-will-change-the-way-we-shop-in-2018/

Shin, J., Park, Y, and Lee, D. 2018. "Who will be smart home users? An analysis of adoption and diffusion of smart homes". *Technological Forecasting and Social Change*, 134, 246-253.

Sievert, C. and Shirley, K 2014. {LDA}vis: A method for visualizing and interpreting topics. In "Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces", (pp.63–70), Baltimore, Maryland, USA.

Silge, J., Robinson, D., Fay, C. *et al.* 2020. tidytext: Text Mining and Analysis Using Tidy Data Principles in R, available at: https://CRAN.R-project.org/package=tidytext (R package version 0.2.6).

Statista 2020. Global Smart Speaker Market Share 2018 and 2019, by platform (published by H. Tankovska, Sep 1, 2020). Available at: https://www.statista.com/statistics/1005558/worldwide-smart-speaker-market-share/

Teo, T. 2010. "Examining the influence of subjective norm and facilitating conditions on the intention to use technology among pre-service teachers: a structural equation modelling of an extended technology acceptance model". *Asia Pacific Education Review*, 11 (2), 253–262.

Tsai, H.S., Shillair, R., and Cotten, S.R. 2015. "Social support and 'playing around': An examination of how older adults acquire digital literacy with tablet computers". *Journal of Applied Gerontology*, 36, 29–55.

Vasilakos, A., and Pedrycz, W. 2006. Ambient Intelligence, Wireless Networking, Ubiquitous Computing. Artech House, Norwood, MA, USA.

Venkatesh, V. Y. L. Thong, J., and Xu, X. 2012. "Consumer acceptance and use of information Technology: Extending the unified theory of acceptance and use of technology". *MIS Quarterly*, 36(1), 157-178.

Venkatesh, V., and Bala, H., 2008. "Technology acceptance model 3 and a research agenda on interventions". *Decision Science*, 39 (2), 273-315.

Venkatesh, V., and Davis, F.D., 2000. "A theoretical extension of the technology acceptance model: four longitudinal field studies". *Management Science*, 46 (2), 186–204.

Venkatesh, V., Morris, M.G., Davis, G. B., and Davis, F.D. 2003. "User acceptance of information technology: Toward a unified view". *MIS Quarterly*, 27(3), 425–478.

Wee, T.T.T., and Ming, M.C.H. 2003. "Leveraging on symbolic values and meanings in branding". *Journal of Brand Management*,10 (3), 208–218.

Weibull, L. 1985. Structural factors in gratifications research. In K. E. Rosengren, L. A. Wenner, & P. Palmgreen (Eds.), Media gratifications research: Current perspectives (pp. 123–157). Beverly Hills, CA: Sage.

Wickham, H., François R., Henry, L., and Müller, K. 2020. dplyr: A Grammar of Data Manipulation", available at: https://CRAN.R-project.org/package=dplyr (R package version 1.0.2).

Yang, H., and Lee, H. 2019. "Understanding user behaviour of virtual personal assistant devices". *Information Systems and e-Business Management*, 17, 65–87.

Yang, H., Lee H., and Zo H. 2017. "User acceptance of smart home services: An extension of the theory of planned behavior". *Industrial Management & Data Systems*, 117 (1), 68–89.

Figure 1. Steps for transforming free-form text into a structured form, and main research questions.

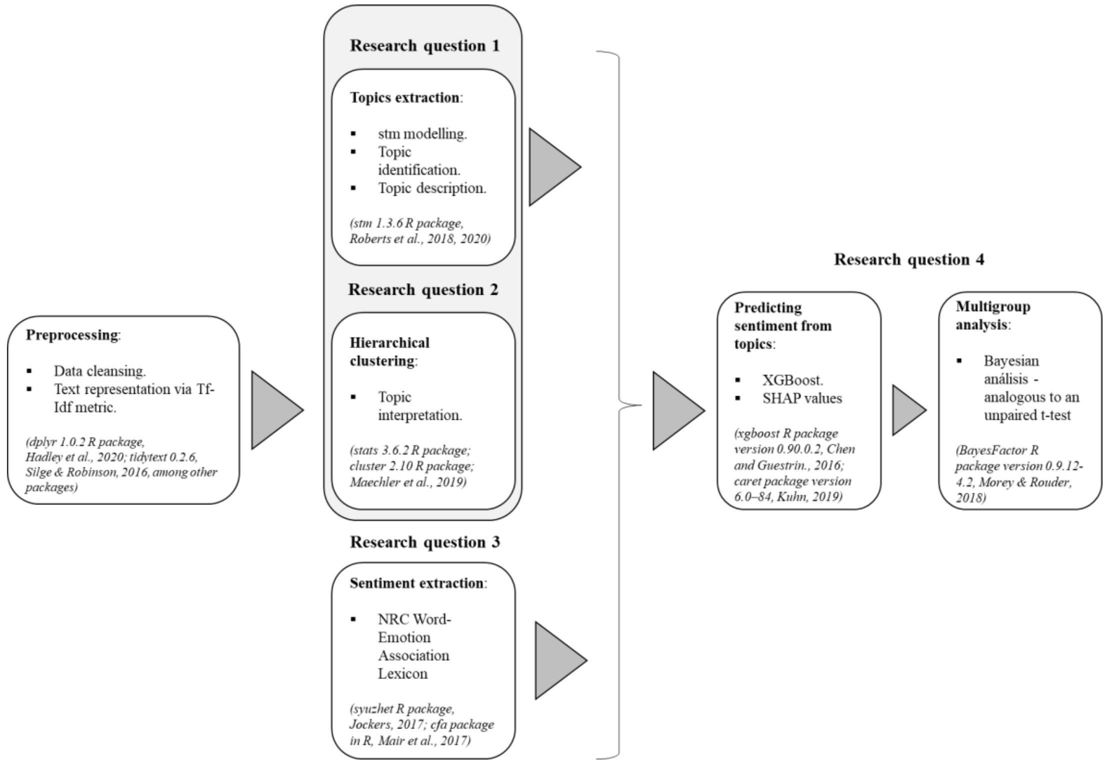


Figure 2. Distribution of the sample by API model.

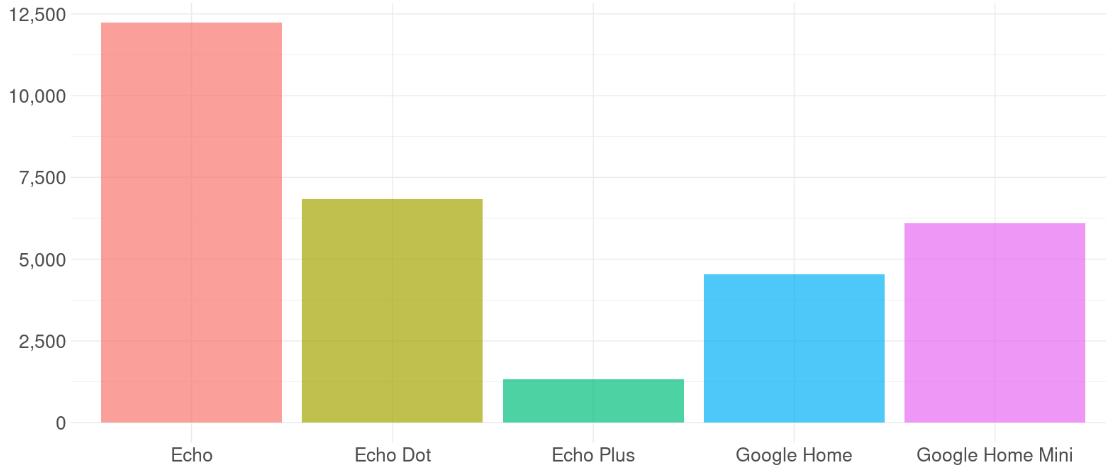
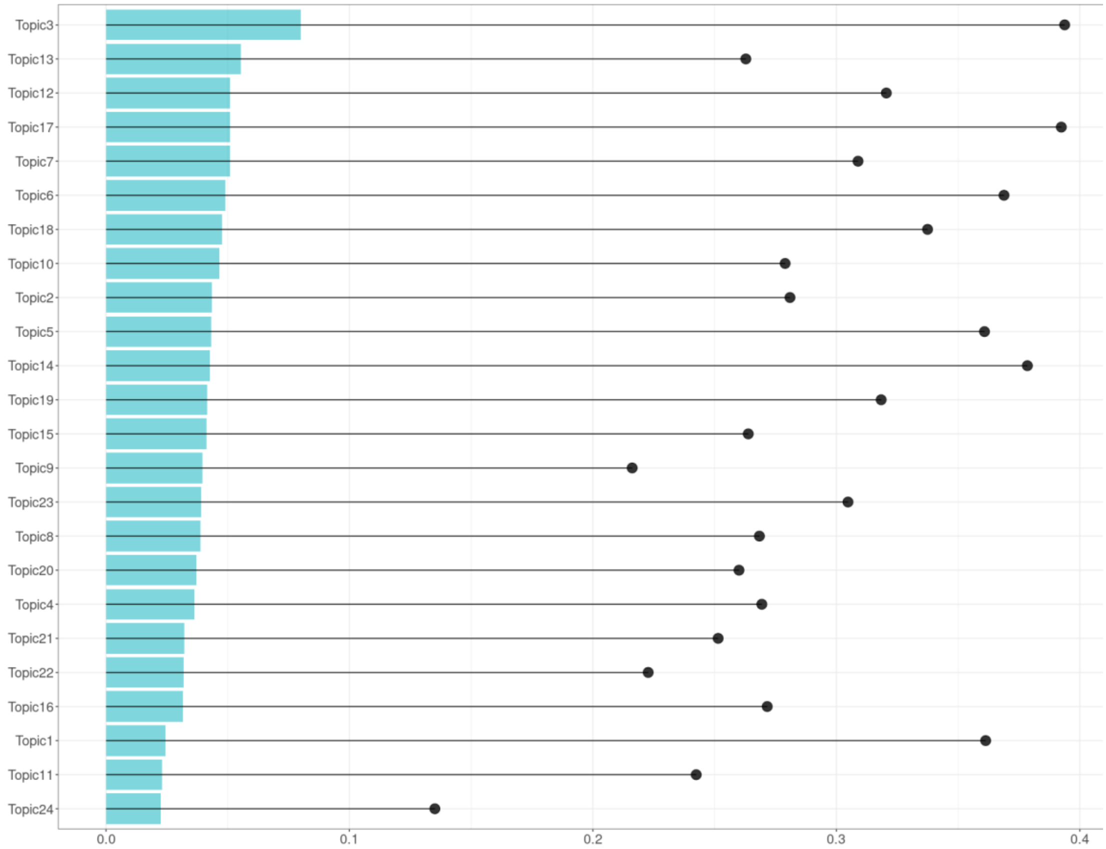


Figure 3.a. Graphical display of estimated topic proportions (bar graph) and Gini coefficients (point graph).



* “λ [here, 0.60] determines the weight given to the probability of term w under topic k relative to its lift (measuring both on the log scale)” (Sievert & Shirley, 2014, p.66).

Figure 3.b. Comparison of topic composition between Amazon’s Echo and Google Home based on estimated topic proportions.

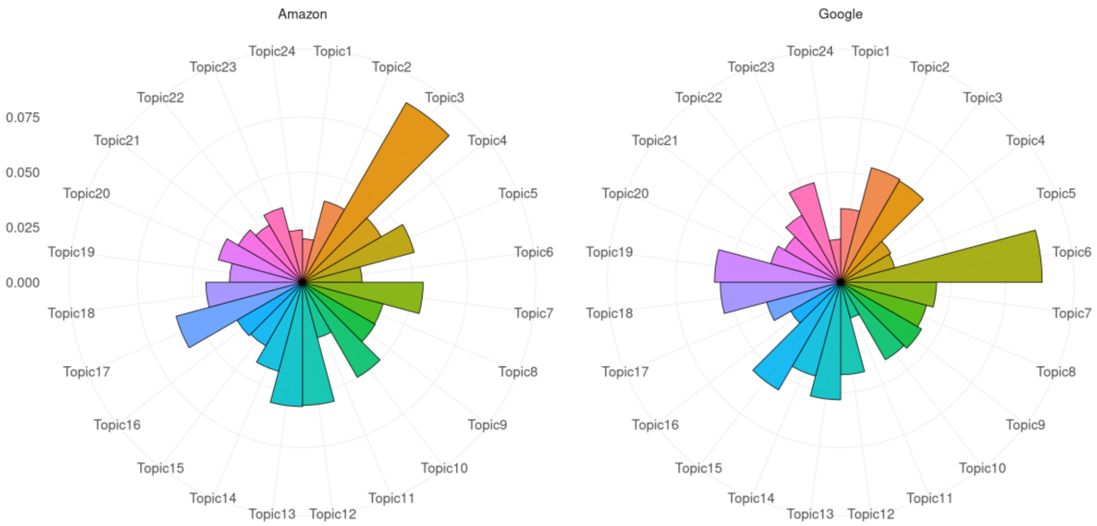
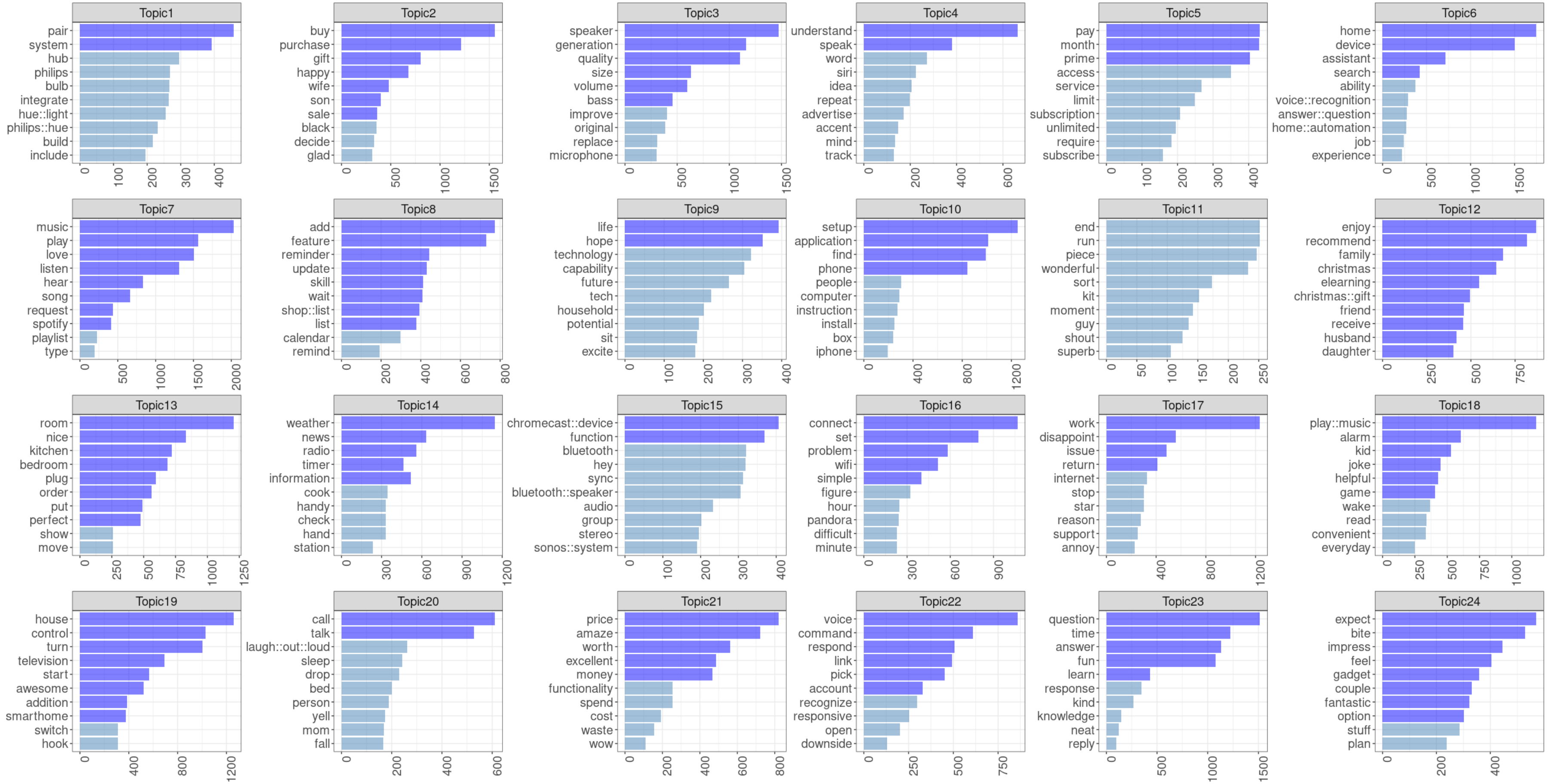
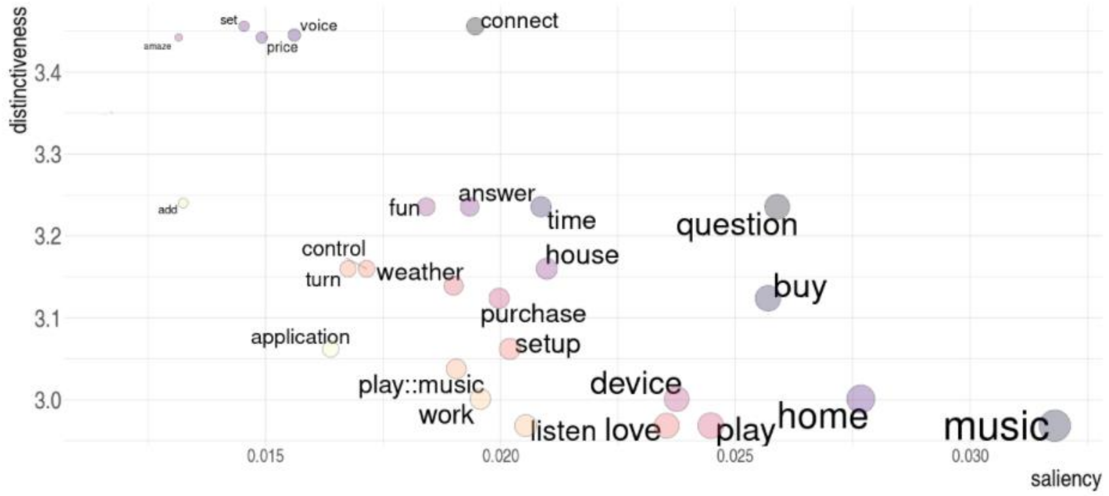


Figure 4. Top 10 terms ordered by relevance values (λ = 0.6).



* "λ [here, 0.60] determines the weight given to the probability of term w under topic k relative to its lift (measuring both on the log scale)" (Sievert & Shirley, 2014, p.66).
** y-axis represents estimated term frequency within the topic *t*.
*** The darker bars represent terms with saliency coefficient above the median of saliency-values.

Figure 5. Most relevant terms based on saliency and distinctiveness coefficients.



* Point size depends on the relevance-value of the term.

Figure 6. Agglomerative hierarchical cluster analysis. Dendrogram.

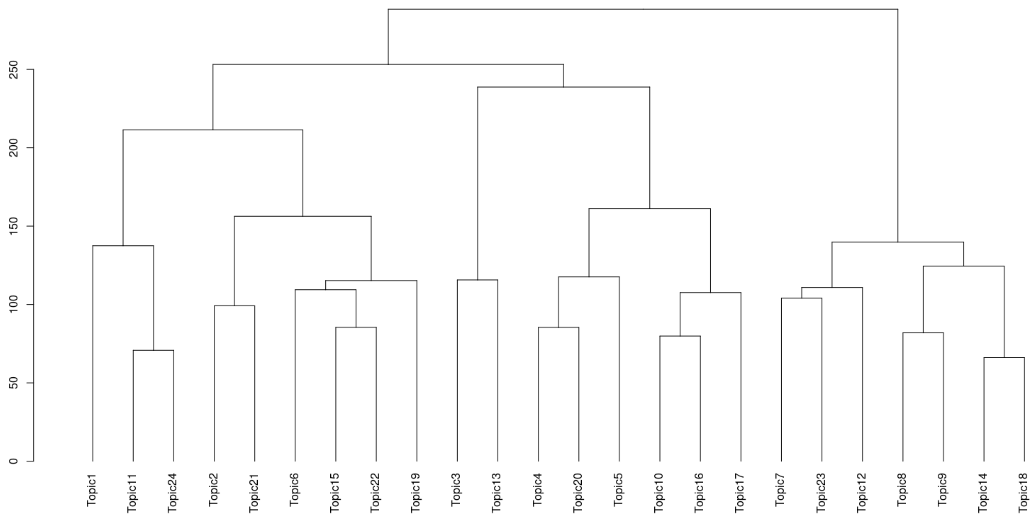


Figure 7. Score- and sentiment-plots by model.

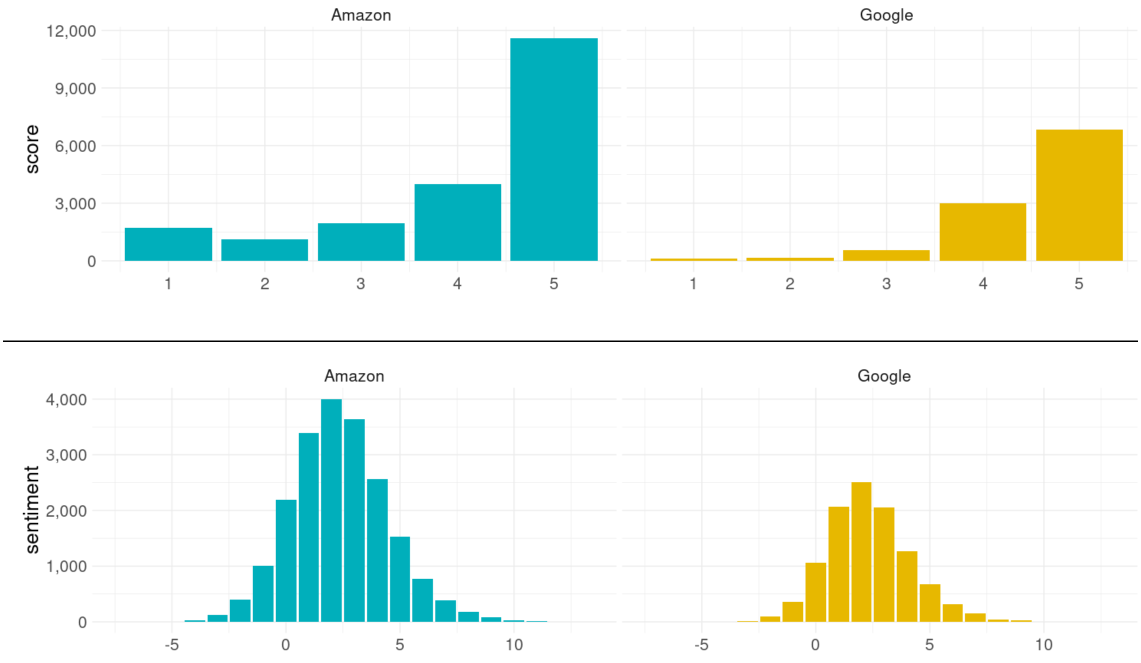


Figure 8. Sentiment plots by model.

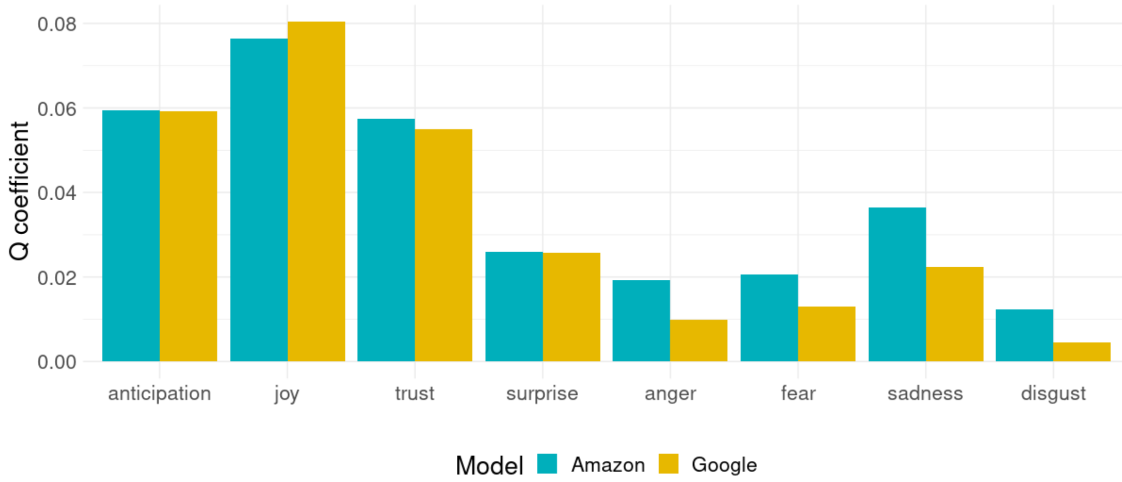


Figure 9. Top frequent terms from reviews to (positive- and negative-valence) sentiment.

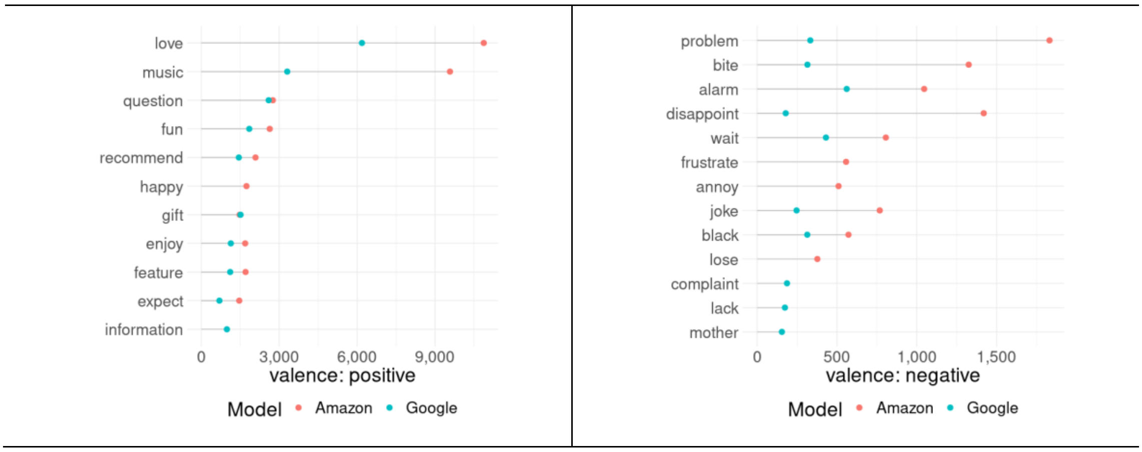


Figure 10. Feature importance displayed by local SHAP values.

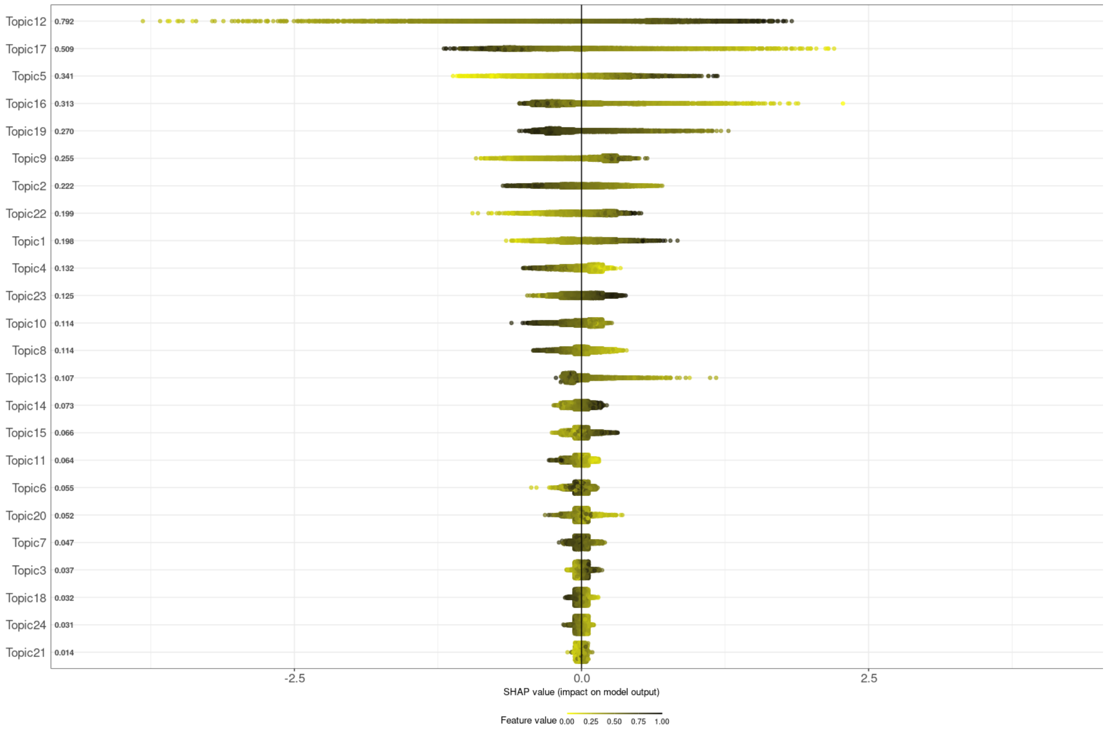


Figure 11. SHAP dependence plots.

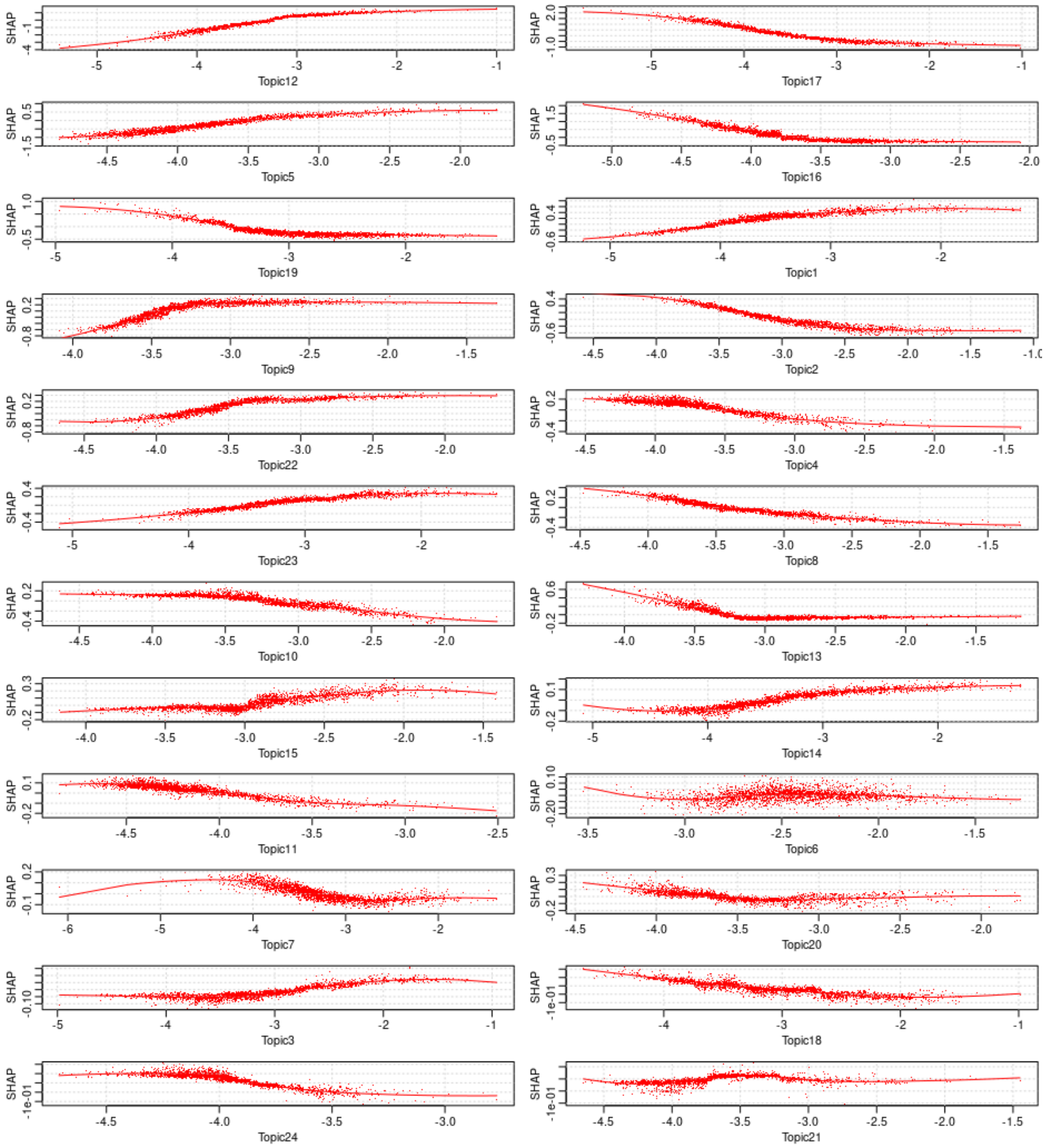


Figure 12.a. Google Home: Density plot with median lines (and marginal rug).

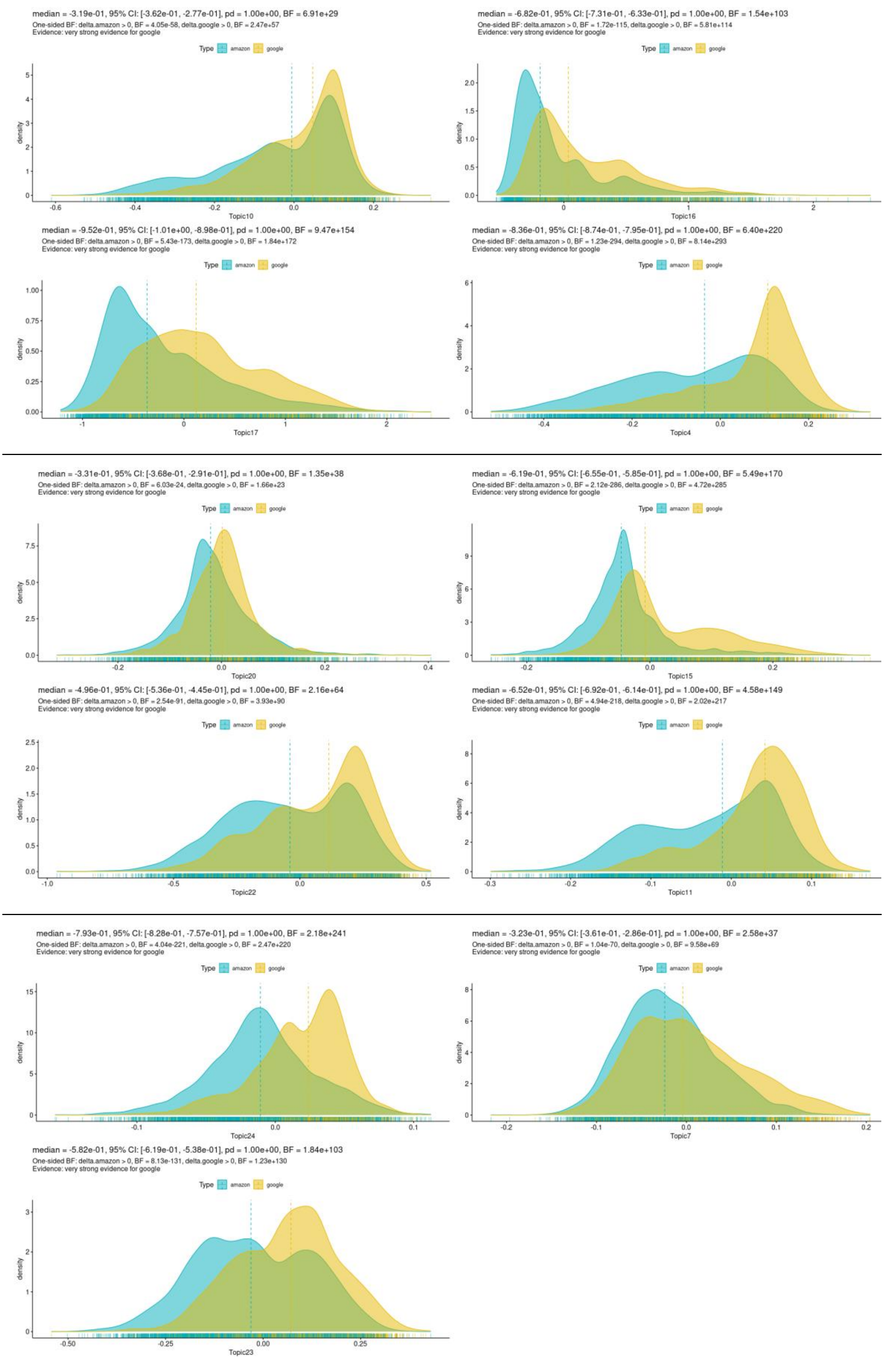


Figure 12.b. Amazon’s Echo: Density plot with median lines (and marginal rug).

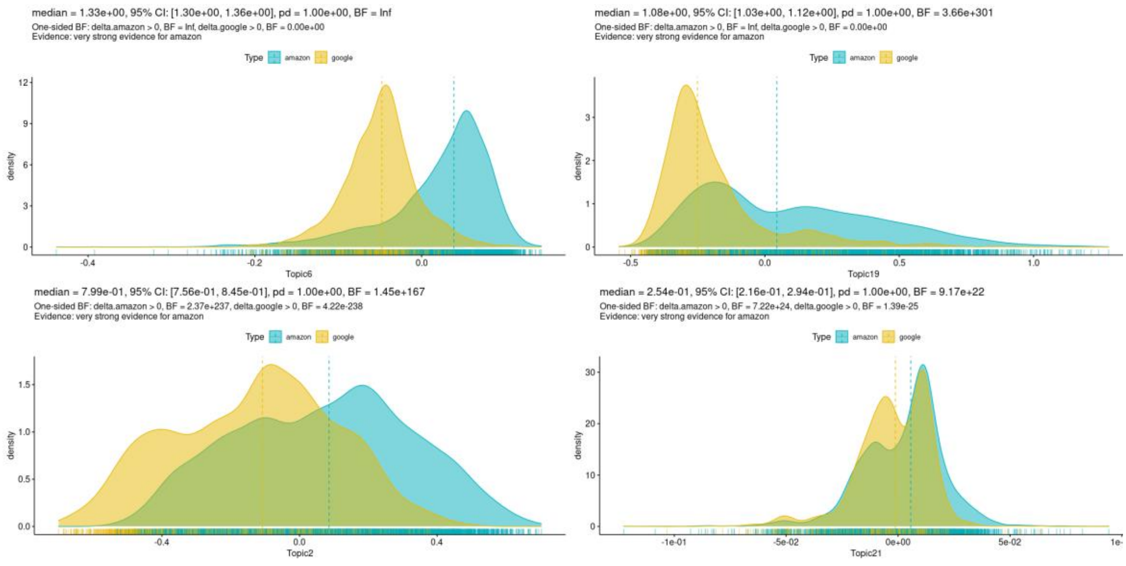


Table 1. Summary statistics: Score vs Sentiment.

Statistics	Score	Sentiment
mean	4.25	2.34
median	5	2
mode	5	2
range	[1,5]	[-7,13]
IQR	1	3
sd	1.14	2.04

Table 2. Results of a CFA test analysing.

Model + Emotion		n	expected	Q	chisq (1)	p.chisq (2)	sig.chisq (3)	z (4)	p.z (5)	sig.z (6)
Positive emotions										
Google	positive	42843	16804	0.3021	40350	0	TRUE	201	0	TRUE
Amazon	positive	39933	16428	0.2715	33632	0	TRUE	183	0	TRUE
Amazon	negative	12798	1287	0.1132	102984	0	TRUE	321	0	TRUE
Google	negative	7432	712	0.0657	63381	0	TRUE	252	0	TRUE
Positive emotions										
Google	joy	22934	2988.9	1.26e-01	133093	0	TRUE	365	0	TRUE
Amazon	joy	21984	3263.2	1.19e-01	107401	0	TRUE	328	0	TRUE
Google	anticipation	16610	1608.7	9.40e-02	139890	0	TRUE	374	0	TRUE
Amazon	anticipation	16770	1849.8	9.37e-02	120341	0	TRUE	347	0	TRUE
Amazon	trust	16126	1677.5	9.06e-02	124442	0	TRUE	353	0	TRUE
Google	trust	15354	1402.4	8.73e-02	138795	0	TRUE	373	0	TRUE
Amazon	surprise	7034	326.0	4.17e-02	138028	0	TRUE	372	0	TRUE
Google	surprise	6991	284.5	4.17e-02	158101	0	TRUE	398	0	TRUE
Negative emotions										
Amazon	sadness	9919	525.2	5.85e-02	168031	0	TRUE	410	0	TRUE
Google	sadness	6103	283.7	3.62e-02	119361	0	TRUE	345	0	TRUE
Amazon	fear	5521	164.5	3.33e-02	174429	0	TRUE	418	0	TRUE
Amazon	anger	5146	132.8	3.11e-02	189283	0	TRUE	435	0	TRUE
Google	fear	3495	91.4	2.11e-02	126705	0	TRUE	356	0	TRUE
Amazon	disgust	3300	48.9	2.02e-02	216107	0	TRUE	465	0	TRUE
Google	anger	2662	60.3	1.62e-02	112240	0	TRUE	335	0	TRUE
Google	disgust	1185	15.4	7.26e-03	88708	0	TRUE	298	0	TRUE

Note 1: df: 2.32e+02 | Chi-square = 2.25e+06 | $p < 0.0001$ | Sum of counts = 1.61e+05

Note 2: (1) Chi squared for the given configuration; (2) p for the chi-squared test; (3) Is it significant (will Bonferroni adjust if the Bonferroni argument is set)?; (4) z-approximation for chi squared; (5) p of z-test; (6) Is it significant (will Bonferroni-adjust if the Bonferroni argument is set)?

HIGHLIGHTS:

- IPA users mainly assess hedonic benefits from interacting with family and friends.
- Facilitating conditions are highly associated with positive sentiments.
- Google Home should emphasise the trade-off between its price and daily benefits.
- Perception of (poor) facilitating conditions is more present in Amazon Echo users.
- Designers should fit users' technical skills and IPA benefits or functionalities.