Three Essays on Big Data Analytics in Marketing

Hamid Shirdastian

A Thesis
In the Department of
Marketing

Presented in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy (Business Administration)

at Concordia University
Montreal, Quebec, Canada

March 2022

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Hamid Shirdastian

Entitled: Three essays on big data analytics in marketing

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Business Administration)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Saif Ullah ______________________ Chair

Dr. Chankon Kim ______________________ External Examiner

Dr. Roberto G. de Almeida ______________________ External to Program

Dr. Michele Paulin ______________________ Examiner

Dr. Yu Ma ______________________ Examiner

Dr. Michel Laroche ______________________ Thesis Supervisor

Approved By: Dr. Cedric Lesage ______________________ Graduate Program Director

____

Dr. Anne-Marie Croteau, Dean of John Molson School of Business

April, 28, 2022

ABSTRACT

Three Essays on Big Data Analytics in Marketing

Hamid Shirdastian, Ph.D.
Concordia University, 2022

This thesis investigates how the immense amount of real-time and retrospective data can contribute to marketing theories and practices in branding, advertising, and communications.

In the first essay, brand management in social media is studied due to its vast potential as well as the companies' interests in utilizing it for branding purposes. More specifically, the research examines the sentiments toward a brand, via brand authenticity, to identify both the reasons for positive or negative sentiments on social media and its polarity. Practically speaking, while firms need insights about users' sentiment towards their brand, knowing just that it is a simple positive or negative sentiment does not provide them with enough information. They might ask themselves why users like us or what the reason is behind the negative sentiment. Using three qualitative studies, along with the latent semantic analysis (LSA) and the support vector machine (SVM), the findings illustrate the effectiveness of the proposed procedure of brand authenticity sentiment analysis to predict both the brand authenticity dimensions and their level of sentiment.

The second essay is concerned with the immense amount of locational and destination-based data for advertising purposes. That comes from the fact that more and more apps have access to users' locations and destinations. As the context, the sharing economy is selected to test if a ridesharing app, i.e. Uber, can alter one's destination by providing relevant destination-based ads. Examining construal levels (i.e. spatial distance and cultural distance) show the effects on the relationships between attitudes towards destination-based advertising and redemption of marketing incentives as well as app reuse intention. Two experimental mock apps that mimic a well-known ridesharing app in North America are used. The findings provide implications for the construal level theory, the theory of planned behavior, and how practitioners can use it to alter planned behavior, i.e. planned destinations.

The third essay is also in the context of the sharing economy. Building on the speech act theory, the sale description impact on consumers' social cognitions of service providers is investigated to find how they could be used to generate new content. Findings suggest the role of linguistic concreteness, service provider type, and sentiment analysis on the perceived level of warmth and competence of the service providers. Using two text mining methods (frequency analysis and modified LDA), the differences between the four kinds of property descriptions and their hosts are explained. Then, the findings are employed to better train the Natural language generation (NLG) algorithm. This shows how Long-Short-Term Memory (LSTM), as well as cleaned input, could help service providers generate more engaging content to communicate their service.

The research findings on each of these streams could contribute to several theories and facilitate further inquiries into big data analytics in marketing. This research also provides marketing practitioners with reliable and valid theories, models and decision support systems to gain insights and propose appropriate strategies to strengthen their firm.

RÉSUMÉ

Trois Essais sur de Grandes Analyses de Données dans le Marketing

Hamid Shirdastian, Ph.D.

Université Concordia, 2022

Cette thèse examine comment la quantité immense de données en temps réel et les données rétrospectives peuvent contribuer aux théories et aux pratiques commerciales dans l'image de marque, la publicité, et les communications.

Dans le premier essai, la gestion de la marque dans les médias sociaux est étudiée en raison de son vaste potentiel ainsi que les intérêts des entreprises à l'utiliser à des fins de marque. Plus précisément, la recherche examine les sentiments vers une marque, par l'authenticité de la marque, afin d'identifier les raisons de sentiments positifs ou négatifs sur les médias sociaux et sa polarité. En pratique, alors que les entreprises ont besoin d'idées sur le sentiment des utilisateurs envers leur marque, sachant simplement qu'il est un positif simple, ou un sentiment négatif ne leur fournit pas suffisamment d'informations. Ils pourraient se demander pourquoi les utilisateurs nous aiment ou ce que la raison est le sentiment négatif. En utilisant trois études qualitatives, ainsi que l'analyse sémantique latente (LSA) et la machine à vecteurs de support (SVM), les résultats montrent l'efficacité de la procédure d'analyse proposée du sentiment d'authenticité de la marque pour prédire les dimensions de l'authenticité de la marque et leur niveau de sentiment.

Le deuxième essai concerne la quantité immense de données et de localisation à base de destination à des fins publicitaires. Cela vient du fait que de plus en plus d'applications ont accès aux emplacements des utilisateurs et leurs destinations. Le contexte, économie du partage, est sélectionné pour tester si une application pour réserver une place, à savoir Uber, peut modifier une destination en leur fournissant des annonces adaptées aux destination pertinentes. Les niveaux de *construal* (à savoir la distance spatiale et la distance culturelle) montrent les effets sur les relations entre les attitudes envers la publicité basée sur la destination et la réponse à la promotion ainsi que l'intention de réutiliser l'application. Deux applications expérimentales qui simulent une application de covoiturage bien connu en Amérique du Nord sont utilisées. Les résultats fournissent des implications sur la théorie du niveau construal, la théorie du comportement planifié, et comment les praticiens peuvent l'utiliser pour modifier le comportement prévu, à savoir les destinations prévues.

Le troisième essai est également dans le contexte de l'économie de partage. S'appuyant sur la théorie des actes de langage, l'impact de la description du vendeur sur les cognitions sociales des consommateurs de services est étudiée pour trouver comment ils pourraient être utilisés pour générer de nouveaux contenus. Les résultats suggèrent le rôle de la concrétude linguistique, le type de fournisseur de services, et l'analyse des sentiments sur le niveau perçu de la chaleur et de la compétence des fournisseurs de services. L'utilisation de deux méthodes d'extraction de texte (analyse de fréquence et LDA modifiée), les différences entre les quatre types de descriptions de propriété et leurs hôtes sont expliquées. Ensuite, les résultats sont utilisés pour mieux former l'algorithme de génération de langage naturel (NLG). Cela montre comment la mémoire à long

court terme (LSTM), ainsi que les commentaires nettoyés, peuvent aider les fournisseurs de services à générer du contenu plus attrayant pour communiquer leur service.

Les résultats de la recherche sur chacun de domaines de recherche contribuent à plusieurs théories et à faciliter une enquête complémentaire sur les grandes analyses de données dans le marketing. Cette recherche fournit aussi aux praticiens du marketing des théories fiables et valides, des modèles et des systèmes d'aide à la décision afin de mieux comprendre et proposer des stratégies appropriées pour renforcer leur entreprise.

ACKNOWLEDGEMENT

Before starting my Ph.D. studies, I remember that I heard a lot that Ph.D. is a journey, but I was not sure what kind of journey it would look like. Long story short, I owe special thanks to so many amazing people in my life who helped me during all the ups and downs of this long journey. Here, I would like to appreciate each and every moment of their support.

First and foremost, I would like to express my deepest gratitude to Dr. Michel Laroche. I was very lucky to work under his supervision and have his unwavering support, invaluable mentorship, and steady belief in me. Without Dr. Laroche, none of this would have happened. Next, I would like to extend my sincere appreciation to my great advisors, Dr. Michèle Paulin, and Dr. Yu Ma. I am so very grateful for their insightful comments and suggestions as well as their treasured support during different stages of my Ph.D. study.

I would also like to thank my Ph.D. thesis examination committee, Dr. Chankon Kim, and Dr. Roberto G. de Almeida, who helped a lot in terms of improving the thesis. Moreover, I am very grateful to all the distinguished faculties in the Marketing department and in the joint program, specifically Dr. Christopher Ross, who I joined the department once he was leading it; Dr. Jooseop Lim and Dr. Darlene Walsh, the department heads who supported in so many different ways my Ph.D. studies, and the nice and supportive team including Mrs. Beverley Benderoff and Mr. George Spanos; Dr. Bianca Grohmann, who mentored me in the first few years and also I learned a lot from her in her retailing and branding classes; Dr. Zeynep Arsel, who also mentored me during the journey; Dr. Pierre-Yann Dolbec, who coached me for the job market; Dr. Emmanuelle Vaast, who ignited my passion for big data analytics in her class; Dr. Sylvain Sénécal, who helped me develop my second essay's idea in his Online Consumer Behavior class; Dr. Saurabh Mishra and Dr. Demetrios Vakratsas whom knowledge in Marketing Strategy and Modeling opened new research perspectives, and so much more amazing professors and teachers that I had from the early days I learned alphabet up to now.

Here I have to extend my great appreciation to the wonderful team in the Ph.D. office. I am very grateful to Dr. Kathleen Boies and Dr. Cedric Lesage for their exceptional leadership and support as well as to Ms. Norma Paradis and Ms. Dolores Mosquera for being so nice and caring.

Besides those awesome people who helped me at school, I would never forget the one at home whom without her tremendous encouragement, understanding, and patience, I could have not accomplished anything. My love of life, Hanyeh, thank you from bottom of my heart for everything you have done for Helma, Zahra, Hossein, and me. I would also like to extend my sincere thanks to my parents, my parents-in-law, my sister and her family, and my grandmother, who not only supported us by all means, long before this journey, but endured the sadness of being thousands of miles apart. They unconditionally loved and kept us in their thoughts and prayers.

This journey could have not come to an end without benefiting from so many nice colleagues, friends, and families who I was able to count day and night. In specific, I would like to thank Dr. Boris Bartikowski, Dr. Mohammd Reza Habibi, and Dr. Mahdi Tajeddin for all they have done, academically and personally, to make crossing the finish line happen.

Finally, I would like to appreciate the generous financial support and grants that I received from Concordia University, CIBC, Luc Beauregard center, SSHRC, and FRQSC.

DEDICATION

To our devoted parents, Mohammad, Amir, Zahra, and Mehri,

our beautiful little ones, Helma, Zehra, and Hossein,

and

my lovely wife, Hanyeh.

CONTRIBUTION OF AUTHORS

I am very much grateful for the contributions of my co-authors, Dr. Laroche, Dr. Bartikowski, and Dr. Richard, as well as my professors and colleagues, Dr. Vaast, Dr. Grohmann, Dr. Sénécal, and Dr. Habibi. I try my best to accurately explain their contributions.

All of these three papers have been conducted under the supervision of Dr. Laroche. He helped me a lot whenever I needed any inputs, comments, or suggestions. He also revised the manuscripts and made this research possible with his generous financial support.

I started the first paper once I took General Research Methods with Dr. Emmanuelle Vaast. The idea was developed more in the Branding course with the help of Dr. Bianca Grohmann. Dr. Richard helped in developing the scale for Brand authenticity. We also hired research assistances to code tweets and bring up its platform for the purpose of training the model.

I developed the idea for the second essay in Dr. Sylvain Sénécal's course on Online Consumer Marketing. In the second paper, Dr. Boris Bartikowski extensively contributed to developing theories and coming up with more focused research once we decided to take out app atmospheric cues from an earlier version. I learned a lot from him in data analysis and theory development, and his revisions and suggestions are incorporated in the text. Dr. Richard has also suggested mediators and moderators. Then, I came up with the current model, design, and analysis.

Regarding the third paper, I and Dr. Habibi talked about the idea of textual analysis with an Airbnb dataset. Then we explored it a bit, and Dr. Habibi suggested some potential constructs. Then I continued working on the paper and developed it thoroughly. Like the first paper, I am thankful for the help that I received from research assistances for the coding parts.

I am also thankful for anonymous reviewers, participants in the conferences, research seminars, and whoever contributed to this thesis.

Table of Contents

List of Figures

List of Tables

GENERAL INTRODUCTION

Right now, hundreds of thousands of new data are generated and added to the accumulated data. This immense amount of real-time and retrospective data has helped change previous paradigms and led to a new one: big data analytics. Big data became widespread as recently as 2011 and attracted many researchers and organizations from different disciplines not only due to its potentially broad range of applications but also because of its high involvement in the day-to-day life of people all over the world, including researchers and organizational stakeholders.

Not surprisingly, like other disciplines, marketing has been engaged in big data analytics to contribute to the knowledge of analyzing big data from its own perspective, and also to make itself able to propose more precise practical implications (Feldman, 2013). Nevertheless, the marketing intelligence literature is still far behind the big data mainstream and needs research in order to gain more insights for marketing practitioners. In today's age of big data (text, numbers, emoji, or video) traditional techniques are becoming obsolete (Sheth 2021). More theoretical and empirical research could help business managers stay one step ahead of the competition by obtaining and analyzing big data which could be converted into both short-term and long-term strategic planning in the new era (Wright and Calof, 2006).

From theory perspective, there are gaps in the existing literature of different marketing streams, i.e. branding, advertising, and communication. So, the first essay investigates brand management in social media due to the vast potential of social media as well as the companies' interests in utilizing it for branding purposes. More specifically, it examines the sentiments toward a brand, via brand authenticity, to identify both the reasons for positive or negative sentiments on social media and its polarity. Practically speaking, while firms need insights about users' sentiment towards their brand, knowing just that it is a simple positive or negative sentiment does not provide them with enough information. They might ask themselves why users like us or what the reason is behind the negative sentiment. Using three qualitative studies, along with the latent semantic analysis (LSA) and the support vector machine (SVM), the findings illustrate the effectiveness of the proposed procedure of brand authenticity sentiment analysis to predict both the brand authenticity dimensions (as the reason) and their level of sentiment (as the polarity).

The second essay is concerned with the immense amount of real-time locations available in mobile marketing and the possibility to exploit the data in location-based advertising, particularly in advertising based on destinations. The study helps in understanding how mobile devices are changing the way customers behave in the context of sharing-based services. Practically in the ridesharing context, the research question is how a ridesharing app like Uber can alter one's destination. In this research, we investigate if spatial distances (low: 50m vs. high: 3000m vs. high $_{\text{with free rerouting}}$ vs. high $_{\text{with non-free rerouting}}$) and cultural distances (high: international app vs. low: Canadian app) make impacts on consumer reactions. Using two experimental mock apps that mimic a well-known ridesharing app in Canada, the research finds strong main effects and interaction effects of DBA in three studies Interestingly, the findings show that users appreciate a culturally congruent app to alter their destination, even 3km away from their planned destination if they find the offer relevant and interesting. we discuss the finding implications for the theory of planned behavior and construal level theory and how practitioners can use it to alter planned behavior, i.e. planned destinations.

The third essay is also in the context of the sharing economy and investigates communication and content generation in the big date era. This research looks into how service-providers should communicate with customers through the textual descriptions to increase their performance level. It argues that the way service providers are communicating would lead to different levels of social cognitions shared via reviews. In this research, we posit that the level of professionalism of the service provider and the level of sentiment they use have also make impacts. To investigate these, we use deep learning methods, i.e. Natural Language Processing (NLP) and Long Short-Term Memory (LSTM) to investigate the existing Airbnb's property descriptions and users' reviews. Applying modified Latent Dirichlet Allocation (LDA), it highlights the related tokens in each of the four categories of service providers based on their linguistic concreteness and professionalism. In the next step, by using Natural Language Generation (NLG) and training the model with a cleaned dataset emerged from the different groups of service providers, we help hosts to generate appropriate content, and so customize their propriety description. The research not only contributes to more easily incorporate linguistic theories in marketing but also advances the literature in terms of understanding how textual communications influence user perceptions. It could also help develop more accurate NLG algorithms to communicate.

The research findings on each of these streams could contribute to several theories and facilitate further inquiries into big data analytics in marketing. Practically speaking, since companies' interest in monitoring and getting insights from big data continues to increase, the research also provides marketing practitioners with reliable and valid theories, models and decision support systems to gain insights and propose appropriate strategies to strengthen their firm. Finally, the thesis highlights the importance of big data ethics. Research at the intersection of business and computer science should look more at this important but usually neglected field of research. Big data's dark-side increasingly shows high potential to disrupt users' lives, companies' brand image, and society well-fare in so many ways. Understanding the dynamic nature of the challenge and its effects on consumer behavior provides novel knowledge for both theory and practice and helps firms implementing appropriate business strategies. Expanding research in these areas would grow the well-being of communities, brands, and consumers in this new big data era.

PAPER 1: USING BIG DATA ANALYTICS TO STUDY BRAND AUTHENTICITY SENTIMENTS: THE CASE OF STARBUCKS ON TWITTER

ABSTRACT

There is strong interest among academics and practitioners in studying branding issues in the big data era. In this article, we examine the sentiments towards a brand, via brand authenticity, to identify the reasons for positive or negative sentiments on social media. Moreover, in order to increase precision, we investigate sentiments polarity on a five-point scale. From a database containing 2,282,912 English tweets with the keyword 'Starbucks', we use a set of 2204 coded tweets both for brand authenticity and sentiment polarity. Firstly, we examine the tweets qualitatively to gain insights about brand authenticity sentiments. Then we analyze the data quantitatively to establish a framework in which we predict both the brand authenticity dimension and its sentiment polarity. Through three qualitative studies, we discuss several tweets from the dataset that classified under the *quality commitment, heritage*, *uniqueness,* and *symbolism* categories. Using latent semantic analysis (LSA), we extract the common words in each category. We verify the robustness of previous findings with an in-lab experiment. Results from the support vector machine (SVM), as the quantitative research method, analyses illustrate the effectiveness of the proposed procedure of brand authenticity sentiment analysis. It shows high accuracy for both the brand authenticity dimensions' prediction and its sentiment polarity. We then discuss the theoretical and managerial implications of the studies.

INTRODUCTION

Social networks (e.g. Facebook), microblogs (e.g. Twitter and Tumblr), blogs (e.g. Blogger and WordPress), social bookmarking (e.g. Delicious and StumbleUpon), and review sites (e.g. Epinions.com, Yelp, TripAdvisor) are considered very important in the big data era (Barbier and Liu, 2011; Gandomi and Haider, 2015). Each of these platforms has many users and fans; however, among these platforms, people all around the world are more engaged with social media, more specifically Twitter and Facebook. Companies invest heavily in developing a social media community not only to strengthen customer–firm relationships, but also to increase the firms' revenues and profits (Kumar *et al.* 2016). In recent years, due to the vast potential of social media, and the companies' interests in utilizing it for branding purposes, social media studies have garnered much attention in the branding and information management literatures (Gensler *et al.* 2013; Habibi, Laroche, and Richard, 2014; Laroche, Habibi, and Richard, 2013; Mahrt and Scharkow, 2013; Naylor, Lamberton, and West, 2012). Today, brands could easily be affected if there is any mismatch between consumer expectations and product characteristics from a variety of sources including social media (Hofacker, Malthouse, and Sultan, 2016). This requires firms to constantly monitor brand health, compare it with that of their competitors, and periodically examine customer mindset measures to guide marketing decisions (Ailawadi, Lehmann, and Neslin, 2003).

This research stream contributed to the literature in terms of illustrating the possibility of analyzing branding issues in social media. However, the literature on brand studies on social media is still limited with regards to brand sentiment analysis. Indeed, most of the existing research studied brand sentiments with a three-point scale (positive, negative and neutral). This simple scale is not able to provide more precise information about the polarity of positive or negative attitudes towards a brand (Hu, Koh, and Reddy, 2014). For instance, it could not show differences between those who like/dislike a brand with those who love/hate a brand. Moreover, it does not provide any insight on the reasons for either a positive, negative, or a neutral sentiment towards a brand. Practically speaking, with these kinds of metrics, a brand manager could not determine which brand characteristics lead to better or worse sentiments. Thus, the analysis should go beyond positive or negative classifications and provide clearer evidence and explanations (Gaspar *et al.* 2016).

In this article, given the necessity of monitoring the perceived value of brand authenticity, to protect a popular brand against the heartbreak of *genericide* (Walsh, 2013), we investigate sentiments towards brand authenticity on Twitter with a five-point scale. This would fill the current gaps in the literature and contribute both in terms of better precision, and of providing firms with valuable insights about the way people interact with their brands. We investigate brand authenticity sentiments both qualitatively and quantitatively. After establishing the descriptions of associated items of brand authenticity dimensions, in study 1, we report and discuss the coded tweets for each of the brand authenticity dimensions. In study 2, we use latent semantic analysis (LSA) to extract the common words in each of the categories. In study 3, we check the robustness of the coding process and the findings from study 1 and 2 with another source of data. Then, in study 4, we train and validate a model, using support vector machine (SVM) analyses, to automatically classify tweets in regard to both their sentiment polarities and their brand authenticity dimensions.

The research findings on brand authenticity sentiment analysis could facilitate further inquiries into sentiment analysis in branding contexts and also in several related domains, such as e-Word-of-Mouth studies. This research also provides marketing practitioners with a reliable and valid instrument to evaluate the level of sentiment towards a brand more specifically, and propose appropriate strategies to strengthen it.

LITERATURE REVIEW

Brand sentiment analysis

There is a growing interest among marketing researchers in studying branding issues on social media. On one hand, more and more companies engage with social media, carefully broadcast sentiments to entertain consumers, and promote brands (Gopaldas, 2014). On the other hand, customers themselves publicly share what they feel and how they evaluate different specifications of brands on social media, forums, and websites. Huberty (2015) suggested that in forecasting consumer behavior, it is reasonable to assume a relatively stable link between online and offline attitudes. In order to provide more personalized offers for customers, marketers should know the users' emotional states toward different aspects of the brand (Ortigosa, Martín, and Carro, 2014). So far, brand managers, both those who have formally engaged with social media and those who have not, are very interested in getting insights about the effectiveness of their branding campaigns through social media contexts.

Arguing that more and more consumers rely on online contents when they want to get information about brands, He *et al.* (2015) focused on brand sentiments and proposed a social media competitive analytics framework. Fuchs, Höpken, and Lexhagen (2014) studied knowledge generation in the tourism field based on customers searching, booking, and also providing feedback in websites and social media. Similarly, Xiang *et al.* (2015) explored the utility of a big database (www.expedia.com reviews) to better understand the relationship between the hotel guest experiences and satisfaction. Smith, Fischer, and Yongjian (2012) illustrated that the brand sentiments of user-generated content is not predictably different across social media platforms (Facebook, Twitter, and YouTube). Lee and Bradlow (2011) justified an automated marketing research model to uncover the customer voice, using six years of online customer reviews for digital cameras. Extending existing knowledge about brand sentiments, we examine the sentiments towards a brand via brand authenticity to address the reasons behind positive or negative sentiments. Without this theoretical basis, one could not provide evidence to address questions about different sentiments. In other words, the brand authenticity could help brand managers to address why some customers love their brand or hate it. Next, we discuss the brand authenticity concept, its associated dimensions and scale.

Brand authenticity concept

According to the Oxford Advanced Learner's *Dictionary*, the word 'authentic' comes from the Greek *"authentikos,"* meaning "principal" and "genuine." The dictionary has also provided the following three definitions for "authentic": 'known to be real and genuine and not a copy', 'true and accurate', and 'made to be exactly the same as the original'.

In the literature, there is little congruency among the proposed definitions, leading to interpreting authenticity in different ways (Choi *et al.* 2015). Beverland and Farrelly (2010) invited researchers to view authenticity as "a socially constructed interpretation of the essence of what is observed rather than properties inherent in an object" (p. 839). Gathering these aspects together, Beverland (2005) suggested that "brand authenticity can be inherent in an object, come from a relation between an object and/or a historical period, an organization form, or nature, or be given to an object by marketers and consumers. Authenticity can also be true and/or contrived" (p. 1006).

Although authors defined authenticity in different ways, the literature is unanimous regarding its significant effects and advantages in marketing and branding. Eggers *et al.* (2013) established the linkages among brand authenticity, brand trust, and SME growth from a CEO perspective. Assiouras *et al.* (2015) found that brand authenticity predicts brand attachment, while brand attachment influences consumer purchase intentions, willingness to pay more, and to promote the brand. Kadirov (2015) focused on the perceived authenticity gap between national brands and private labels to explore whether and how this factor influences the effect of marketing and manufacturing variables on willingness to pay. Johnson, Thomson, and Jeffrey (2015) believed that if consumers judge brands to be less authentic, the brand is considered to be of lower quality, less socially responsible, and they are less likely to join the corresponding brand community. Arguing that what customers want are memorable experiences rather than products, Gilmore and Pine (2007) suggested that the success of brands, such as Starbucks, no longer depends on its operational prowess or taste superiority; it relies solely on sustaining coffee drinkers' perceptions of the Starbucks experience as being authentic (p. 2).

Brand authenticity dimensions

According to Fournier and Avery (2011), brands have a strategy of openness in social media to establish their authenticity. This strategy led to an increased attention towards the factors, which brings perceived authenticity for customers. Here, we conceptualize four brand authenticity dimensions, and then we use exploratory and confirmatory factor analysis (EFA and CFA) to establish them.

First, Gilmore and Pine (2007) believed that while customers previously perceive low quality products as *junks*, today they do not tolerate products with poor quality and call them *fakes* (p. 2). As such, Napoli *et al.* (2014) established brands' quality commitment as an important factor of brand authenticity. The authors suggest that, i.e. producing to the most exacting standards and making the products by a master craftsman would be signals for customers to perceive *quality commitments*, and therefore authenticity for a brand. As such, we use *quality commitments* as the first dimension for brand authenticity.

Second, brands may have connections to particular places, times and also specific methods of production, designs, and styles, which reflect their concrete referents and cultural associations (Spiggle, Nguyen, and Caravella, 2012). Being called *heritage* in the literature, it could be achieved by using marketing-mix variables that invoke the history of a particular brand, including all its personal and cultural associations (Brown, Kozinets, and Sherry, 2003). Moreover, previous research suggests that through building links to cultural events and also drawing on historical and past events associations, brands could be perceived authentic (Beverland 2005). Therefore, in this research, *heritage* is the second corresponding factor for brand authenticity.

Third, customers may believe a brand is different from competing brands. *Uniqueness* refers to the extent to which customers feel the relative distinction between a brand and its competitors (Netemeyer *et al.* 2004). Lewis, and Bridger (2001) suggested that consumers, by emphasizing on brand authenticity, and even when their purchases are not the same, still expect unity in terms of product uniqueness and originality. Schallehn, Burmann, and Riley (2014) found brand individuality, the unique way in which a brand fulfills its promise, as a factor of brand authenticity. Accordingly, we expect that *uniqueness* also forms the brand authenticity construct.

Finally, *symbolism* reflects the symbolic quality of a brand that define consumers who they are or who they are not. Morhart *et al.* (2015) found that authentic brands reflect values that customers consider important and may thus help shape who they are. Napoli *et al.* (2014) expected that symbolism should also form part of the brand authenticity construct, and called for further research. Hence, we select *symbolism* as the fourth dimension of the brand authenticity construct.

In summary, motivated by Napoli *et al.* (2014)'s call for further studies about the factors of brand authenticity, we conducted factor analysis (both exploratory and confirmatory) to establish the brand authenticity's factorial model, using adapted items from the existing literature namely: *quality commitment* (9 items) and *heritage* (7 items), *uniqueness* (3 items), and *symbolism* (6 items). Before going through the four main brand authenticity sentiment studies on Twitter, here we present the EFA and CFA study, and its results.

Brand authenticity factorial model

Two hundred undergraduate students (51% males) in exchange for extra course credit participated in the study. Participants rated their level of agreement with 25 items on a seven-point scale about the Starbucks brand (1 = strongly disagree, 7 = strongly agree). Previous research shows different level of perceived authenticity among Starbucks customers (Thompson, Rindfleisch, and Arsel, 2006). We removed the data from participants who indicated they were unfamiliar, not knowledgeable, or inexperienced with the Starbucks (with a mean score of less than 2 on the five-point scale: "Please indicate your level of experience with Starbucks": strongly unfamiliar/ strongly familiar, not knowledgeable at all/strongly knowledgeable, strongly inexperienced/ strongly experienced, $\alpha = .89$), resulting in a final sample size of 188.

Using principal component EFA with oblimin rotation, a four-factor model with eigenvalues greater than one obtained, *symbolism* (8.75), *quality commitment* (1.93), *uniqueness* (1.26), and *heritage* (1.17). The scree plot examination also confirmed the existence of four major factors. The four factors explained 62.43% of the total variance, 41.65, 9.21, 6.02, and 5.55 respectively, which is in the acceptable range (Hinkin 1998). We removed three items with low factor loadings on their main factors (< .4) and one item with high cross-loading (> .2). Then, we conducted series of CFA to see which model fit the data best (See Table 1.1 for the details). Comparing the fit indices of different measurement models shows the presence of a four-factor correlated model ($\chi^2(183) = 372.91$, $p < .01$, $\chi^2/df = 2.04$, CFI = .945, NNFI = .938, GFI = .924, SRMR = .061, RMSEA = .072; Bollen 1989). As a result, we used this model and its corresponding items for further steps in brand authenticity sentiment analysis. Table 1.2 shows and defines the emerged items of each brand authenticity dimension.

Table 1.1: Fit indices results for different factorial models

Model Number	Model name	χ^2	df	P	χ^2/df	CFI	NNFI	RMSEA	SRMR
0	Null-model	*	*	*	*	NA	NA	NA	NA
1	One-factor (Q-H-U-S)	673.73	189	< .01	3.56	.735	.706	.117	.085
2a	Two-factor uncorrelated (Q-U and H-S)	194.24	54	< .01	3.60	.855	.823	.118	.220
2b	Two-factor correlated (Q-U and H-S)	132.99	53	< .01	2.51	.917	.897	.090	.064
3	Three-factor correlated (Q, U, H-S)	184.80	87	< .01	2.12	.920	.903	.078	.062
4a	Four-factor uncorrelated (Q, U, H, S)	646.00	189	< .01	3.42	.750	.723	.114	.286
4b	Four-factor correlated (Q, U, H, S)	372.91	183	< .01	2.04	.945	.938	.072	.061

Q= Quality commitment, U= Uniqueness, H= Heritage, and S= Symbolism. NA= not applicable
Model 1: All items forced to load on one factor. Model 2a and 2b: Due to some similarities between Q and U, and also H and S, we forced the items of these four dimensions to load on two, the first one made from Q and U, and the second one from H and S. Model 3: Due to some similarities between H and S, we forced the items of these two under one factor and items of Q and U on two other factors. Model 4a and 4b: Items of each dimension loaded on its own factor.

Table 1.2: Brand authenticity items and dimensions

Items	Dimension	Adapted from
Only the finest ingredients/materials are used in the manufacture of Starbucks's products.	Quality commitment	Napoli *et al.* (2014)
Quality is central to the Starbucks.		
Starbucks's product is made to the most exacting standards, where everything is aimed at improving quality.		
Starbucks' products are manufactured to the most stringent quality standards.		
Artisan skills and customized manufacturing processes used in the production of Starbucks' products.		
Starbucks' products are made by a master craftsman who pays attention to detail and is involved throughout the production process.		
Starbucks has strong connections to a historical time period, a culture and/or a specific region.	Heritage	Napoli *et al.* (2014)
Starbucks has a strong link to the past, which is still perpetuated and celebrated to this day.		
Starbucks reminds me of a golden age.		
Starbucks exudes a sense of tradition.		
Starbucks reinforces and builds on long-held traditions.		
I miss or remember a regular habit that I had in the past with Starbucks.		
The way Starbucks fulfills its brand promise is very different from competing brands.	Uniqueness	Schallehn *et al.* (2014)
The way Starbucks fulfills its brand promise is unique.		
Starbucks fulfills its brand promise in a distinct way.		
Starbucks is a brand that adds meaning to people's lives.	Symbolism	Morhart *et al.* (2015)
Starbucks is a brand that reflects important values that people care about.		
Starbucks is a brand that connects people with their real selves.		
Starbucks is a brand that connects people with what is really important.		
Starbucks is a brand that cares about protecting people's identity.		
Starbucks is a brand that fulfills my life in providing my favorite product.		

RESEARCH DESIGN
Twitter as the research platform

To reach our research goals, the brand authenticity sentiments were analyzed from contents generated on Twitter, which is one of the most popular social media platforms. In 2006, two years after Facebook, Twitter introduced itself as a microblogging social media platform and recently surpassed 500 million tweets per day on average (Internet live stats, 2015). On this platform, users are able to share publicly with their followers on a variety of devices up to 140-character texts (tweets). This limited amount of characters leads users to express their updates (thoughts, news, emotions and so on) in smaller phrases, which could help content analyzing researchers deal with lesser amounts of unnecessary information (Milstein *et al.* 2008).

Starbucks Coffee as the brand

Given that our research is about brand related issues, and that we are going to illustrate the practical effectiveness of our proposed procedure, we had to select a brand for this study. Considering the *Business Week*'s Top 100 Brands and the American Customer Satisfaction Index, and taking into account statistics regarding the most popular brands on Twitter (Social bakers, 2015), we selected Starbucks Coffee.

The Starbucks Corporation, commonly known as Starbucks Coffee, opened its first coffeehouse in Seattle (US) in 1971. Today, it owns a vast coffeehouse chain in many countries (21,536 stores in 64 countries) and it has become the largest coffeehouse company in the world ahead of its UK rival, Costa Coffee (Wikipedia, 2015). Starbucks Coffee joined Twitter in 2006, and at the time of our research, was the second most popular brand on Twitter with 8.3 million followers. It received replies (comments), retweets (forwarding), favorites (likes), updates (tweets) and mentions (listed) much more than even the first popular brand, Samsung Mobile (Twittercounter, 2015). Previous research found Starbucks' Twitter account as a place for a combination of customer testimony, complaining, feedback, and QandA (Jansen *et al.* 2009). This means that people communicate with Starbucks Coffee more than with other brands. In other words, users use social media to express feelings and attitudes towards this brand (Lee, Han, and Suh, 2014), which is exactly the main focus of our research.

Data collection

There are two methods for gathering the required data from Twitter. The first, which is free of charge, is provided by Twitter (Twitter Application Programming Interface (API) version 1.0 and recently version 1.1). Through Twitter's streaming API, one can obtain real-time access to tweets, replies, and mentions created by public accounts in a sampled and filtered form (Bifet and Frank, 2010). Driscoll and Walker (2014) suggested using the 'fire hose' of other data providers (e.g. Gnip PowerTrack) as the second method of tweet collection, and illustrated some advantages compared with tweets provided APIs, mostly in terms of tweets volume.

Although authors (e.g. Boyd and Crawford, 2012) expressed some uncertainty about how tweets are provided through Twitter API, and introduced *rich big data* and *poor big data* terms as a symbol of the differences between those who have access to more and fewer data in the new era, Twitter API is commonly used by scholars and no one reported unexpected results due to using this method (Shi, Rui, and Whinston, 2013). Thus, we collected stream tweets with the 'Starbucks' keyword through Twitter API version 1.1 fire hose, using the Python programming language. We collected the data for two months, starting on July 20, 2015 at 10:00:00 p.m. (GMT) to Sep 20, 2015 at 9:59:59 p.m. (GMT). We obtained 2,988,560 tweets, which contain the keyword *Starbucks*, approximately 50000 per day. The data was messy but absolutely rich, including information regarding its created time, sender's name, ID, location, language, number of followers, number of friends, number of favorites, number of statuses if the tweet is in reply to others, who sent the tweet if it was a retweet, when the sender joined Twitter and so on. Figure 1.1 shows a sample of received raw data for the following two tweets:

- *"I've always loved and will remain loving Starbucks!"*
- *"@Starbucks the worker was very rude too yikes I'm not going back to that one store".*

First, we created a query to filter the tweets based on its language to remove non-English ones. That led to the total number of 2,282,912 English tweets for further analyses. As we were interested in examining user sentiments, we ignored tweets by Starbucks accounts such as @Starbucks, @Starbucks – OA, @Starbucks Arizona, @Starbucks Australia, @Starbucks Help, @Starbucks Partners, @Starbucks UK, @Starbucks view, @Starbucks Belleville, @Starbucks Rosental. Removing these tweets from the database gave us 2,242,231 tweets. From the final dataset, we randomly chose 3000 tweets, representative of the whole period, for our qualitative and quantitative studies. Next, we provide details regarding the proposed procedure for brand authenticity sentiment analysis.

Brand authenticity sentiment analysis procedure: coding algorithm

While some authors used a prepared lexicon (e.g. Jansen *et al.* 2009), we did not exploit these available sources because they are not able to properly address our research objectives, due to its novelty in terms of studying brand authenticity sentiments rather than just brand sentiments in general, and also in terms of the use of the five-point scale rather than a three-point one. Thus, two graduate students evaluated individually the tweets through the following algorithm, step by step and tweet by tweet.

First, they judged sentiments towards Starbucks on a five-point scale: 2= strongly positive, 1= positive, 0= neutral, -1= negative, -2= strongly negative. This evaluation helps us have an overall judgment about the brand sentiment before investigating authenticity sentiment dimensions.

Second, following a detailed guideline on how items on each dimension of brand authenticity (*quality commitment*, *heritage*, *uniqueness*, and *symbolism*) should be captured from the tweets, they categorized each tweet under the proper dimension among the four dimensions of brand authenticity. The detailed guideline was written based on the items and descriptions of each brand authenticity dimension, previously introduced in the Table 1.2. There were three more options for tweets which could not be classified under the four dimensions. For tweets about brand authenticity but not classifiable under one of the above mentioned dimensions, we created a fifth category, n*one of them*. The sixth category was for tweets, which did not provide sufficient information to judge the users' brand authenticity sentiments. Also, for those tweets which were irrelevant to brand authenticity, we created another category.

****
{"created_at":"Tue Jul 21 21:46:25 +0000 2015" id:623610002524274689 id_str:"623610002524274689" text:"I've always loved & will remain loving Starbucks! \ud83d\ude18
https:\/\/t.co\/iguETV1nz1" source:"\u003ca href=\"http:\/\/twitter.com\/download\/iphone\" rel=\"nofollow\"\u003eTwitter for iPhone\u003c\/a\u003e" truncated:false in_reply_to_status_id:null
in_reply_to_status_id_str:null in_reply_to_user_id:null in_reply_to_user_id_str:null in_reply_to_screen_name:null user:{"id":860191964 id_str:"860191964" name:"Anthony Goehring "
screen_name:"AnthonyGoehring" location:"" url:null description:null protected:false verified:false followers_count:136 friends_count:250 listed_count:0 favourites_count:184
statuses_count:1371 created_at:"Wed Oct 03 19:05:36 +0000 2012" utc_offset:null time_zone:null geo_enabled:false lang:"en" contributors_enabled:false is_translator:false
profile_background_color:"C0DEED" profile_background_image_url:"http:\/\/abs.twimg.com\/images\/themes\/theme1\/bg.png"
profile_background_image_url_https:"https:\/\/abs.twimg.com\/images\/themes\/theme1\/bg.png" profile_background_tile:false profile_link_color:"0084B4"
profile_sidebar_border_color:"C0DEED" profile_sidebar_fill_color:"DDEEF6" profile_text_color:"333333" profile_use_background_image:true
profile_image_url:"http:\/\/pbs.twimg.com\/profile_images\/622562126650023937\/q513jFLQ_normal.jpg"
profile_image_url_https:"https:\/\/pbs.twimg.com\/profile_images\/622562126650023937\/q513jFLQ_normal.jpg"
profile_banner_url:"https:\/\/pbs.twimg.com\/profile_banners\/860191964\/1437265922" default_profile:true default_profile_image:false following:null follow_request_sent:null
notifications:null} geo:null coordinates:null place:null contributors:null quoted_status_id:622127994199515136 quoted_status_id_str:"622127994199515136" quoted_status:{"created_at":"Fri Jul
17 19:37:27 +0000 2015" id:622127994199515136 id_str:"622127994199515136" text:"Top 6 Mind-Blowing Starbucks Facts You Never Knew http:\/\/t.co\/VYmOvfhUSJ
http:\/\/t.co\/PAvhcPM1S3" source:"\u003ca href=\"https:\/\/about.twitter.com\/products\/tweetdeck\" rel=\"nofollow\"\u003eTweetDeck\u003c\/a\u003e" truncated:false in_reply_to_status_id:null
in_reply_to_status_id_str:null in_reply_to_user_id:null in_reply_to_user_id_str:null in_reply_to_screen_name:null user:{"id":3070181257 id_str:"3070181257" name:"Alex"
screen_name:"Alex122131" location:"" url:null description:"Fun Guy with Big Dreams." protected:false verified:false followers_count:2763 friends_count:0 listed_count:1 favourites_count:0
statuses_count:4 created_at:"Mon Mar 09 18:16:16 +0000 2015" utc_offset:19800 time_zone:"Chennai" geo_enabled:false lang:"en" contributors_enabled:false is_translator:false
profile_background_color:"C0DEED" profile_background_image_url:"http:\/\/abs.twimg.com\/images\/themes\/theme1\/bg.png"
profile_background_image_url_https:"https:\/\/abs.twimg.com\/images\/themes\/theme1\/bg.png" profile_background_tile:false profile_link_color:"0084B4"
profile_sidebar_border_color:"C0DEED" profile_sidebar_fill_color:"DDEEF6" profile_text_color:"333333" profile_use_background_image:true
profile_image_url:"http:\/\/pbs.twimg.com\/profile_images\/581425659240558592\/OcoDssCq_normal.jpg"
profile_image_url_https:"https:\/\/pbs.twimg.com\/profile_images\/581425659240558592\/OcoDssCq_normal.jpg" default_profile:true default_profile_image:false following:null
follow_request_sent:null notifications:null} geo:null coordinates:null place:null contributors:null retweet_count:0 favorite_count:0 entities:{"hashtags":[] trends:[]
urls:[{"url":"http:\/\/t.co\/VYmOvfhUSJ" expanded_url:"http:\/\/bit.ly\/1TpBeVO" display_url:"bit.ly\/1TpBeVO" indices:[50 72]}] user_mentions:[] symbols:[]
media:[{"id":622127990718099457 id_str:"622127990718099457" indices:[73 95] media_url:"http:\/\/pbs.twimg.com\/media\/CKI-HIOUcAEfMPA.jpg"
media_url_https:"https:\/\/pbs.twimg.com\/media\/CKI-HIOUcAEfMPA.jpg" url:"http:\/\/t.co\/PAvhcPM1S3" display_url:"pic.twitter.com\/PAvhcPM1S3"
expanded_url:"http:\/\/twitter.com\/Alex122131\/status\/622127994199515136\/photo\/1" type:"photo" sizes:{"medium":{"w":600 h:333 resize:"fit"} small:{"w":340 h:189 resize:"fit"}
large:{"w":1009 h:562 resize:"fit"} thumb:{"w":150 h:150 resize:"crop"}}}]} extended_entities:{"media":[{"id":622127990718099457 id_str:"622127990718099457" indices:[73 95]
media_url:"http:\/\/pbs.twimg.com\/media\/CKI-HIOUcAEfMPA.jpg" media_url_https:"https:\/\/pbs.twimg.com\/media\/CKI-HIOUcAEfMPA.jpg" url:"http:\/\/t.co\/PAvhcPM1S3"
display_url:"pic.twitter.com\/PAvhcPM1S3" expanded_url:"http:\/\/twitter.com\/Alex122131\/status\/622127994199515136\/photo\/1" type:"photo" sizes:{"medium":{"w":600 h:333 resize:"fit"}
small:{"w":340 h:189 resize:"fit"} large:{"w":1009 h:562 resize:"fit"} thumb:{"w":150 h:150 resize:"crop"}}}]} favorited:false retweeted:false possibly_sensitive:false filter_level:"low"
lang:"en"} retweet_count:0 favorite_count:0 entities:{"hashtags":[] trends:[] urls:[{"url":"https:\/\/t.co\/iguETV1nz1"
expanded_url:"https:\/\/twitter.com\/alex122131\/status\/622127994199515136" display_url:"twitter.com\/alex122131\/sta\u2026" indices:[56 79]}] user_mentions:[] symbols:[]} favorited:false
retweeted:false possibly_sensitive:false filter_level:"low" lang:"en" timestamp_ms:"1437515185372"}
****
{"created_at":"Tue Jul 21 21:48:02 +0000 2015" id:623610407870029824 id_str:"623610407870029824" text:"@Starbucks the worker was very rude too yikes im not going back to that one store"
source:"\u003ca href=\"http:\/\/twitter.com\/download\/iphone\" rel=\"nofollow\"\u003eTwitter for iPhone\u003c\/a\u003e" truncated:false in_reply_to_status_id:623609658159140864
in_reply_to_status_id_str:"623609658159140864" in_reply_to_user_id:1535183485 in_reply_to_user_id_str:"1535183485" in_reply_to_screen_name:"shiningdeming" user:{"id":1535183485
id_str:"1535183485" name:"cindy" screen_name:"shiningdeming" location:"dtx" url:null description:"i like my eyelashes as dark as my soul and as long as the list of people i hate" protected:false
verified:false followers_count:521 friends_count:771 listed_count:2 favourites_count:4853 statuses_count:12511 created_at:"Thu Jun 20 22:22:58 +0000 2013" utc_offset:-18000
time_zone:"Central Time (US & Canada)" geo_enabled:false lang:"en" contributors_enabled:false is_translator:false profile_background_color:"080108"
profile_background_image_url:"http:\/\/pbs.twimg.com\/profile_background_images\/497853840369016832\/Mp03vhCE.jpeg"
profile_background_image_url_https:"https:\/\/pbs.twimg.com\/profile_background_images\/497853840369016832\/Mp03vhCE.jpeg" profile_background_tile:true profile_link_color:"7AD0D6"
profile_sidebar_border_color:"000000" profile_sidebar_fill_color:"DDEEF6" profile_text_color:"333333" profile_use_background_image:true
profile_image_url:"http:\/\/pbs.twimg.com\/profile_images\/622625867278086144\/2mjdkJUA_normal.jpg"
profile_image_url_https:"https:\/\/pbs.twimg.com\/profile_images\/622625867278086144\/2mjdkJUA_normal.jpg"
profile_banner_url:"https:\/\/pbs.twimg.com\/profile_banners\/1535183485\/1436332553" default_profile:false default_profile_image:false following:null follow_request_sent:null
notifications:null} geo:null coordinates:null place:null contributors:null retweet_count:0 favorite_count:0 entities:{"hashtags":[] trends:[] urls:[] user_mentions:[{"screen_name":"Starbucks"
name:"Starbucks Coffee" id:30973 id_str:"30973" indices:[0 10]}] symbols:[]} favorited:false retweeted:false possibly_sensitive:false filter_level:"low" lang:"en"
timestamp_ms:"1437515282014"}
****

Figure 1.1: Raw data for two tweets

Third, based on the n-gram analysis technique (Sidorov *et al.* 2014), evaluators detected the most meaningful uni-gram (i.e. #*amazing*), bi-gram (i.e. #*Starbucks_addiction*), and tri-gram (#*do_not_like*) and assessed its brand authenticity sentiments towards Starbucks Coffee on a five-point scale (2= strongly positive, 1= positive, 0= neutral, -1= negative, -2= strongly negative).

Finally, based on the provided guidelines regarding the specifications of each dimension of brand authenticity and the other three categories, evaluators classified each tweet under one of the seven categories.

Figure 1.2 shows the user-friendly PHP platform in which the coding process took place. One of the authors trained the coders how to work with the platform and continuously monitored their tasks. In the platform, evaluators read the tweet, decided on its overall sentiments about the brand, chose its class among the seven categories, evaluated its brand authenticity sentiments, and picked its most meaningful unigram, bigram, and trigram from the database just by typing its initials. Evaluations which were not unanimous between the evaluators were discarded, resulting in 2204 tweets incorporated into the final dataset, and showing a high level of inter-rater agreement, 0.73. To verify the coding process, one of the authors checked the final coded data, and found the evaluations completely in accordance with the provided guidelines. Our analysis of brand authenticity sentiment consists of two phases, the qualitative and the quantitative ones.

In the next section, we provide details for the steps of these two phases.

Qualitative studies

In order to analyze brand authenticity sentiments, we conducted three qualitative studies. First, we explored the dataset to ensure the reliability and validity of the classification and coding process. During this step, as we noticed a great passion towards Pumpkin spice latte (PSL) among users, we decided to focus the study 1 on the first week of Fall 2015 (September 1 to September 8), the days around the time Starbucks returned that popular product to the menu. Previous research also found strong sentiment towards PSL but didn't investigate its reasons (Ghiassi, Zimbra, and Lee, 2016). This qualitative study helped us go beyond common sentiment analysis by providing many more insights about the brand authenticity sentiments. We discussed all tweets of this time period, whether they are related to PSL or not.

Second, in order to establish the key common themes of the tweets, to support the coding process and also the study 1 analysis, we approached the whole data with a text mining tool. we used latent semantic analysis (LSA) (Deerwester *et al.* 1990) to determine associated words in each of the seven categories. Evangelopoulos (2011) compared latent dirichlet allocation (LDA) and LSA, and found LSA results more accurate.

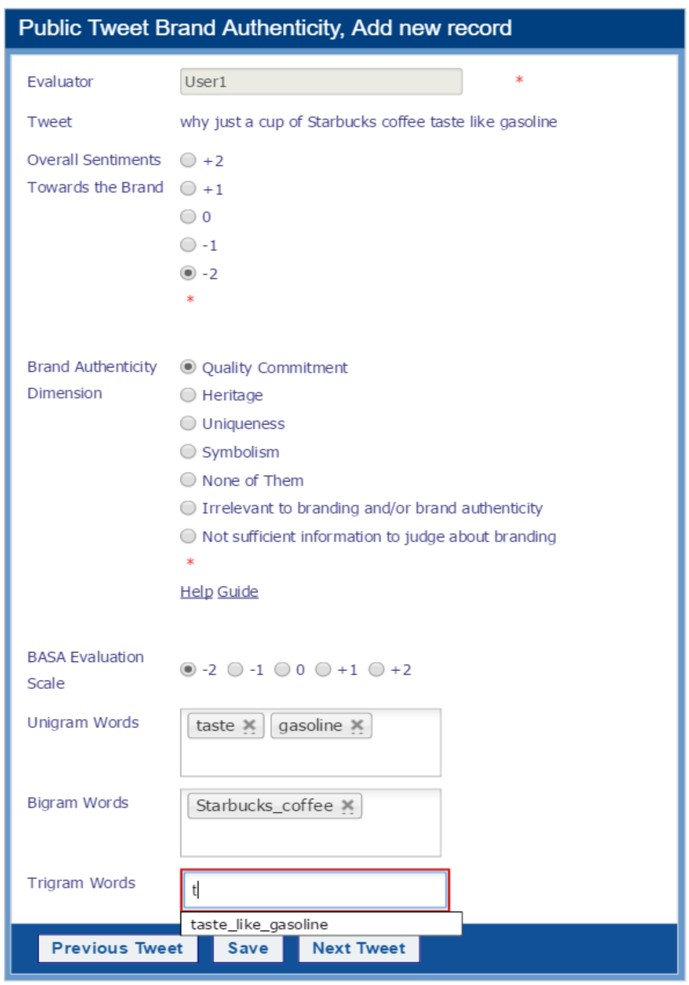


Figure 1.2. The platform for coding process of brand authenticity sentiment analysis

LSA, instead of counting the frequency of words, co-occurrences, or simple correlations in usage, creates a new semantic space where deeper relations among words/documents are inferred (Ahmad and Laroche, 2017). Once we conducted our LSA, we set the singular value decomposition (SVD) to seven. We report and discuss the extracted words in each category in the results and discussion section.

As the last step in the qualitative phase, we validated our findings with data from a laboratory setting. One hundred eighty undergraduate students participated in study 3 in exchange for having a chance of winning one of six Amazon's $25 gift cards. At the beginning of each of the four sessions, one of the authors asked the participants to think about their own experience with Starbucks, and answer to each of the following four questions about the brand. Then, they saw following scenario on the screen:

"Suppose you want to share your sentiments and attitudes (either very positive, positive, negative, very negative or neutral) towards Starbucks on social media (i.e. Twitter, Facebook, Instagram, or so on). What are the words and/or phrases, which you might use about each of the following situations? (There is no right or wrong answer! Also, do not worry about possible spelling errors!)

- *You want to share something about Starbucks's product quality:*
- *You want to share something about Starbucks's products, values, behaviors, and so on that adds meaning to your life:*
- *You want to share something about Starbucks's ordinary/unique characteristics:*
- *You want to share something about Starbucks's connection to a historical period in time, culture and/or specific region:"*

We also asked demographic questions along with questions regarding their level of social media usage, and the social media platforms, which they have account, use at least once a day, and at least once a week. Finally, they thanked and debriefed. We present and discuss the findings from this study in the results and discussion section.

Quantitative study

In the quantitative study, and for the last step in the brand authenticity sentiment analysis, we used a machine learning algorithm to analyze the efficiency and accuracy of features, which were created in the coding process. The object in this phase was to automatically categorize the tweets in accordance with the sentiment polarity, which has five possibilities, and the brand authenticity which has seven categories. In fact, the dependent variable in this phase was the class designation of each tweet among the possible five categories for brand sentiment polarity, and the seven options for brand authenticity dimensions. To this end, we used support vector machine (SVM) analysis (Cortes and Vapnik 1995). SVM is recognized as a fundamental algorithm for classification and regression problems (Chen and Zhang, 2014). Findings from different fields, specifically in sentiment analysis, (Marafino *et al.* 2014; Mostafa, 2013; Sidorov *et al.* 2014) suggest its usefulness and acceptable results accuracy. Once SVM is trained by the training data set, it can map a new data point into a space, which it belongs to, and classify into a specific category (Chan, Li, and Zhu 2015).

Practically, it finds hyperplanes that separate the classes with the largest margins. In the simplest case, suppose the training data set is $\{x_i, y_i\}$, i=1, 2, 3, …, l, $y_i \in \{-1, +1\}$, $x_i \in \mathrm{R}^d$, and we have the following hyperplane, which divides the positive from the negative examples

$$\vec{w}.\vec{x} + b = 0 \qquad (1)$$

where the $\vec{x}$ are those which lie on the hyperplane, $\vec{w}$ is normal to the hyperplane, $|b|/||\vec{w}||$ is the perpendicular distance from the hyperplane to the origin, and $||\vec{w}||$ is the Euclidean norm of $\vec{w}$ (Burges 1998). As depicted in Figure 1.3, H_1 represents all vectors that fit in the $\vec{x}_i \cdot \vec{w} + b = +1$ and H_2 shows all vectors in the $\vec{x}_i \cdot \vec{w} + b = -1$ equation. All data points should fall outside of these hyperplanes, and not inside the margin. In other words, the data points should satisfy the following inequality:

$$y_i (\vec{x}_i \cdot \vec{w} + b) - 1 \geq 0 \ \forall i \qquad (2)$$

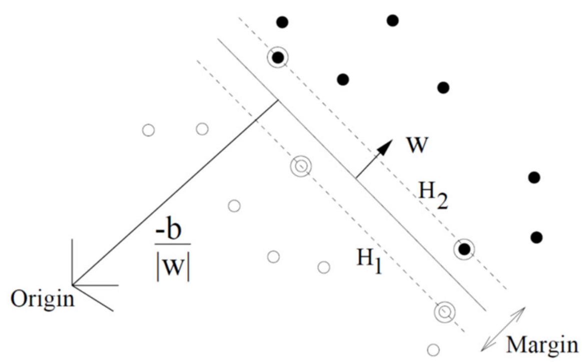


Figure 1.3: An example of support vector machine classification. The support vectors are circled (Burges 1998)

The SVM job is to maximize the margin. The distance of positive vectors (the black ones in the Figure 1.3), from the origin is $\frac{|1-b|}{||\vec{w}||}$. Similarly, the distance of negative vectors (the white ones), from the origin is $\frac{|-1-b|}{||\vec{w}||}$. So, the distance between the positive and negative pair vectors, which is the margin, is $\frac{2}{||\vec{w}||}$. So by minimizing $||\vec{w}||$, or $\frac{1}{2}||\vec{w}||^2$ for mathematically convenience, we would maximize the margin (Burges 1998).

In order to minimize the margin subject to inequality (2) constraint, using the Lagrange multipliers is a solution. To form the Lagrangian, the constraint should be multiplied by positive Lagrange multipliers (α_i) and subtracted from the objective function. This gives:

$$L = \frac{1}{2}||\vec{w}||^{2} - \sum_{i=1} \alpha_i [y_i(\vec{x}_i \cdot \vec{w} + b)] - 1] \qquad (3)$$

Then in order to find the extremum, partial derivatives in respect to $\vec{w}$ and b should set to zero.

$$\frac{\partial L}{\partial \vec{w}} = \vec{w} - \sum_1 \alpha_i y_i \vec{x}_i = 0 \Longrightarrow \vec{w} = \sum_1 \alpha_i y_i \vec{x}_i \qquad (4)$$

$$\frac{\partial L}{\partial b} = -\sum_1 \alpha_i y_i = 0 \Longrightarrow \sum_1 \alpha_i y_i = 0 \qquad (5)$$

Substituting equations of (4) and (5) in the Largrangian equation gives:

$$L = \frac{1}{2} - (\sum_1 \alpha_i y_i \vec{x}_i).(\sum_1 \alpha_j y_j \vec{x}_j) - (\sum_1 \alpha_i y_i \vec{x}_i).(\sum_1 \alpha_j y_j \vec{x}_j) - b \sum_1 \alpha_i y_i + \sum_1 \alpha_i$$

$$L = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j (\vec{x}_i . \vec{x}_j) \qquad (6)$$

As shown in equation (6), maximizing the margin is only depended on the dot product of the vectors, with respect to α_i and its positivity, and also subject to constraints (5) and with solution given by (4) for a linear separable support vector training (Cortes and Vapnik 1995). It is also the case for non-linear separable support vectors (Burges 1998).

In the next section, first we present and discuss the results of the three qualitative studies. Then we demonstrate the results of the quantitative study through the support vector machine.

Results and discussion

Study 1: Qualitative analysis of tweets

As stated, once we noticed a great passion towards PSL, we brought it into focus. We received the first tweet about PSL on July 21:

- "*I'm so excited for Starbucks to have their red Christmas cups. And for pumpkin spice lattes to be in season*".

Similar to the previous qualitative studies on Twitter (Gaspar *et al.* 2016), we narrowed down the study 1 window to have the chance to review all tweets one by one. Although we chose days around the official return of PSL to Starbucks menu, we did not constrain ourselves to PSL related tweets. In fact, we analyzed all tweets in that time frame, totaling 254 tweets, even those which belonged to other brand aspects or products. Here, we present and discuss some tweets by categorizing them in their appropriate brand authenticity dimensions.

Quality commitment

Building on the emerged items of the quality commitment dimension (as shown in Table 1.2), tweets which deal with product ingredients, improving quality, quality standards, artisan skills, customized manufacturing processes, and having involved a master craftsman (barista) are categorized under the quality commitment dimension. Customers usually share on social media their positive or negative evaluations of Starbucks's quality commitment in-store through their smart phones or tablets, i.e.:

- *"Tried my first ever caramel macchiato from Starbucks. Yummy considering I don't usually like coffee."*
- *"Congratulations Starbucks. This is the weirdest shet I've ever drank ever. #tasteslikepoison... https://t.co/3FrlKNoh1T"*

They also talk about the ingredients and how they taste, i.e.:

- *"White chocolate, coconut and lime cookie from Starbucks is the best thing in the entire world..."*

At the time of data collection, Starbucks was promoting the PSL by emphasizing its coming back "with real pumpkin." This led to questioning PSL quality in previous years. For example, one wrote:

- *"The Pumpkin Spice Latte is back at Starbucks! Now made with actual pumpkin in it! Wait, what have I been drinking all these years?"*

That might even be the reason for negative comments regarding the new PSL taste, i.e.:

"This year's #PSL recipe doesn't taste as good as last years! Sorry, @Starbucks! At least this saves me a lot of money"

- *"Pumpkin scones @Starbucks are different this season. No taste. Zero. #ButIAteWholeThing"*
- *"I don't get why Starbucks had to start using real pumpkin this year ...it was way better last year"*

However, PSL still had its own fans. For example, someone shared about their first ever PSL, and another one about their first PSL of the year:

- *"@Starbucks just tried the new #PSL. Love the new recipe! http://t.co/psSArukYuJ"*
- *"Had the #PSL today for the first time ohh my god it's amazing now I know why everyone likes it haha @Starbucks #tobeapartner"*
- *"Just had my first @Starbucks PSL of the year. Decaf and skimmed milk mind due to remaining a good girl. Still tastes amazing"*

Considering both the positive and negative tweets, one could argue about the importance of quality consistency over time. If quality varies and does not match customer expectations, a consumer will be more likely to switch to another product or brand (Kim and Sullivan 1998).

There were also several tweets complaining about service and/or product quality. For instance:

- *"TERRIBLE SERVICE...don't ask for personal drink order...they don't read them... #disappointed (@ Starbucks) https://t.co/C93Rd34kMu"*
- *"I'm not one to complain, but I just found a finger/toe nail in my @Starbucks this morning. So gross."*
- *"So this happened at @Starbucks this morning. I mobile ordered, it wasn't ready and then I get this. #disappointed http://t.co/yyEGutNj82"*

In most of the complaint cases, Starbucks Help (@starbuckshelp) came back to the customers in a timely manner, and asked for more details regarding the complaint to follow up on it. However, a proficient store manager or barista could handle the case even better. This may not only impede or prevent the decrease in perceived quality commitment, but may also bring positive tweets like:

- *"This lady at Starbucks was so boss at taking care of a customer complaint."*

Heritage

Based on the descriptions about the heritage dimension (as shown in Table 1.2), tweets which show Starbucks's connection to a historical period in time, culture and/or specific region, or tweets which remind the customers of a golden age, and highlight Starbucks' authenticity to its espoused values are categorized under the heritage dimension. Tweets in this category mostly came from shares about Seattle, Washington, where Starbucks was founded in 1971, or about the Seahawks, a professional American football franchise based in Seattle. For example:

"Starbucks first store at Pike Place #Starbucks #Seattle #coffee @Starbucks #Washington #sightseeing... https://t.co/yrxuVIWUdI"

"Seattle Seahawks Starbucks. https://t.co/CCLF7CVO83"

In line with another heritage item, we found that customers also tweet about missing a regular habit they had in the past in Starbucks. They also share what contact with Starbucks highlights from their past memories, i.e.:

"I've been back in LA for 36 hours and have already stumbled upon a live reading in the back of a Starbucks. I think I maybe missed this?"

"Working from Starbucks, sipping a London fog, cloudy day… makes me feel like I'm in college again which is kinda tight"

Moreover, some customers question Starbucks about its espoused value, i.e. "embracing diversity of every kind." For example:

"So supposedly #Starbucks "embraces diversity of every kind"; except a #Christian diversity? Sounds to me like (con't) http://t.co/ggCGKnoWmc"

While we expected more tweets about brand heritage, the coding process resulted in a few ones, totaling 66 tweets. This may be due to the fact that in social media, users are more engaged with present events for a brand, like Starbucks, that they are in touch with on a daily basis. We validated this explanation with the data from the laboratory study.

Uniqueness

Following the guidelines about the uniqueness dimension (as shown in Table 1.2), we classified the tweets concerning the ways Starbucks is fulfilling its brand promises in comparison with competing brands under the uniqueness dimension. Practically, we examined if customers perceive Starbucks services and products in a very different, unique or distinct way from competitors. Customers usually mention the competing brand, use the *than* conjunction to compare, and express their attitudes and reasons for the comparison. For example:

- *"@TheRealPSL @Starbucks I don't need one at @DunkinDonuts I enjoyed my nice Iced pumpkin latte and greeted w/ [=with] awesome smiles #mademyday"*
- *"@DutchBrothers is FAR superior to Starbucks anyway! Better coffee and much friendlier, unpretentious customer service! https://t.co/5N7eXPal1O"*
- *"RT @JuicyZac: Tim Hortons over Starbucks anytime anywhere any day"*
- *"Nothing better than Starbucks and a shower ... and music ... Hella feeling myself right now"*
- *"RT @PreventionMag: This homemade pumpkin spice latte is healthier, tastier, and even quicker than Starbucks: http://t.co/FQjZ20WkX6"*

In many cases, when users reveal their product desires, they highlight their sentiments towards Starbucks' uniqueness by using the phrase *from Starbucks.* For instance, in the following tweet, the customer orders any frozen drink as long as it comes from Starbucks and not from anywhere else. This specification is also observed in another tweet about requiring a Frappuccino with chocolate cake.

- *"A frozen drink from Starbucks sounds great right now."*
- *"If someone can bring me a vanilla bean frappe with 2 chocolate cake pops from Starbucks, that'd be great"*

The other stream of tweets in this category comes from price sensitive customers by comparing Starbucks prices or by stating their intentions to go elsewhere as they could not afford Starbucks costs. For example:

- *"When you're too broke for Starbucks so you go to McDonalds for ice coffee"*
- *"Might have to take out a loan to be able to afford Starbucks"*

The last aspect of Starbucks brand uniqueness is directly related to the way this brand exploits the power of social media. To influence consumers, positive e-WOM is considered a powerful marketing medium for companies (Jansen, Zhang, Sobel, and Chowdury, 2009). That is exactly what Starbucks utilizes from its popularity on social media. The way that Starbucks customers share their positive sentiments on social media, not only promotes the brand uniqueness itself, but also attracts other users to purchase Starbucks products. The following tweets support this argument:

- *"Starbucks don't need ads cuz [=because] teenage girls Instagram accounts are their ads..."*
- *"Social media users go nuts for return of Starbucks pumpkin spice latte #Durham http://t.co/S0m0bUE3Fn"*
- *"Genius marketing by @Starbucks cuz [=because] I'm gonna go and get a pumpkin spice latte 1st thing in the morning http://t.co/jyex1lPmID"*

Symbolism

In line with our conceptualization of symbolism as one of the brand authenticity dimensions, we explored our dataset to catch tweets that have indicated the following items. These tweets suggest Starbucks is/is not adding meaning to people's lives, reflecting important values that people care about, connecting people with their real selves, connecting people with what is really important, caring about protecting people's identity, and fulfilling one's life in providing his/her favorite product. In this essence, customers purchase Starbucks beverages and foods not just for consuming but for adding meaning to their lives. We observed several tweets that appreciate a Starbucks product or the Starbucks itself because of wishing for a better life. For example:

- *"RT @KaileyDaggitt: Life is so much better when Starbucks has pumpkin spice."*
- *"When Starbucks is life. @jiffpom http://t.co/69mfuemTbf"*
- *"#PSL is back at Starbucks.... Fall weather just needs to come, my life will be complete"*

Based on identity theory (Stryker and Burke, 2000), one could expect that customers care about the way they are named and called by salespersons. Previous research reports how students interpret misspelling or mispronouncing by cashiers or workers at an on-campus Starbucks store (Kanemoto and Dai, 2015). In accordance with that research, we found several customers wishing to be known by baristas or complaining about being named wrongly. For example:

- *"Life goal is the Starbucks worker to know my name and order"*
- *"Starbucks has really outdone themselves with misspelling my name...#starbucksfail #notmyname #fail... https://t.co/SNIbVd5sJ3"*

As the last stream of tweets in this dimension, we noticed several calls to boycott Starbucks because they believe that Starbucks does not reflects values they care about. For example, in response to Starbucks CEO's statement against shareholders who support traditional marriage, we found many tweets containing invitations to boycott Starbucks. For instance:

- *"Boycott Starbucks... Let your wallet speak your values and faith. Time for Starbucks to learn a business and lesson http://t.co/Ez8mC1Pd4G"*
- *"THIS IS ASININE IT'S TIME 2 BOYCOTT. Starbucks CEO: If You Support Traditional Marriage We Don't Want Your Business http://t.co/0yDC5ACuqo"*
- *"@D_Rob317 @donna_jacobsen @FreeAmerican100 @Starbucks I'm asking conservative Christians to boycott Starbucks. As the CEO said"*
- *"RT @estera8763: AFTER 16 YEARS OF WHITE CHOCOLATE MOCHA AND WARM BUTTERED CROISSANT, I AM DONE. #BOYCOTT https://t.co/Wttzivxqm1"*

These tweets show how consumers react to a brand once they perceive their values are not being respected. Since it is well documented that boycotting devalues customer perceptions of a brand (Klein, Smith, and John, 2004), managers should be aware of their states and behaviors in order to avoid boycott campaigns.

None of them

In addition to the four dimensions of brand authenticity−*quality commitment, heritage, uniqueness,* and *symbolism*−we added an alternative option for our coders to choose from. When they found that a tweet is related to brand authenticity but the four dimensions are not able to capture it, they chose the *none of them* category. They labeled 69 tweets in the *none of them* category which means 3.1% of all tweets did not belong to any of the four dimensions.

Tweets in this category are covering a wide range of issues, from difficulties in finding stores, stores being crowded or having long lines, to appreciating Starbucks because of product unrelated issues. Here are some examples of these tweets:

- *"Why is it that I get the same drink every time I go to Starbucks, but it's always a different price depending on where I go?"*
- *"So, @Starbucks why can't I login to the website to manage my card?!"*
- *"When Starbucks is so crowded you don't even want to go in"*
- *"There are like no Starbucks anywhere"*
- *"I'd like to thank Starbucks for calling 911 and giving me all the stuff to give to the doctor on scene! You guys rock!"*
- *"Loving the new look of our favorite Starbucks! @Starbucksph https://t.co/fUZEWzBqJC"*

Irrelevant to brand authenticity

Since the research object was to analyze the tweets' brand authenticity sentiments, we grouped the tweets which were not about brand authenticity in the *irrelevant to brand authenticity* category. This helped us have a lexicon for tweets which could be ignored in sentiment analysis. The focus of these tweets was not about Starbucks while they talk about it. For example:

"When your best friend knows your exact order at Starbucks and brings it to you after a long night out. Thanks @_TheChosenJuan_ #TrueMVP"

"Cam fell asleep in a chair in Starbucks and twitched so hard he woke himself up lmao"

There were also some tweets in this category where Starbucks is introduced as a prize or a Starbucks gift card is offered. For instance:

- *"2 Winners for Starbucks! Also join online community to earn Amazon! At @CrunchyBchMama #giveaway! http://t.co/SkYg72TTBm"*
- *"L #Gifts #Cards New 2015 Starbucks Coffee Taiwan Gift Card FRAPPUCCINO... #Coupon #BuyItNow... http://t.co/ihFff767X6"*

Not sufficient information to judge brand authenticity

The last category consists of tweets where coders could not judge the brand authenticity sentiments due to a lack of sufficient information or cues. This might come from having few words or it might be because of our evaluation outside the situation of the tweets. For instance, in the following tweets, while we know that the user is sharing feelings towards Starbucks, we do not have any cues that help us categorize them under a proper brand authenticity dimension:

- *"@KIRSTlN I love you more than Ariana and Nathan loves Starbucks"*
- *"RT @Wonder_Buns: @LGP4july @Just_a_Texan @helpsosme same here! I tell everyone I know that @Starbucks is a piece of crap company"*
- *"RT @hansclm49: That s easily fixed. Bye Bye* Starbucks! *https://t.co/xA582BnXEt"*
- *"At* Starbucks *https://t.co/y3Js1QvJb7"*

Study 2: Latent semantic analysis (LSA)

As stated, we used LSA to group the common words in each category. This helped us support the coding process and also the findings from the qualitative study using another perspective. For the *quality commitment* dimension, we found that the established common words are related to perceived quality issues, such as: great, make, crave, taste, really, good, love, pumpkin, coffee, and like. These words are representative of the tweets, which we discussed earlier in the section regarding the quality issue. Similarly, the words such as: day, life, basic, boycott, fall, white_girl, first, back, pumpkin_fall, and name remind us of the positive or negative attitudes towards Starbucks in terms of *symbolism*; adding meaning to its customer lives, or reflecting important values that customers care about. Likewise, the common words in the *uniqueness* dimension, present its singularity by for example coffee comparison and admitting addiction. Those words are: better, addict, just, coffee, girl. For the *heritage* dimension, as expected, we found Seattle, Seahawks, miss, and remember as the frequent words. Finally, results of LSA were also consistent with those of qualitative analysis for the tweets which were *Irrelevant to brand authenticity*. We found gift, coupon, and eBay as the common words for this dimension, supporting the discussed tweets earlier in this section. Table 1.3 shows the most common words in each dimension. Due to the wide range of tweets in the *not sufficient information to judge* *brand authenticity* and *none of them* categories*,* we did not get common words for these two categories.

Table 1.3: Most common words in each dimension

Quality commitment	Symbolism	Uniqueness	Heritage	Irrelevant to brand authenticity
great	day	better	Seattle	gift
make	life	addict	Seahawks	coupon
crave	basic	just	miss	eBay
taste	boycott	coffee	remember	
really	fall	girl		
good	white_girl			
love	first			
pumpkin	back			
coffee	pumpkin_fall			
like	name			

Study 3: Laboratory setting brand authenticity analysis

As described earlier, we asked one hundred eighty undergraduate students (52% female) to share their sentiments about Starbucks, answering four questions, related to each of the four brand authenticity dimensions. Our sample had 3.76 social media accounts (including Facebook, Instagram, Twitter, LinkedIn, Google+, Pinterest, Tumblr, and Snapchat among others) in average (median=4, S.D. 1.66). Regarding time spending on social media, 61% of our sample use social media more than one hour (0-30 mins: 10%, 31-60 mins: 29%, 61- 90 mins: 22%, 91-120 mins: 13%, 121- 150 mins: 10%, and more than 150 mins: 16%). In the next sections, we quote some of the shared sentiments by our participants respecting Starbucks's perceived quality commitment, heritage, uniqueness, and symbolism.

Starbucks's product quality

We received one hundred fifty valid responses about the situation, in which respondents want to share something about Starbucks's product quality on social media. In line with our findings from the tweets, there are different sentiments towards products' quality. For example:

- "Their coffee doesn't taste good and I think Starbucks is overrated."
- "I'm a coffee lover and I love the taste of their coffees."
- "Very disappointing: at first, I thought that Starbucks was a good brand, but then I heard that they ally with Monsanto, so the quality of the products is more than disappointing, it's poisoning us. It really changed the way I like Starbucks."
- "The Starbucks product quality is amazing! I am a person that likes strong coffee and Starbucks is able to fulfill my consumer needs!"

Although we gathered this data in early spring, far from the fall, which is the season of PSL, one of our participants express her interest in the PSL. She wrote:

- "Nothing better than a Pumpkin Spice Latte!"

As another source of perceived quality commitment, several participants appreciate its consistency across different branches, i.e.

- "Starbucks' product quality is really good, with a consistent product across branches. They source their ingredients from a variety of geographic regions, and are open about it, which I appreciate."
- "Consistent, however not the best. You always know what you're getting."

, however, someone reports a contradictory observation:

- "Not consistent across countries. Way better quality in the Middle East than in North America."

Taking these examples to consideration and comparing them to the tweets (discussed earlier), and the extracted words (shown in Table 1.3), we see consistency between findings from Study 1 and 2 and the current study. This shows validity of our findings about quality commitment across different samples and different research methods.

Starbucks's connection to a historical time, culture and/or region

One hundred eleven valid responses received under the question regarding sharing about Starbucks's connection to a historical period in time, culture and/or specific region. Among the answers, near one third were declines of knowledge about Starbucks's heritage. We consider the following quotes as the reasons of lower amount of heritage related tweets in the database.

- *"I never heard about Starbucks historical time or culture."*
- *"Starbucks has become a trendy hang out place for the younger generation. Don't know much about the past history."*
- *"To me, Starbucks is a global brand that doesn't really reflect a specific history."*
- *"I don't see any link to this. It really doesn't feel like a "traditional" brand to me. Very modern and in today's time. Not any historical period."*

However, rest of the participants mentioned its connection with the American culture, the 1970s time period, and its first store in Seattle's historic Pike Place Market. For example:

- *"Starbucks does have a connection to the 1970s era when coffee became highly stipulated into the North American lifestyle and has been ever present since then."*
- *"It's a good cultural reference to Italian "bars" in Italy, where they serve coffee and espressos. They took that concept and brought the cafe experience to North America."*
- *"Starbucks's linked with North American's culture, people don't have time to sit and drink a coffee in the morning. In North America (vs. European take the time in the morning), so Starbuck's did think about a new and practical way to facilitate worker's life."*
- *"They still have their first Starbucks in the Pike Place Market, Seattle, and it's very lovely."*

Responding to this question, as expected, percipients also shared about remembering or missing an experience with Starbucks. For example, one of the participants wrote about her first experience with Starbucks, and another one shared her favorite childhood sweet roll:

- *"Always keeps its traditions but still caters to new cultures and traditions that form and develop every day. First time I entered a Starbucks was unlike any other experience. It was a very relaxed environment. My order was made personal and I felt as part of the community. They say you never forget first impressions and this experience is no exception. Every time I pass by a Starbucks, I remember the first time I got something from there."*
- *"Starbucks is the only brand that makes Cinnamon rolls, which reminds me of my childhood in Sweden and when my mom would make them later in France. As I am not a coffee drinker, my main connection to Starbucks is to this sole product. The quality has remained constant, not as good as home baked, but good enough for me to buy them frequently."*

In summary, in this study we checked the robustness of our findings about heritage from Study 1 and 2. We also reported more explanations for the relatively low heritage related tweets.

Starbucks's ordinary/unique characteristics

We got one hundred thirty valid responses to sharing something about Starbucks's ordinary/unique characteristics. A few participants mentioning not being aware of unique characteristics for Starbucks, e.g.:

- *"Starbucks is not unique in my opinion."*
- *"Not much uniqueness, standard business model."*
- *"Starbucks is not that unique of a brand, and the price of their food is pretty expensive."*

However, most of the responses contained different kinds of sentiments in regard to Starbucks uniqueness. As expected, they usually compared Starbucks characteristics with other coffeehouses. For example:

- *"The ambiance is what differentiates them from the Second cup and others."*
- *"Starbucks was my favorite store previously. Their coffee was better than other brands and the slightly more expansive coffee was worth it for the atmosphere but it has greatly changed lately."*
- *"I prefer Tim Hortons to be honest and barely go to Starbucks."*

Moreover, some participators mentioned product's customization, its friendly atmosphere, providing seasonal drinks, changing cups' color for the holiday season, and also its marketing strategies as features for its uniqueness, e.g.:

- *"The drinks are customizable and are flexible to your needs, which not many coffee places have."*
- *"Good environment, like a living room space where you can relax, talk with people conduct business or do work."*

- *"I like how they change the color of their cups for Christmas."*
- *"Very very good marketing is the best characteristic they have."*

Participants also specified the higher price of Starbucks products as another source of its perceived uniqueness. For instance:

- *"I love Starbucks but I find it very expensive, therefore it is rare that I can afford it."*
- *"So expensive, God damn!"*
- *"Expensive addiction, expensive habit."*

In sum, Study 3 helped us to replicate findings from Study 1 and 2 with another data, which provides support for the coding schema.

Starbucks adding meaning to consumers' life

We received one hundred thirty-one respondents to the question about Starbucks's products, values, behaviors, and so on that adds meaning to one's life. Few people stated that Starbucks is a coffeehouse, offering coffee and tea, so they do not think that it would add anything to their lives. As an example:

- *"I don't believe Starbucks products add meaning to life as it is only coffee and teas etc."*

Moreover, we found that being a brand's consumer is not sufficient to feel a brand's symbolism. But, if someone considers a brand is adding meaning to his life, he would also being attached to that brand. For example:

- *"I am a Starbucks consumer but I don't consider myself the attached customer that only has Starbucks products. So it doesn't really add meaning to my life."*

In this study our participants shared the cases they perceive Starbucks is adding meaning to their life. For instance:

- *"Starbucks cookies are life."*
- *"Makes my day brighter."*
- *They also mentioned several observations supporting their feelings. For example:*
- *"Potential links to political factors put a damper on its alleged values and good behaviour."*
- *"Some of the shares of their products go to charities."*
- *"I'd like to see more giving back to communities."*
- *"The only reason I go to Starbucks is for one specific product. I try to boycott since I believe they have sacrificed freshness and quality for profits."*
- *"It provides yummy vegan drinks!"*
- *Similar to what we previously found about the importance of being correctly named and called, we got several approaches to this subject. For example:*
- *"I love the "name" feature, makes people feel like individuals and not consumers."*
- *"When you order a drink at Starbucks, the baristas write your name on the cup which makes the customer feel special and nice."*

- *"The cashier often messes up the names of their clienteles... Although I spell out my name instead of simply saying it, they end up misspelling the name. But the cashiers and baristas are still very nice and welcoming."*

Finally, we found several quotes indicating consuming Starbucks in a conspicuous manner, or calling the girls going there as the white girls, e.g.:

- *"Starbucks in known to be more luxurious than the other companies (it's more expansive). Also, when a girl now buys from Starbucks people may call her a basic white girl."*
- *"Starbucks coffees are expensive, bad quality and are only a way for people to associate themselves with the brand and show off!"*
- *"Never been to Starbucks, but I guess they do something right with the product since every white girl go."*

Taking findings from Study 3 into account, we provide support for findings from Study 1 and 2 from a new perspective.

Study 4: Quantitative study through support vector machine (SVM)

Study 1, 2 and 3, using different samples and/or methods, show the robustness of the coding schema for the brand authenticity sentiment analysis. While the findings have several theoretical and managerial implications, which we discuss them further in the next section, these studies are still qualitative in nature and might be cumbersome for big data analysis. In other words, while we think without conducting previous studies, one could not check the robustness of the coding process, for a big data study we need for a framework to perform the analysis automatically. Therefore, the object in Study 4 is to predict automatically the brand sentiments both in terms of associated authenticity dimension–among the seven options−and its polarity–on a scale between -2 to +2. We used the algorithm of support vector machine as the supervised machine learning technique.

Building on the results from the coding algorithm, and based on the frequency of containing the labeled words–unigram, bigram, and trigram−we obtained for each tweet a number containing 35 digits. Each five numbers respectively represent the frequency of words in a dimension from strongly negative to strongly positive. In fact, we have five sets of frequencies; respectively for *heritage*, *none of them*, *symbolism*, *uniqueness*, *irrelevant to brand authenticity*, *quality commitment*, and for *not sufficient information to judge brand authenticity*. As an example, for a tweet like "*#PSL is back at Starbucks.... Fall weather just needs to come, my life will be complete*", we have the following code: [(0, 0, 0, 0, 0); (0, 0, 0, 0, 0); (0, 1, 0, 4, 5); (0, 0, 0, 3, 1); (0, 0, 0, 0, 0); (0, 0, 0, 2, 2); (0, 0, 0, 0, 0)]. It means that the tweet does not contain any of the words associated with *heritage* (0, 0, 0, 0, 0), *none of them* (0,0,0,0,0), *irrelevant to brand authenticity* (0, 0, 0, 0, 0)*, and not sufficient information to judge brand authenticity* (0, 0, 0, 0, 0). However, its n-grams are matched with some coded n-grams, which are associated with sentiment strength of *symbolism* (0, 1, 0, 4, 5), *uniqueness* (0, 0, 0, 3, 1), and *quality commitment* (0, 0, 0, 2, 2).

Since most of the real datasets are not 100% linearly separable, Cortes and Vapnik (1995) introduced positive slack variables ξ_i for the constraints. A constant C is also introduced as the positive multiplier for the $\sum_i \xi_i$, $0 < \alpha_i < C$. C represents trades off misclassification of training examples against simplicity of the decision surface. A low C makes the decision surface smooth, while a high C aims at classifying all training examples correctly by giving the model freedom to select more samples as support vectors. We used Statistica 13 software to conduct SVM analysis.

In this study, we used the one-versus-all type, in which the winner-takes-all strategy is applied (Zhou and Tuck, 2007). It means that for each classification, the classifier with the highest output value is selected as the corresponding class label (Yun, Sim, and Kim, 2000).

We randomly selected 51% of the data for training and 49% for testing. Applying v-fold cross validation (v-value=10), we came into the optimized values for C and gamma in Kernel type of Radial Basis Function of 10.00 and 0.02857 respectively.

The results show a high level of recall both for the training set and the testing set. We got 92.2% accuracy for the training set classification and 89.7% for the testing set, which means 91.0% overall. To derive some confidence on the performance estimate, we ran the model multiple rounds, and did not get significant differences. Table 1.4 summarizes the correct and incorrect classification for each category.

The highest accuracy, 98.4%, comes from the *quality commitment* class. However, the lowest one comes from *not sufficient information to judge brand authenticity.* The low accuracy in this class might be due the wide range of tweets in this category, from a single word to a long sentence with lots of similar n-grams with other categories.

Table 1.4: Classification summary

	Total	Correct	Incorrect	Correct (%)	Incorrect (%)
Heritage	16	12	4	75.0	25.0
Irrelevant to brand authenticity	101	89	12	88.1	11.9
None of them	43	31	12	72.1	27.9
Not sufficient information to judge brand authenticity	60	18	42	30.0	70.0
Quality commitment	737	725	12	98.4	1.6
Symbolism	538	511	27	95.0	5.0
Uniqueness	362	303	59	83.7	16.3

The confusion matrix in Table 1.5 shows the number of confused classifications for each dimension. For example, from the total of 362 tweets in the *uniqueness* category, one is incorrectly classified under *irrelevant to brand authenticity,* 49 under *quality commitment* and nine under *symbolism*. The elements on the main diagonal represent the tweets which were correctly classified.

Table 1.5: Confusion matrix

	H	I	NOT	NSI	QC	S	U
Heritage	**12**	0	0	0	3	1	0
Irrelevant to brand authenticity	0	**89**	0	0	11	1	0
None of Them	0	1	**31**	0	8	1	2
Not sufficient information to judge about branding	0	2	0	**18**	23	16	1
Quality Commitment	0	0	0	1	**725**	2	9
Symbolism	0	0	0	0	20	**511**	7
Uniqueness	0	1	0	0	49	9	**303**

To visualize the model's prediction accuracy, we graphed the observed brand authenticity vs. predicted vs. its accuracy with a surface plot. As exhibited in Figure 1.4, a large amount of data is placed around z=0, which translates to the model's high prediction accuracy.

After obtaining predictions of the associated brand authenticity dimension, we ran another SVM algorithm to analyze the overall sentiments towards the brand. This time we used SVM regression in order to predict the sentiments. Using ten-fold cross validation, we obtained C=10, epsilon=0.20 and gamma=0.029 for Kernel type of Radial Basis Function.

Our results establish a high level of correlation between the observed and predicted values for our dependent variable, brand sentiment. The training set and the testing set results show correlations of 0.848 and 0.827. For the whole set, the correlation is 0.837. Table 1.6 provides more details about the regression findings.

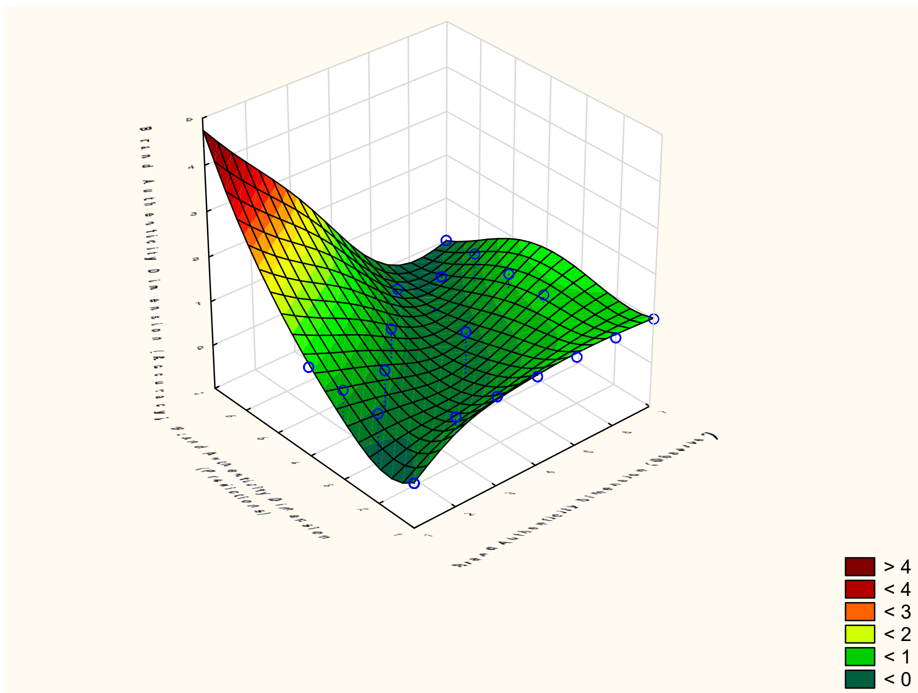


Figure 1.4: Observed brand authenticity dimension (x) vs. predicted brand authenticity dimension (y) vs. its accuracy (z)

Table 1.6: Regression results

	Training sample	Testing sample	Overall
Observed mean	0.5523	0.4396	0.4970
Predictions mean	0.5033	0.4443	0.4744
Observed S.D.	1.0011	1.0384	1.0209
Predictions S.D.	0.7903	0.7963	0.7936
Mean squared error	0.2879	0.3448	0.3158
Error mean	0.0490	-0.0048	0.0226
Error S.D.	0.5346	0.5875	0.5616
Abs. error mean	0.4302	0.4569	0.4433
S.D. ratio	0.5340	0.5657	0.5501
Correlation	0.8475	0.8268	0.8372

GENERAL DISCUSSION

All in all, we contributed to the literature in terms of analyzing brand authenticity in social media. Moreover, we extended the current knowledge about brand sentiment analysis by addressing the reasons for getting positive or confronting negative sentiments. Furthermore, we improved the sentiment measurement accuracy by labeling and predicting the sentiments on a five-point-scale. Practically speaking, we obtained high accuracies in training and testing the dataset, suggesting the usefulness of the proposed algorithm for predicting both brand authenticity dimensions and its sentiment polarity. Taking into account the results from both Study 1 and Study 2, we provide important managerial implications.

First, changes to the specifications of popular products should be made with great caution. Our findings show that while PSL has its own fans−waiting for fall throughout the year−once they found its taste was different from their expectations, they began to share their negative sentiments about the PSL and the Starbucks brand. Second, companies should find ways to engage customers in sharing positive e-WOM on social media. The results suggest that it could be done by working on improving different aspects of the perceived brand authenticity. For companies like Starbucks, with many followers on social media, highlighting services and products to reinforce ties with customers would be an appropriate approach (Chu and Kim, 2011). Third, while some issues might look ordinary, i.e. misspelling customer names by Starbucks baristas, firms should be aware of their potential viral effects on social media and the possible negative impacts on customer's perceived authenticity. Fourth, managers should be aware of their statements and behaviors specifically about issues that are strongly related to social values and beliefs. In this study, we found evidence that compromising these issues not only decreases perceived brand authenticity, but it also might lead to calls for boycott campaigns on social media. Finally, firms should not only continuously monitor sentiments toward their perceived brand authenticity, but they should also look for ways to improve it in comparison with their competitors. To this end, our framework provides a useful way of understanding the tweet's related brand authenticity dimensions and sentiment polarity. This would help brand managers effectively observe what is happening on social media about themselves or their rivals, gain insights from the data in order to take proper actions to increase customer perceptions, and finally, formulate new insights for designing short-term and long-term strategies.

Conclusion

In this research, we propose a new algorithm to analyze brand sentiments on social media. We empirically shed light on how brands could investigate sentiments towards them in terms of perceived brand authenticity. Results from both the qualitative and the quantitative studies demonstrate its usefulness. Through the qualitative study, we discussed several tweets from our dataset which our coders classified under the *quality commitment, heritage*, *uniqueness,* or *symbolism* categories. We found a few tweets which were about brand authenticity but not associated with our four categories, placed in the *none of them* class. We also provided examples about the tweets that were categorized under *irrelevant to brand authenticity*, and *not sufficient information to judge brand authenticity*. Using LSA, we extracted the common words in each of the categories. The extracted words supported the coding process and also our discussion. However, due to the diversity of tweets in *none of them* and *not sufficient information to judge brand authenticity* categories, we did not receive significant common words for these two categories. The results of SVM analyses show high accuracy for the prediction of both the brand authenticity dimensions and its sentiment strength. With the prepared lexicon and the training dataset, the recall for both analyses was above 0.83.

In summary, this research contributed both theoretically and managerially to the brand sentiment literature. The proposed procedure, and the research findings on brand authenticity sentiment analysis could facilitate further inquiries into sentiment analysis for all other brand constructs and also in several related domains, such as e-Word-of-Mouth studies. Practically speaking, this research could provide marketing practitioners with a reliable and valid instrument to evaluate the level of sentiments towards a brand more specifically and more accurately, which could lead to proposing appropriate strategies to strengthen their brand authenticity.

Limitations and future research

This study has some limitations that offer opportunities for further research. First, we studied brand authenticity sentiments with cross-sectional data, while we acknowledge that customer sentiments could even change over a short period of time. Thus, future research could examine how the sentiments towards brand authenticity vary over time. Second, we did not take into account the brand's follow-up interventions about the shared tweet. We think it could be interesting to see if this kind of intervention has positive effects on perceived brand authenticity. Moreover, we expect that this potential benefit would vary depending on the time lapse between the sharing moment and the brand's follow-up. Third, as we collected the data in a real-time manner, we did not have the data about the amount of likes and retweets a post received. Future studies could use retrospective data to access this sort of data to see if there are sentiment differences between popular tweets and others.

PAPER 2: HOW UBER CAN ALTER DESTINATIONS? IMPLICATIONS FOR THE RIDESHARING INDUSTRY

ABSTRACT

This article extends knowledge in the novel field of destination-based advertising (DBA). As the context, we chose the ridesharing industry to test how consumers react to DBA, and how it changes their planned behavior. Practically speaking, the research question is how a ridesharing app, i.e. Uber, can alter a destination. Building on the construal level theory, we examine if spatial distance (low: 50m vs. high: 3000m), need for cognition (NFC) (low vs. high), cultural distance (high: international app vs. low: Canadian app), and marketing incentives (low: non-free rerouting vs. high: with free rerouting) matter in DBA. Moreover, we study how susceptibility to apps mediates the interactive effects of spatial distance and need for cognition. We develop two experimental mock apps that mimic a ridesharing app in Canada. The results of four studies support the effects of spatial, NFC, cultural distance, and incentives, as well as their interactions. More specifically, we show that when users get DBAs, their decision of whether to alter their planned behavior is transmitted through their level of openness and curiosity (susceptibility). We also confirm that adding marketing incentives to high spatial distance makes them as interesting as the ads near the planned destination. Implications for the theories of planned behavior and construal level, and how practitioners can use DBA to alter planned behavior, i.e. planned destinations, are discussed.

INTRODUCTION

Given the ever-increasing proliferation of mobile devices, mobile marketing has received significant attention from academics and practitioners (Appel *et al.* 2020; Kim *et al.* 2015; Liu *et al.* 2012). Location-based advertising (LBA) is a recent form of mobile marketing that allows managers to provide consumers with tailored mobile advertisements based on their current location (Bruner and Kumar 2007; Izquierdo-Yusta *et al.* 2015). For example, Whole Foods places geofences around its stores to target ads and send special offers to mobile devices which pass by the locations, achieving a 4.69% post-click conversion rate increase (Beaconstac 2019). Studies show optimistic forecasts about LBA (Lin 2016). In the U.S. alone, LBA ad spending is expected to grow from 17.1 billion USD in 2017 to 38.7 billion USD in 2022 (Statista 2019).

While there is a growing body of literature studying consumer behavior in location-based advertising, novel LBA methods have not yet received much attention in the marketing literature. For example, Ghose *et al.* (2019a) showed the effectiveness of using trajectory data to predict customer responses to mobile coupons. However, none of their treatment groups received mobile ads based on their destinations. Another field study shows that, among users of a public transit app, those who received mobile coupons while commuting between home and work are more likely to redeem food and beverage offers than those not commuting (Ghose *et al.* 2019b). Nevertheless, the research is neither concerned with the spatial distance between users' location/destination and the advertised chain stores, or users' abandoning their trip. In a recent study by Bernritter *et al.* (2021), while researchers studied behaviorally targeted location-based mobile ads, they did not consider sending ads based on consumers' destinations. In other words, existing research about location-based advertising has not investigated destinations as locations. While the destination is not totally independent from a user's current location, some mobile applications, i.e. ridesharing apps, not only have access to users' current location but also to their planned routes and destinations. In fact, the purpose of destination-based advertising is not only to create interest in the advertised product or retail store, but also to alter planned destinations. (Luo *et al.* 2014). The research question is how a ridesharing app, i.e. Uber, can alter a consumer destination by providing relevant ads tailored to the destination. This is heightened by the fact that consumers are using more geo-enabled mobile gears, including wearable devices. In this research, building on construal level theory, we examine how spatial distance, moderated by the effects of need for cognition (NFC), plays a role in altering planned behavior via susceptibility. We also look into app design in regards to cultural distance (high: international app vs. low: Canadian app), and marketing incentives (high: free rerouting vs. low: non-free rerouting) that could be used to make ads far from planned destinations more pleasant. We develop two experimental mock apps mimicking a ridesharing app in Canada. Drawing on the Theory of Planned Behavior (TPB: Ajzen 1991), the results of four consecutive studies support the role of spatial distance and NFC. We also shed light on susceptibility to apps as the underlying mechanism. More specifically, lower psychological distances (i.e. cultural distance) and spatial distances lead to more positive attitudes about the ad, higher purchase intention and app reuse intention, and a higher chance of altering planned behavior; while higher marketing incentives compensate for the effects of the increased unexpected cognitive load produced by higher psychological distances.

The research findings facilitate further inquiries into mobile advertising contexts and in several related domains, such as location-based targeting in social media (Lee and Hong 2016;

Steinhoff *et al.* 2019), and developing more effective locational recommender systems (Noguera *et al.* 2012). We discuss the implications for the TPB and the construal level theory, and shed light on how practitioners can use it to alter planned behavior, i.e. planned destinations. The research also provides marketing practitioners with insights about the factors to consider when they are going to communicate with customers through this new medium. More importantly, considering their access to big datasets of locations and destinations, they could utilize the findings of this research to foster their user's journey experience while gaining the benefits of a new source of revenue: destination-based advertising.

HYPOTHESES DEVELOPMENT

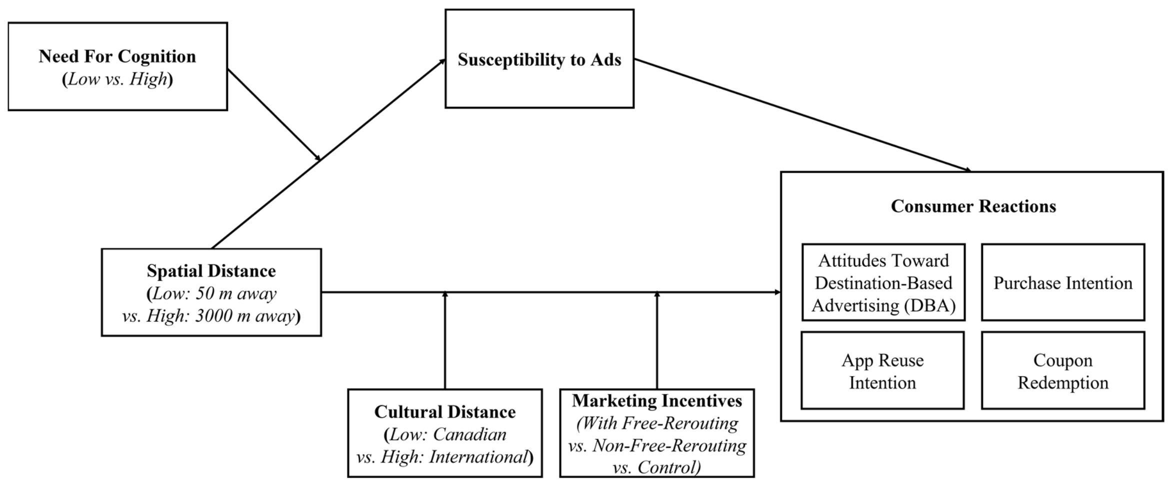


Figure 2.1: Conceptual model for the four studies.

Study 1 examines the interactive effects of spatial distance and need for cognition via susceptibility to apps on coupon redemption.

Study 2 investigates the interactive effects of spatial distance and cultural distance on app reuse intention and coupon redemption.

Study 3 tests the role of cultural distance and marketing incentives for high spatial distance condition on consumer reactions.

Study 4 tests if low spatial distance without marketing incentives differs from high spatial distance with marketing incentives within different cultural-laden app designs.

Ads that address customers' preferences of what they exactly want –when, where, and how (Pine *et al.* 1995) – are known as customized offers. Xu *et al.* (2008) find that users are more satisfied if advertising messages are relevant, and sent to them at the right time and in the right location. This right location could be users' destination rather than their current location because when someone is leaving an existing location, they may be no longer interested in receiving offers related to a place where they would not be (Xue *et al.* 2013).

Building on the theory of planned behavior (Ajzen 1991), we consider coupon redemption as an unplanned behavior that happens if users accept the destination-based ad. In the context of the ridesharing industry, not redeeming the coupon at the "original" destination is a planned behavior, whereas either redeeming the coupon at the "original" destination or heading to the "alternative destination" is an unplanned behavior. DBAs may trigger consumers to opt for these unplanned behaviors. Changing the planned behavior may simply happen by redeeming the coupon at their destination, somewhere near to it, or even at an altered destination.

Construal level theory and psychological distances

Construal level theory (Trope and Liberman 2010) suggests that the way people think about distant events, objects, situations is not the same as how they consider events or objects happening close to them. Subsequently, they interpret, react, and behave differently based on their perceived proximal (vs. distal) distance (Danziger *et al.* 2012). These *near* and *far* distances not only refer to geographic locations (*spatial*) but also involve time spans (*temporal*) and social contacts (*social*), among other psychological distances (i.e. *cultural*). Liberman and Trope (2014) identify traversing psychological distance (TPD) as a forming belief about the substitutability of near (vs. far) events or objects. They also speculate its implication for the purpose of understanding how consumers choose between the alternatives. Research suggests that construal level leads to schematic, abstract, and purpose-focused qualities for high-level events or objects while it brings detailed, concrete, and specific qualities for low-level ones (Rogers and Bazerman 2008; Bolton *et al.* 2021). In this research, and in line with the ridesharing industry characteristics, we focus on spatial (H1- H3) and cultural distances (H4) as the two psychological elements that make an impact on the consumer cognition process.

Spatial distance

According to construal level theory (Trope and Liberman 2010), spatial distance is the physical distance between two events. If an event is taking place near one's location ("*here*"), the information is processed at a low level. However, for an event happening in a far distant location ("*somewhere else*"), the information is processed at a high level. This level difference in processing the information leads to evaluating the events concretely (vs. abstractly) for those occurring near (vs. far). For example, once users receive an ad near their location, they evaluate it thoroughly, consider every detail of it, and see if that meets their expectations. However, for ads far from their location, they broadly consider it without going into details.

In fact, communication through DBA, even about a relevant exciting event, or a greatly customized offer, which is far away from the destination, would not be warmly welcomed, or have a modest effect on ad attractiveness as it gets evaluated abstractly (Ketelaar *et al.* 2017). Supporting this argument, Lee *et al.* (2015) show that sending ads near the current location of users is effective in inducing positive attitudes toward LBA, as it reduces users' effort to assess the ad usefulness. This positive perception of DBA happens by triggering relevant knowledge structures to evaluate the ad more easily with the existing cognitive resources. The developed systems which enable the localization of clients within tens of centimeters (Vasisht *et al. 2*016) heighten the need for customization of spatial distance. Therefore, we expect that the extent to which customers perceive the higher spatial distance between their planned destinations and alternative destinations, the DBA outcomes decrease. Therefore, we hypothesize that:

H1: *Higher (vs. lower) spatial distance to the advertised destination decreases (increases) consumers' willingness to redeem coupons at that destination.*

Susceptibility as the mediator and need for cognition as the moderator

Susceptibility is defined as the state or fact of being influenced by extraneous factors. It varies across consumers with different characteristics and traits (McGuire 1968; Bearden *et al.* 1989). For example, people with higher self-esteem show a higher level of openness to the influences of other factors or persons (Janis 1954). The consumers who openly welcome an ad, deliberately process its messages to not easily miss them (Burke and Srull 1988). This impact on these open mined recipients is similar to the effect of scarcity messages. Limited offers are appreciated more by consumers leading to enhanced value perception of the messages. This stream of research suggests that including desirable objects in advertising leads to triggering self-serving motivation of thinking to be clever if persuaded by the ad. In the context of advertising, we expect offering ads based on their destination would be enticing for consumers, and so it should have a similar provoking self-serving effect. Therefore, the impact of spatial distance on consumer reactions, i.e. coupon redemption and app reuse intention, is an indirect outcome of the level of consumer openness (susceptibility level) to the role of spatial distance. Research shows that the assumption of having more preventive action from higher perceptions of susceptibility is not valid (Weinstein 1984). When consumers are provided with ads near their destination (vs. far from their destination), their level of accepting the ads is strengthened (vs. attenuated) due to changes in their perception of vulnerability against the persuasion of the ads. The increased personal susceptibility could enhance consumer reactions to the ad, i.e. ad adoption or purchase intention (Eisend 2008).

Although the diffusion of spatial distance through susceptibility level promises to evoke consumer reactions to the destination-based ads, it may be still associated with their individual cognitive tendency (Cacioppo and Petty 1982). Previous research shows differences among consumers in their level of preference in being involved in tasks that requires thinking. This is conceptualized as need for cognition (NFC), which is “an individual's intrinsic enjoyment of and motivation to engage in effortful cognitive information processing” (Zhang 1996, p.16). According to Batra and Stayman (1990), consumers with a low level of NFC process messages more heuristically than consumers with a high level of NFC. The latter enjoys completing tasks that require a lot of mental effort and assesses messages concretely. Haugtvedt *et al.* (1992) finds that even when consumers should not compare alternatives, the level of NFC affects attitude change. Individuals with high NFC evaluate product attributes more than consumers with low NFC. Moreover, the attitudes of low (high) NFC individuals are more (less) shaped by incidental cues inherent in the ads.

Regarding how different levels of NFC lead to consumer reactions, Inman *et al.* (1997) show that consumers with low NFC have more intention to purchase from ads with scarcity messages (i.e. limited quantity of on-sale products). This happens because by just knowing that they might miss the promotion opportunity, low NFC consumers would more easily skip different thoughtful tasks like a comparison between alternatives and consideration of features. Similarly, past research finds that low NFC consumers have higher preference for bundle vs. individual items because it helps feeling less overwhelmed by search efforts (Harris and Blair 2012). Building on the existing literature on NFC and construal level theory, we predict that low (vs. high) NFC will enhance the impact of low (vs. high) spatial distance on consumer reactions and their level of susceptibility. Construal level theory suggests that when consumers receive an ad far from their destination, they do not need to process the message concretely. So, although it is far from their planned destination, its evaluation does not exert more effort (i.e. answering questions like whether it is worth going to a new destination rather than the planned destination). Therefore, low NFC consumers tend to adapt their planned destination for the benefit of the destination-based ads. While those with high NFC prefer their life to be filled with different puzzles that they have to solve (Cacioppo and Petty 1982), the joy of evaluating, processing, and exploring the new opportunity would not exceed these pleasures, and so they would show less positive reactions. Past research also suggests that curiosity would facilitate taking riskier behavior in the digital environment because of both more openness to explore and less perceived cognitive burden to adopt technologies (Benenson *et al.* 2017; Moody *et al.* 2017). As such, we expect that changes in consumers' susceptibility transfer the effects of spatial distance and NFC. More formally, we hypothesize the moderating role of need for cognition and the mediation role of susceptibility as follow:

H2: *The effect of high (vs. low) spatial distance of destination-based ads from planned destinations is moderated by the level of need for cognition (NFC): low (vs. high) NFC will enhance the positive impact of low (vs. high) spatial distance on consumer reactions.*

H3: *The interactive effect of spatial distance and need for cognition (NFC) is mediated by the level of susceptibility to the ads: when the spatial distance is low (vs. high) and NFC is low (vs. high), more susceptibility to apps results in more positive consumer reactions.*

Figure 2.1 shows the research conceptual framework.

METHODOLOGY

Spatial distance manipulation

The spatial distance of the ads with users' destinations was manipulated based on the results from a pretest. We asked two hundred and fifty-two undergraduate students how many meters an ad can be far from them to be considered based on their location. The destination average was 1774 meters, and more than 90% of the respondents considered an ad being 3000m away from them to be incongruent. So, we chose 3000m as the high spatial distance stimuli. We primed low spatial distance to 50 meters from their destination (the City Mall), which again was mentioned by more than 90% of respondents.

Study 1

Study 1 examines the effects of spatial distance on the *planned behavior* (coupon redemption) as well as app reuse intention. In fact, we test H1, H2, and H3.

Experimental design

The experimental design was a 2 (spatial distance: low = 50m vs. high = 3000m away) × 2 (need for cognition: low vs. high) between-participants design. Two hundred Mturk respondents (57.5% female, M_{age} = 37.18, SD = 10.23) filled the questionnaire.

Data collection

We removed four incomplete responses and rejected twenty-three in-progress responses because of missing the attention check questions, leaving a sample of two hundred responses. Participants were told that the purpose of the study was to investigate how mobile devices are changing the way customers behave. After accepting to participate, they read the following scenario:

"Suppose you take a rideshare using the RideShare4You app on your device and are going to the City Mall as your destination to purchase clothes, accessories, and/or to eat lunch or drink. While you are sitting in the car and going to the City mall, you receive a welcome note on your rideshare app in regard to knowing the offers based on your destination. Now, please browse the app, see the home page, read the welcome note, go to the contact page, and finally go to destination-based ads. Please do not forget to read the coupons' terms and conditions' page as well. Enjoy your ride!"

Next, they randomly entered one of the two conditions for spatial distance (high vs low). Assignment of the respondents to NFC (high vs low) was based on their responses to the NFC scale (Zhang and Buda 1999). Then, they saw pictures of the ridesharing app (its home page, welcome, contact, and the destination-based ads). On the destination-based ads page, they received the same 30% off in all conditions for *food and drinks*. Then, they filled out the questionnaire. All the constructs were measured through a 7-point Likert scale and adapted from existing measurement scales. Cronbach's alpha coefficients validated the reliability of the scales ($\alpha > .88$) (see Appendix A).

Findings

First, we asked if the scenario and app screenshots helped them imagine they were using their own cellphone and using a ridesharing app. Results showed that it was successful and significantly higher than the midpoint ($M_{imagination} = 6.18$, SD = 1.20, $t(199) = 19.70$, $p_{one\text{-}sided} < .001$). Participants also perceived the ads in the low spatial distance situation more based on their destination rather than the high spatial distance ($M_{50m} = 4.60$, SD = 1.50 vs. $M_{3000m} = 3.67$, SD = 1.55; $F(1, 198) = 18.42$; $p < .001$; $\eta^2 = .085$). This provides support for the effective manipulation of spatial distance.

Coupon redemption

We examined the intention to alter planned behavior by consumer's coupon redemption in the low spatial distance (50m away from the City Mall) and high spatial distance (3000m away) conditions. Running an ANOVA with spatial distance (high vs. low) and NFC (high vs. low) as independent variables, and coupon redemption as the dependent variable, showed a significant negative main effect of spatial distance ($M_{low} = 5.06$, SD = 1.04 vs. $M_{high} = 3.41$, SD = 1.11; $F(1,196) = 125.14$; $p < .001$; $\eta_p^2 = .390$), supporting H1. Furthermore, results revealed a marginal positive main effect of NFC ($M_{low} = 4.19$, SD = 1.27 vs. $M_{high} = 4.24$, SD = 1.27; $F(1,196) = 3.90$; $p = .05$; $\eta_p^2 = .020$) on coupon redemption as well. The negative main effect of destination and positive effect of NFC were qualified by a significant interaction effect ($F(1,196) = 4.27$; $p < .05$; $\eta_p^2 = .021$). When participants received an ad far from their destination, those with low NFC expressed higher intention to redeem the coupon ($M_{low} = 3.66$, SD = 1.15) than those with high NFC ($M_{high} = 3.04$, SD = .94; $F(1,100) = 8.21$; $p = .005$; $\eta^2 = .076$). However, this was not significant for those receiving the ad near their destination ($M_{low} = 5.05$, SD = .94 vs. $M_{high} = 5.07$, SD = 1.11; $F(1,96) < 1$; $p > .10$). More importantly, it suggests low NFC would be helpful for ads far from user destination. These findings support the moderating role of NFC on users' intention to redeem coupons (H2).

We also studied the moderation-mediation analysis considering the role of NFC and susceptibility to apps in the near and far destination conditions. We previously expected that when the spatial distance is low and consumers have low NFC levels, being more susceptible to ads leads to higher coupon redemption. A moderation-meditation analysis including spatial distance as the independent variable, NFC as the moderator, susceptibility as the mediator, and coupon redemption as the dependent variable (PROCESS model 7, 5000 bootstrap samples; Hayes, 2013) provided support for the expected effects. More specifically, we found that high NFC (as compared to low NFC) decreased the negative impact of spatial distance on susceptibility to apps (β_{low} = -1.86, SE = .22, $p < .001$ vs. β_{high} = -1.17, SE = .21, $p < .001$). Moreover, higher susceptibility led to more coupon redemption (β = -.144, SE = .07, $p < .05$). Also, susceptibility meditated the effects of spatial distance on coupon redemption. As captured by the index of moderation-mediation (index = .099, SE = .068, 95% CI[.003, .252]), the indirect effects were different between users with low NFC (β_{low} = -.268, SE = .133, 95% CI[-.542, -.016]) and users with high NFC (β_{high} = -.168, SE = .089, 95% CI[-.357, -.009]).

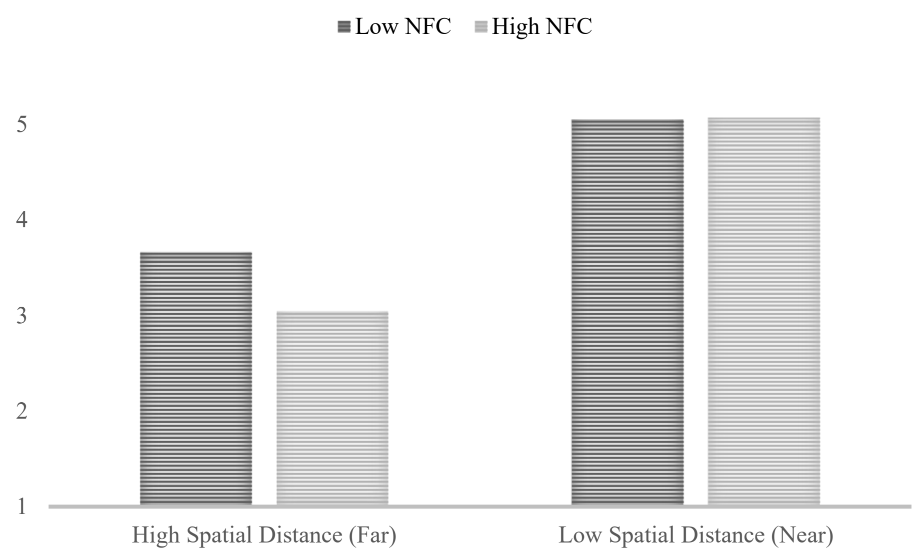


Figure 2.1: Low NFC increases coupon redemption for ads far from the planned destination (high spatial distance), but the interaction effect is not significant for ads near the destination (low spatial distance).

Discussion

Study 1 shows how destination-based ads affect coupon redemption. As predicted, spatial distance had a negative direct effect on consumers' intention to redeem coupons. Results revealed that ads far from consumer destinations could be more helpful for low NFC consumers as it burdened less their cognitive load, but we did not find the same effect for ads far from their destination. Moreover, we found support for the mediating role of susceptibility to apps. When users get destination-based ads, their decision of whether to accept changes in their planned behavior or decline is affected through their level of openness and curiosity. As expected, their susceptibility is depended on their NFC level. The more they enjoy scrutinizing each and every task that they are confronted with, the less chances of being vulnerable would be, and so less chance of altering their planned behavior for the purpose of redeeming the unexpected coupons.

In summary, Study 1 illustrated the negative impact of spatial distance on consumer reactions to destination-based advertising. More importantly, we verified the contingency of this relationship considering the role of NFC, mediated by users' susceptibility. Building on the results of Study 1, more specifically higher coupon redemption for low NFC in ads far from users' destination, one may argue that if apps decrease the cognitive load, they could have a better chance of getting benefits of DBA adoption by their users (we discuss these benefits more in the managerial implications). Therefore, in Study 2, and in line with our theoretical grounding in construal level theory, we investigate the role of cultural distance.

Cultural distance

To create the right customer experience in online settings, the role of culture must be considered (Bagozzi *et al.* 1999, Shobeiri *et al.* 2018). There is evidence suggesting that Canadians (Asians) are more detail-oriented (holistically-minded) as a result of their low (high) construal levels (Nisbett *et al.* 2001). Like evaluations of spatial distances, if consumers perceive a higher cultural distance with ads, they need more cognitive resources to assess them. However, evaluating ads that are culturally matched brings less cognitive load, which makes heuristic low effort processing available (Wilson *et al.* 2013). In terms of the cultural match, previous research finds that cultural incongruency leads consumers to exert more effort to process different kinds of information, i.e. ads (Ko *et al.* 2015).

Therefore, app and website designers should minimize the cognitive load by tailoring the ads and platforms. For example, feelings aroused by cultural background, feelings of entertainment, and likeability of a website for Canadians, and feelings of behavioral control over the website for Chinese customers have not been the same (Mazaheri *et al.* 2011). If customers find DBAs culturally irrelevant to them, even though they receive them in locations near the vendors, it could not have the same positive effects as culturally relevant ones (Lin *et al.* 2016). Therefore, it is suggested that local companies publicize themselves in a culturally congruent way even to minority populations within a country for better targeting. While research shows that Tunisian minority customers are inclined toward web atmospheric cues with a French look (Bartikowski *et al.* 2016), local companies should still be aware of the business opportunities concerning the majority population (Bartikowski and Singh 2014). We speculate that members of a mainstream culture will react differently to app atmospherics with its own cultural symbols as compared to international ones. For example, if a Canadian app contains Canadian elements (i.e. maple leaf), Canadians would be more inclined to it (i.e. adoption of its DBAs) as the design could reduce the perceived cognitive load. Thus, we expect:

H4: *The cultural distance high (vs. low) decreases (increases) the negative effects of spatial distance on consumer reactions.*

Study 2

Study 2 examines the interactive effect of cultural distance and spatial distance on the *planned behavior* (coupon redemption), as well as app reuse intention (H4) in a lab setting with student samples and more importantly with mock apps rather than app screenshots. We also added another product category to see if the product category makes an impact on coupon redemption.

Experimental design

The experimental design was a 2 (cultural distance: low = Canadian culture vs. high = international) × 2 (spatial distance: low = 50m vs. high = 3000m away) between-participants design. Two hundred and eighty-two undergraduate students (53% female) from a major Canadian university came to the research lab, browsed the developed apps on mobile devices, and filled a questionnaire. Spatial distance manipulation was the same as Study 1. Below, we provide details on cultural distance manipulation.

Cultural distance manipulations

We created two experimental mock apps that mimic a ridesharing app in Canada. Both apps were navigable on the cellphones with four pages in each condition (home, welcome, contact us, and destination-based ads). App content in terms of product categories, advertising descriptions, app logic and structure, and others were constant. We manipulated cultural distance as follows: one version (hereafter, the Canadian version) displayed typical cultural markers of the Canadian culture. The other version featured a typical app design without any cultural marker. The Canadian version used the official country's font type, the national colors in the background, as well as depicting one of its major city's downtowns as the header, the country's symbols, and a website address with .ca domain. The international version used non-cultural font (Cave age and Kaushan script font), an ancient street as the header without any association with Canadian culture, a website address with .com domain. Both versions displayed Canadian contact details (physical address, phone numbers) to suggest that the ridesharing app is used in Canada and targets consumers in this area.

We pretested the effectiveness of the cultural distance manipulation with data collected on MTurk from one hundred and fifty Canadian respondents (52% females; 70% were 25 years or older). The study participants evaluated cultural distance concerning Canada for both versions of the experimental app (assigned in random order) using three items ($\alpha > 0.839$) measured on 5-point Likert scales ("This app reflects typical Canadian/international aspects" — "The images, colors, and symbols on this app remind me of a Canadian/international app" — "This app is designed to target Canadian/international consumers in Canada"). The app cultural distance composite scores show that the participants perceived the Canadian version as more culturally congruent with the Canadian culture than the international version ($M_{Canadian} = 4.34$ vs. $M_{international} = 2.71$; $p < .01$). Conversely, they perceived the cultural free version as more international than the Canadian one ($M_{international} = 3.72$ vs. $M_{Canadian} = 2.0$; $p < .01$). The analysis supports the effectiveness of cultural distance manipulation.

Data collection

We removed six incomplete responses along with 17 respondents because of not identifying themselves as a Canadian ($M_{Canadian} = 5.64$, $SD = 1.66$), leaving the sample of two hundred and fifty-nine responses. The respondents were self-identified Canadian undergraduate students (54% female) from a large Canadian university who participated in the research in exchange for partial course credit. The study took place in a research lab, and participants were told that their attitudes towards advertising were being studied. After signing the consent form, they choose between an IOS device (iPhone 7) and an android one (Galaxy S7) to help them better imagine using their own cellphone. Then, they read the same scenario as Study 1.

Next, they randomly entered one of the four conditions. After browsing the assigned app for at least 3 minutes, they scratched the e-coupon and received the same 30% off in all conditions for the two product categories, *clothes and accessories* and *food and drinks* (Appendix B). Then, they filled out the questionnaire. All constructs were measured through a 7-point Likert scale and adapted from existing measurement scales. Cronbach's alpha coefficients validated the reliability of the scales ($\alpha > .73$).

Findings

Participants perceived the ads in the low spatial distance situation more based on their destination rather than the high spatial distance ($M_{50m} = 4.98$, SD = 1.52 vs. $M_{3000m} = 3.53$, SD = 1.64; $F(1,257) = 54.47$; $p < .01$; $\eta^2 = .175$). This shows an effective manipulation of spatial distance. Similarly, results show that the manipulation of cultural distance was successful. Participants reported the psychologically far design was less associated with the Canadian culture than the international one ($M_{\text{Canadian app}} = 4.87$, SD = .98 vs. $M_{\text{international app}} = 3.36$, SD = 1.09; $F(1,257) = 132.04$; $p < .01$; $\eta^2 = .339$).

App reuse intention

We investigated whether the intention to reuse the app varies based on the spatial distance of the communicated ad (low vs. high), and the app's design (Canadian vs. international). An ANOVA with spatial distance (high vs. low) and cultural distance (Canadian vs. international) as independent variables, and app reuse intention as the dependent variable revealed a significant main effect of cultural distance ($M_{\text{Canadian}} = 4.26$, SD = 1.66 vs. $M_{\text{international}} = 3.73$, SD = 1.68; $F(1,255) = 6.018$; $p = .015$; $\eta_p^2 = .014$). We also found a negative main effect of spatial distance ($M_{50m} = 4.12$, SD = 1.70 vs. $M_{3000m} = 3.82$, SD = 1.66) on users' intention to reuse the app but the effect was not statistically significant ($F(1,255) = 2.32$; $p > .10$). We did not find a significant interaction effect ($F(1,255) = 2.12$; $p > .10$). The results confirmed the moderating role of cultural distance on customers' intention to use the app in the future. Surprisingly, the role of receiving a destination congruent ad to persuade in using the app in the future was not supported. This provides important managerial implications that we discuss later.

Coupon redemption

Finally, we checked the effects of spatial distance and cultural distance on the user's decision to redeem the coupons. We previously predicted that participants who received low spatial distance (vs. high) ad and browsed low cultural distance app (vs. high) would be more willing to embrace the ad and redeem the coupon toward their purchase. A repeated-measure ANOVA with product category (clothes and accessories vs. foods and drinks) as the within-participants factor, coupon redemption as the dependent variable, and spatial distance and cultural distance as between-participants factors confirmed the prediction. Those who received an ad near their destination (50m away) were more likely to redeem the coupon than participants offered an ad far from their destination ($M_{50m} = 5.58$, SD = 1.40 vs. $M_{3000m} = 4.93$, SD = 1.64; $F(1,255) = 11.431$; $p = .001$; $\eta^2 = .043$). Moreover, the Canadian version of the app led to higher coupon redemption than the international version ($M_{\text{Canadian}} = 5.49$, SD = 1.49 vs. $M_{\text{international}} = 5.08$, SD = 1.61; $F(1,255) = 4.286$; $p < .05$; $\eta_p^2 = .017$). Also, the interaction effect of the product category with either spatial distance or cultural distance was not significant ($F(1,255) = .792$; $p > .30$, and $F(1,255) = .14$; $p > .70$, respectively). This shows that coupon redemption does not vary as a function of spatial distance or cultural distance among the two product categories. Similarly, the interaction effect of cultural distance and spatial distance was not significant ($F(1,255) = .154$; $p > .60$). However, for those who received the ad near their destination, we found a significant effect of cultural distance on coupon redemption ($M_{\text{international}} = 5.37$, SD = 1.59 vs. $M_{\text{Canadian}} = 5.84$, SD = 1.19; $F(1,130) = 3.61$; $p = .03$; $\eta_p^2 = .027$). This effect was not significant for those who received the high spatial distance ads ($M_{\text{international}} = 4.79$, SD = 1.59 vs. $M_{\text{Canadian}} = 5.12$, SD = 1.69; $F(1,125) = 1.20$; $p = .275$; $\eta_p^2 = .010$).

Therefore, designing a culturally congruent app is critical for the enjoyment of the ridesharing experience, and further brings higher coupon redemption at the advertised destination-based locations. However, this is only the case for low spatial distance ads and it does not have an impact on ads far from the planned destinations.

Discussion

Study 2 shows how destination-based ads affect app reuse intention and coupon redemption. We checked whether sending destination-congruent ads and providing users with a culturally congruent app design moderates the effects. We expected that lower spatial distance and cultural distance need less cognitive resources to process the information. We found that for the effects on app reuse intention, users of the Canadian version of the ridesharing app, who received an ad near their destination, had more favorable attitudes compared to those who used the international version. That was also the case for the user's intention to redeem the coupon. However, we did not find a significant interaction effect of spatial distance and cultural distance on coupon redemption for those receiving the ad far from their destination. In other words, cultural distance makes the ads near the users' destination more appealing and heightens the actual change in the planned behavior (to redeem the coupon). However, it does not offset the abstraction of ads in high spatial distances (3000m away).

In summary, Study 2 illustrated the negative impact of spatial distance on consumer reactions to destination-based advertising. More importantly, we verified how designing culturally congruent apps could be beneficial. While the cultural distance enhanced the effects of low spatial distance ads (50m away), it did not have significant effects for ads with high spatial distance (3000m away). However, in the context of destination-based ads, it is equally important to find ways to make ads far from planned destinations interesting enough for customers to convince them to head towards the unplanned destination. That would lead users to alter their destinations to the advertised one, even if it is far from their planned destination. In Study 3, and in line with our rideshare context, we investigate if some benefits (i.e. free rerouting) could make a high spatial distance ad appealing to ridesharing users. While receiving an ad with a new destination comes with some sort of surprise, we expect users to have more favorable attitudes, intentions, and willingness to accept the ad (via altering their planned destination and redeeming the coupon). Therefore, we hypothesize that:

H5: *Marketing incentives such as free rerouting have positive effects on consumer reactions.*

Thus, Study 3 tests the moderating effect of marketing incentives (H5), and the interaction between cultural distance and marketing incentives in order to attract users to the ads with high spatial distance.

Study 3

In Study 3, we examined the interaction effect of cultural distance and marketing incentives on attitudes towards destination-based ads, purchase intention, app reuse intention, and coupon redemption of ads that are far from users' destination, and investigated if they are interesting enough to alter their destinations.

Experimental design

The experimental design was a 2 (cultural distance: low = Canadian vs. high = international) × 3 (marketing incentives: high = 30% off coupon with free rerouting vs. low = 30% off coupon without free rerouting vs. control = no information provided) between- participants design. Four hundred and thirty-one undergraduate students (54% female) from a major Canadian university came to the research lab, browsed the developed apps on mobile devices, and filled a questionnaire.

Data collection

We removed three incomplete responses along with nine respondents because of not identifying themselves as Canadian ($M_{Canadian}$= 5.66, SD = 1.62), leaving a sample of four hundred and nineteen responses. The respondents were undergraduate students (55% female) from a large Canadian university and participated in the research in exchange for partial course credit. The research design was the same as for Study 2 except the fact that there has been no low spatial distance condition considered in Study 3.

Findings

First, we verified whether the manipulations of cultural distance and marketing incentives were successful. Participants perceived the ads in the Canadian version more culturally congruent with the Canadian culture than the international one ($M_{Canadian\ app}$ = 4.97, SD = .96 vs. $M_{international\ app}$ = 3.33, SD = 1.00; $F(1,417) = 289.70$; $p < .001$; $\eta^2 = .410$). This shows a successful manipulation of cultural distance. To verify the marketing incentives manipulation, we measured the level of deal proneness. Deal proneness is not only related to price attention, but it also includes shopping smartly and getting a better deal on a quality item (Steenkamp and Maydeu-Olivares 2015). Those who received more incentives (the coupon and free rerouting) expressed a higher sense of deal proneness (Lichtenstein *et al.* 1995) than the respondents in the less incentives (the coupon and non-free rerouting) and control conditions ($M_{non\text{-}free\ rerouting}$ = 5.25, SD = 1.37 vs. $M_{control}$ = 5.31, SD = 1.38 vs. $M_{free\ rerouting}$ = 5.62, SD = 1.15; $F(2,416) = 3.21$; $p < .05$; $\eta_p^2 = .015$), thus the incentive manipulation was successful.

Attitudes towards destination-based ads

We studied the attitudes towards destination-based ads in three conditions (marketing incentives: high = 30% off coupon with free rerouting vs. low = 30% off coupon without free rerouting vs. control = no information provided). Running an ANOVA with marketing incentive (high vs. low vs. control) and cultural distance (Canadian vs. international) as independent variables and attitudes towards the destination-based ads as the dependent variable provided a significant main effect of marketing incentives ($M_{non\text{-}free\ rerouting}$ = 4.30, SD = 1.37 vs. $M_{control}$ = 4.22, SD = 1.33 vs. $M_{free\ rerouting}$ = 4.60, SD = 1.30; $F(2,413) = 3.60$; $p < .05$; $\eta_p^2 = .017$).

Furthermore, findings revealed a negative main effect of cultural distance ($M_{Canadian}$ = 4.52, SD = 1.33 vs. $M_{international}$ = 4.26, SD = 1.34; $F(1,413) = 3.45$; $p = .032$; $\eta_p^2 = .016$) on attitudes toward destination-based ads. The main effects of incentives and culture were qualified by a significant interaction effect ($F(1,255) = 3.18$; $p = .045$; $\eta^2 = .012$). Participants who used the Canadian version of the app had more favorable attitudes toward the destination-based ad with the free rerouting offer than those who were asked to pay for the rerouting and the control conditions ($M_{non\text{-}free\ rerouting}$ = 4.32, SD = 1.38 vs. $M_{control}$ = 4.24, SD = 1.43 vs. $M_{free\ rerouting}$ = 4.94, SD = 1.11; $F(2,189) = 5.61$; $p = .004$; $\eta_p^2 = .056$). However, this was not significant for those receiving the ad on the international version of the app ($F(2,224) < 1$; $p > .10$). Figure 2.2 shows these results. Therefore, the findings validated the significant role of marketing incentives on how users feel about ads that are based on their planned destination, but far from it. We also found a negative main effect of cultural distance in this matter. More importantly, cultural distance significantly heightened the effects of marketing incentives for the app that was culturally congruent with the Canadian culture. This comes from the lower psychological distance of culturally congruent apps that helps users perceive the advantage of incentives with less cognitive load.

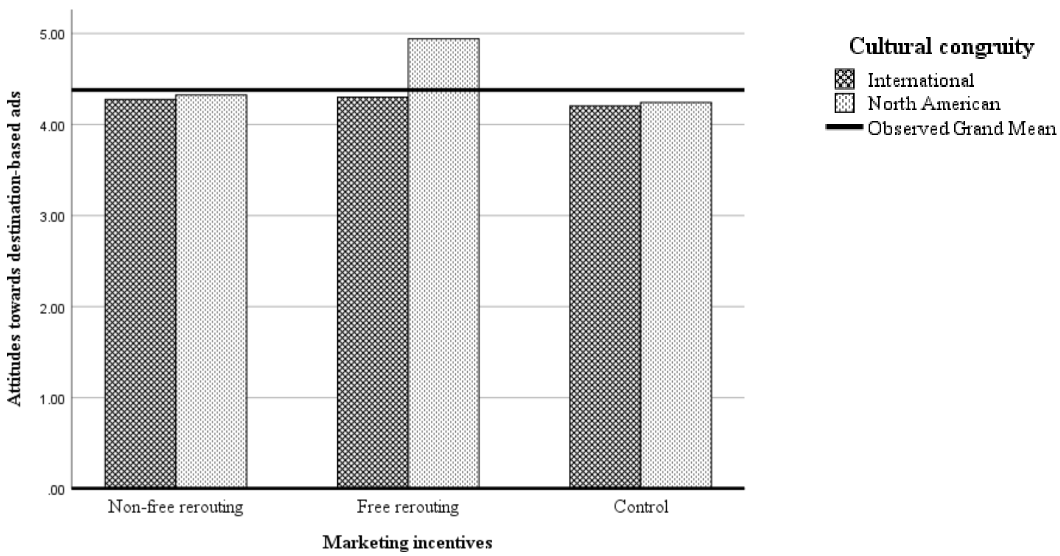


Figure 2.2: An app with low cultural distance increases attitudes towards destination-based ads for ads with higher marketing incentives, but the interaction effect is not significant for non-free rerouting ads or the control group.

Purchase intention

We investigated the purchase intentions from the advertised store in three marketing incentives' conditions. An ANOVA analysis with marketing incentives (high vs. low vs. control) and cultural distance (Canadian vs. international) as independent variables and purchase intention as the dependent variable did not find a significant main effect of marketing incentives, while the purchase intention was slightly higher for the high marketing incentive condition ($M_{non\text{-}free\ rerouting}$ = 3.43, SD = 1.35 vs. $M_{control}$ = 3.43, SD = 1.23 vs. $M_{free\ rerouting}$ = 3.70, SD = 1.40; $F(2,413) = 2.32$; $p = .10$; $\eta_p^2 = .011$). We also obtained a significant main effect of cultural distance ($M_{Canadian}$ = 3.70, SD = 1.39 vs. $M_{international}$ = 3.37, SD = 1.28; $F(1,413) = 5.43$; $p = .020$; $\eta^2 = .013$). Furthermore, the results showed a significant interaction effect of marketing incentives and cultural distance ($F(2,413) = 4.43$; $p = .013$; $\eta_p^2 = .021$) on purchase intention. More specifically, participants who *(1)* received more marketing incentives (both the 30% off and the free rerouting), and *(2)* used the Canadian version of the ridesharing app had higher intention to purchase from the advertised locations than those using the international version ($M_{Canadian} = 4.12$, SD = 1.39 vs. $M_{international} = 3.59$, SD = 1.40; $F(2,189) = 5.44$; $p = .005$; $\eta^2 = .054$). However, the interaction effect was not significant for the ads with fewer marketing incentives and the control condition. Therefore, the findings provide support for the role of cultural distance and marketing incentives on purchase intentions. The interaction effect also showed the positive moderating effect of cultural distance for the marketing incentive ads. Like Study 2, we showed that providing users with a culturally congruent app plays an important role in altering planned destinations (even to far locations), which leads to higher purchase intention.

App reuse intention

Next, we examined the effects of marketing incentives and cultural distance on users' intention to reuse the app in the future. An ANOVA with marketing incentives (high vs. low vs. control) and cultural distance (Canadian vs. international) as independent variables and app reuse intention as the dependent variable revealed that the lower cultural distance is, the higher is participants' intention to reuse the app in the future are ($M_{Canadian} = 4.12$, SD = 1.58 vs. $M_{international} = 3.81$, SD = 1.58; $F(1,413) = 3.65$; $p = .023$; $\eta_p^2 = .009$). Moreover, we found a positive main effect of marketing incentives ($M_{non\text{-}free\ rerouting} = 3.67$, SD = 1.62 vs. $M_{control} = 3.82$, SD = 1.67 vs. $M_{free\ rerouting} = 4.36$, SD = 1.49; $F(2,413) = 8.29$; $p < .001$; $\eta_p^2 = .039$) on users' intention to reuse the app. The positive main effects of cultural distance and marketing incentives were qualified by a significant interaction effect ($F(2,413) = 3.29$; $p = .038$; $\eta_p^2 = .016$). Participants who received the free rerouting offer (high marketing incentives) and were using the Canadian version of the ridesharing app expressed higher intention to reuse the app ($M_{Canadian} = 4.79$, SD = 1.32) than those using the international version ($M_{international} = 3.98$, SD = 1.53; $F(1,144) = 11.51$; $p = .001$; $\eta^2 = .074$). However, the interaction effect was not significant for either those who did not receive the free rerouting offer (low marketing incentives) or the control group ($F(1,125) < 1$; $p > .10$). Furthermore, the results showed that for participants who used the international version of the app, marketing incentives did not significantly increase the app reuse intention ($M_{non\text{-}free\ rerouting} = 3.73$, SD = 1.60 vs. $M_{control} = 3.72$, SD = 1.62 vs. $M_{free\ rerouting} = 3.98$, SD = 1.53; $F(2,224) = .661$; $p > .10$; $\eta^2 = .06$). The results confirmed the moderating role as well as the interaction effect of cultural distance and marketing incentives on customers' intention to reuse the app in the future.

Coupon redemption

We tested whether the decision to alter the planned destination from the City Mall to a location that is 3000m further varies based on different levels of marketing incentives and cultural distance. We expected participants who received low cultural distance (vs. high cultural distance) ads and those who received higher marketing incentives (free rerouting vs. non-free rerouting) would be more willing to alter their destination and redeem the coupon towards their purchase. A repeated-measure ANOVA with product category (clothes and accessories vs. food and drink) as the within-participants factor, coupon redemption as the dependent variable, and marketing incentives and cultural distance as between-participants factor supported our hypothesis. Participants who received an ad with higher marketing incentives (30% off plus the free rerouting offer) were more willing to redeem the coupon at a distance of 3000m from their planned destination than the control group, as well as participants offered an ad coming with fewer marketing incentives (just 30% off) ($M_{non\text{-}free\ rerouting}$ = 4.45, SD = 1.78 vs. $M_{control}$ = 5.17, SD = 1.81 vs. $M_{free\ rerouting}$ = 5.71, SD = 1.36; $F(2,413) = 26.843$; $p < .001$; $\eta_p^2 = .115$). We also found no significant effect of cultural distance on coupon redemption ($M_{Canadian}$ = 5.02, SD = 1.59 vs. $M_{international}$ = 4.79, SD = 1.57; $F(1,413) = 2.521$; $p = .113$; $\eta_p^2 = .006$) which replicates the findings of Study 2 that showed cultural distance was not effective for ads far from users' destination. More importantly, the product category did not have a significant interaction effect with marketing incentives ($F(2,413) = 1.300$; $p > .20$). This shows that coupon redemption does not vary as a function of the moderator among the two product categories.

Table 2.1: Main effects and interactions for Study 3

		Attitudes toward DBA	**Purchase intention**	**App reuse intention**	**Coupon redemption**
Cultural Distance	**Marketing incentives**				
Main effects					
-	Low (non-free rerouting)	4.30 *(1.37)*	3.43 *(1.35)*	3.67 *(1.62)*	4.45 *(1.78)*
-	Control	4.22 *(1.33)*	3.43 *(1.23)*	3.82 *(1.67)*	5.17 *(1.81)*
-	High (free rerouting)	4.60 *(1.30)*	3.70 *(1.40)*	4.36 *(1.49)*	5.71 *(1.36)*
Low (Canada)	-	4.52 *(1.33)*	3.70 *(1.39)*	4.12 *(1.58)*	5.02 *(1.59)*
High (International)	-	4.26 *(1.34)*	3.37 *(1.28)*	3.81 *(1.58)*	4.79 *(1.57)*
Interactions					
Low (Canada)	Low (non-free rerouting)	4.32 *(1.38)*	3.59 *(1.40)*	3.60 *(1.66)*	4.20 *(1.59)*
	Control	4.24 *(1.43)*	3.34 *(1.25)*	3.93 *(1.73)*	5.12 *(1.69)*
	High (free rerouting)	4.94 *(1.11)*	4.12 *(1.39)*	4.79 *(1.33)*	5.75 *(1.06)*
High (International)	Low (non-free rerouting)	4.28 *(1.37)*	3.30 *(1.30)*	3.73 *(1.60)*	4.28 *(1.63)*
	Control	4.21 *(1.26)*	3.51 *(1.22)*	3.73 *(1.62)*	4.80 *(1.59)*
	High (free rerouting)	4.30 *(1.38)*	3.33 *(1.31)*	3.98 *(1.53)*	5.29 *(1.32)*

Mean *(SD)*

Discussion

Study 3 illustrates how an advertiser can effectively provide users with ads far from their planned destinations. We studied the role of two moderators, a relevant incentive in the context of the ridesharing industry, by offering free rerouting along with designing culturally low distant apps. Intriguingly, comparing the results of the control condition and non-free rerouting offer revealed that users are less interested in the destination-based ads once they were explicitly asked to pay for going to the alternative destination. However, results showed that attitudes toward destination-based ads would increase even for ads far from planned destinations if it comes with higher marketing incentives. Using the Canadian version of the app led to higher positive attitudes toward the destination-based ad for users who received the free rerouting offer than for those who were asked to pay for the rerouting, and the control group. Users of the Canadian version of the ridesharing app, who received more marketing incentives (both the 30% off and the free rerouting) had more favorable attitudes compared to those who used the international version. Similarly, the purchase intention was higher for users who received more marketing incentives and used the Canadian version of the ridesharing app than those using the international version. Regarding intention to reuse the app, only if a free rerouting offer gets advertised on a culturally congruent app, will users express higher intention to come back to the app in the future. Like study 2, there is no effect of cultural distance in the no-free rerouting and control conditions. Also, it does not increase the redemption of ads far from the user destination. However, higher marketing incentives would successfully alter planned destinations to redeem the coupons.

In summary, Study 3 showed that changes in the relationships between the attitudes, intentions, and behavior for high spatial ads (3000m away) are based on the cultural distance and marketing incentives. In line with the findings of Study 2, we found no significant effect of cultural distance on coupon redemption for ads that are far from the destination. Results showed that the free rerouting offer has no effect on attitudes toward ads, purchase intention, and app reuse intention in the conditions with the international app design. However, it has positive effects for apps with the Canadian design (= low cultural distance). We also illustrated how providing a free rerouting offer could bring opportunities for a firm like Uber to advertise not only near a destination, but far from it. In this way, ridesharing companies not only get benefits from riding fees, but they also can make money out of destination-based advertisements even for places thousands of meters away. However, they may be still undecided about which strategy to use, advertising near the planned destination, or advertising far from it. In Study 4, we examine whether these two strategies differ.

Study 4

Study 4 compares the effectiveness of destination-based advertising near the planned destination and far from it when the latter comes with marketing incentives, but no incentive is provided with the ads near the planned destination.

Experimental design

The experimental design was a 2 (spatial distance-incentive: near; low incentives vs. far; high incentives) × 2 (cultural distance: Canadian culture vs. international) between- participants design. We did not consider the two conditions with predictable results (near; high incentives vs. far; low incentives) and focused on the conditions that might cause a dilemma for practitioners. Two hundred and eighty-six undergraduate students (54% female) from a major Canadian university came to the research lab, used the developed apps on mobile devices, and then filled a questionnaire.

Data collection

We excluded one incomplete response along with seven respondents because of not identifying themselves as Canadian ($M_{Canadian}$= 5.57, SD = 1.56), leaving a sample of 278 responses. The respondents were undergraduate students (54% female) from a large Canadian university who participated in the research in exchange for partial course credit. The research design was the same as for Study 2 and Study 3.

Findings

Before testing for the effects of cultural distance and destination-incentive, we checked if the three manipulations were successful. Like previous studies, participants perceived the culturally congruent app (Canadian version) more associated with the Canadian culture than the culturally incongruent app (international version) ($M_{Canadian\ app} = 4.93$, SD= .94 vs. $M_{international\ app} = 3.35$, SD = 1.14; $F(1,276) = 154.10$; $p < .001$; $\eta^2 = .358$). Moreover, participants reported the first type of ads (the low spatial distance without higher marketing incentives situation; 50m away) more based on their destination rather than the other type of ads high spatial with higher marketing incentives one (3000m away) — ($M_{near;\ low\ incentives} = 4.90$, SD = 1.70 vs. $M_{far;\ with\ high\ incentives} = 3.81$, SD = 1.71; $F(1,276) = 28.36$; $p < .001$; $\eta^2 = .093$). Therefore, the manipulation of spatial distance was successful. Finally, respondents who received higher marketing incentives (the coupon and free rerouting) far from their destination expressed a higher sense of deal proneness than the respondents who did not receive such an offer near their destination ($M_{near;\ low\ incentives} = 5.70$, SD = .91 vs. $M_{far;\ with\ high\ incentives} = 5.18$, SD = 1.06; $F(1,276) = 18.94$; $p < .001$; $\eta^2 = .015$) providing support for the successful manipulation of the incentives.

Cultural distance as the first moderator

We examined how the attitudes towards destination-based ads, intentions (for purchase and reusing the app), and behavior (coupon redemption) changed based on the effects of cultural distance. Like Study 2 and Study 3, results showed that the main effects for all of the four constructs were significantly positive. Attitudes towards destination-based ads, purchase intention, app reuse intention, and coupon redemption were higher for the culturally congruent app (the Canadian version) than the culturally incongruent app (international version) ($p < .001$). The results provide further support for the effects of cultural distance in this regard.

Spatial distance -incentive as the second moderator

We tested whether the combination of spatial distance and marketing incentives (50m without higher marketing incentives vs. 3000m with higher marketing incentives) makes a difference on any of the four constructs. Interestingly, we found none of the main effects of the four elements significant. More specifically, attitudes towards destination-based ads, purchase intention, and coupon redemption were higher for destination-congruent ads without higher marketing incentives than for destination-incongruent ads with higher marketing incentives. However, the difference was not statistically significant ($p > .6$). We also found the app reuse intention was higher for the ads 3000m away with higher marketing incentives, but the effect was not significant ($p > .2$) (Table 2.2).

Table 2.2: Main effects and interactions for Study 4

		Attitudes toward DBA	Purchase intention	App reuse intention	Coupon redemption
Cultural Distance	**Spatial Distance-Incentive**				
Main effects					
-	Near; low incentives	4.63 *(1.35)*	3.77 *(1.40)*	4.12 *(1.71)*	5.58 *(1.43)*
-	Far; with high incentives	4.60 *(1.30)*	3.70 *(1.40)*	4.36 *(1.49)*	5.50 *(1.22)*
Low (Canada)	-	4.32 *(1.32)*	3.37 *(1.32)*	3.87 *(1.64)*	5.33 *(1.45)*
High (International)	-	4.95 *(1.24)*	4.15 *(1.37)*	4.68 *(1.43)*	5.79 *(1.12)*
Interactions					
Low (Canada)	Near; low incentives	4.97 (1.39)	4.19 *(1.36)*	4.55 *(1.55)*	5.80 *(1.24)*
	Far; with high incentives	4.94 *(1.11)*	4.12 *(1.39)*	4.79 *(1.32)*	5.75 (1.06)
High (International)	Near; no incentives	4.34 *(1.26)*	3.41 *(1.34)*	3.74 *(1.76)*	5.37 *(1.59)*
	Far; with incentives	4.30 *(1.38)*	3.33 *(1.31)*	3.98 *(1.53)*	5.33 *(1.27)*

Mean *(SD)*

Discussion

In study 4, we mixed spatial distance with incentives and focused only on the two conditions that practitioners may doubt for which to choose. Study 4 showed that users perceive a balanced trade-off between the value of the incentive and the cost of heading to a far destination. Thus, firms in the ridesharing industry would see no significant difference in users' feelings, intentions, and behavior by providing ads near the planned destination or far from it once the latter came with marketing incentives. In other words, they have similar choices of collaborating with destination-based advertisers wherever their location is if they provide higher marketing incentives (i.e. free rerouting) for locations that are far from the planned destinations. Therefore, there is no difference in choosing either of the advertising strategies, advertising near the planned destinations or far from it once it comes with some sort of incentives.

GENERAL DISCUSSION

The location-based service of mobile communication has been identified as one of the most important trends in the development of electronic commerce (Kuo, Chen, and Liang 2009). Four studies examined the impact of spatial distance, NFC, cultural distance, and marketing incentives, the underlying mechanism of susceptibility, and how altering planned behavior happens by advertising based on users' destinations. Findings in Study 1 show that the level of spatial distance proximity (farness) of the ad to the planned destination enhances (diminishes) keeping the planned behavior. Moreover, we found that having a low NFC not only helps avoid the situation of in-depth thinking about altering planned behavior, but also decreases the negative impact of high (vs. low) spatial distance on coupon redemption. This interaction between spatial distance and NFC is found via the level of curiosity and openness to the apps with the destination-based ads.

Building on Study 1 findings on the role of cognition load, Study 2 projects that lowering psychological distance by decreasing cultural distance (Canadian version) could make the ads near the user's destination more appealing. Receiving an ad near (50m) the destination was no different in terms of intention to reuse the app in the future with ads far from it. This means that ridesharing apps should not worry about the retention of their users if they are exposed to ads far from their destination. We also found that the users of the Canadian version of the app expressed higher intention to reuse the app and higher coupon redemption than those using the international version since the Canadian design was perceived as more culturally congruent. More importantly, those who get the ad near their destination would have higher coupon redemption when using the Canadian version than the international one. However, these three interaction effects were not significant for ads far from their destination. In other words, cultural distance does not play the same role for ads near vs. far from users' destinations.

Thus, in Study 3, we explored ways to make high spatial ads (3000m) more attractive for users to alter their planned destinations to the advertised ones. Sending higher marketing incentives (i.e. discounts plus a free rerouting offer) enhances attitudes toward destination-based ads, purchase intention, and coupon redemption more than only offering discounts (control group) and lower marketing incentives (i.e. discounts without a free rerouting offer). However, there is no difference in terms of intention to reuse the app. In fact, while providing the option of going to a new destination is something considerable for users (either with or without more incentives), there is no harm to the rideshare app's reuse intention. This type of ad enhances users' favorable attitudes toward destination-based ads, purchase intention, and intention to reuse the app when the Canadian version of the app is used compared to the international version. Furthermore, participants who used the Canadian version of the app have more favorable attitudes toward the destination-based ad if it came with the free rerouting offer, than those who are asked to pay for the rerouting and the control conditions. However, this is not significant for those receiving the ad on the international version of the app. Participants who receive more marketing incentives (both the 30% off and the free rerouting) and use the Canadian version of the ridesharing app have more intention to purchase from the advertised locations than those using the international version. Moreover, the users of the Canadian version express a higher intention to reuse the app than those using the international version. However, the interaction effect is not significant for either those who do not receive the free rerouting offer (low marketing incentives) or the control group. The result broadly shows that cultural distance outweighs the positive effects of incentives.

In Study 4, we found that there is no moderating role for the combination of spatial distance-incentive if destination incongruent ads came with higher marketing incentives, and destination congruent ads came without higher marketing incentives. More specifically, the findings show that none of the main effects of the four variables (attitudes towards destination-based ads, purchase intention, app reuse intention, and coupon redemption) are significant. In other words, users do not find any difference between getting ads near their destination or far from it when the destination-incongruent ads have higher marketing incentives. Therefore, they are equally interested in both types and have the same desire towards them as the different levels in the spatial distance are offset by the benefits of higher incentives. This has important practical implications that we mention in the managerial implications section. Table 2.3 summarizes the three studies and key findings.

Table 2.3: Summary of the studies

Study	Purpose	Key findings
Study 1	Examining the interactive effects of spatial distance and NFC on consumer reactions to DBA (H1-H2) Studying the mediation roe of susceptibility and looking into changes in the mediation effect in high vs. low NFC levels	✓ High spatial distance ads redeemed more for low NFC consumers, but we did not find the same effect for ads near their destination. ✓ When users get DBAs, their decision whether to alter their planned behavior is transmitted through their level of openness and curiosity (susceptibility).
Study 2	Investigating if decreasing the cognition load (by cultural cultural-laden app design), could lead to higher DBA adoption by the users	✓ Users who receive an ad near (50m away) their destination, have higher coupon redemption than those who receive an ad far from their destination (3000m away). However, there is no difference in terms of app reuse intention in the future. ✓ Users who use the Canadian version of the app, express more intention to reuse the app, and higher coupon redemption than those who browse the international version. ✓ Those who get the ad near their destination would have higher coupon redemption when browsing the Canadian version than when using the international one. However, the interaction effects are not significant for ads far from the planned destination.
Study 3	Making destination-incongruent ads more attractive for customers to alter destinations to the advertised one, even if the ad is far from their planned destination.	✓ Users who get higher marketing incentives have more favorable attitudes toward destination-based ads, higher app reuse intention, and higher coupon redemption than the control group as well as those who receive an ad with less marketing incentives. However, there is no difference in terms of intention to purchase intention. ✓ For ads far from the planned destination, participants who use the Canadian version of the app, express more favorable attitudes toward destination-based ads, higher purchase intention, and more intention to reuse the app than those who browse the international version. ✓ Participants who browse the Canadian version of the app have higher favorable attitudes toward the destination-based ad once it comes with the free rerouting offer than those who are asked to pay for the rerouting and control conditions. However, this is not significant for those receiving the ad on the international version of the app. ✓ Participants who receive more marketing incentives (both the 30% off and the free rerouting) and use the Canadian version of the ridesharing app have more intention to purchase from the advertised locations than those using the international version. Moreover, that group expresses a higher intention to reuse the app than those using the international version. However, the interaction effect in both studies is not significant for either those who do not receive the free rerouting offer (low marketing incentives) or the control group.
Study 4	Studying the moderating role of spatial distance-incentive to see if ads far from the planned destination differ from ads near the planned destination if only the earlier one comes with higher marketing incentives.	✓ None of the main effects of the four dependent variables (attitudes towards destination-based ads, purchase intention, app reuse intention, and coupon redemption) was significant. So, there was no statistically significant difference between low spatial distance ads without higher marketing incentives and high spatial distance with higher marketing incentives.

Theoretical implications

Theoretically, this research is the first one to examine location-based advertising not for the current location but for the destination. The research examines the impact of construal level theory on altering planned behaviors (Ajzen 1985) by taking into account the effects of NFC, susceptibility, cultural distance, and marketing incentives. Nysveen *et al.* (2005) considered the perceived enjoyment of using a mobile service as an important intrinsic motivation for behavioral intention toward mobile services adoption. In fact, we do not expect the unplanned behavior, i.e. coupon redemption, to happen if users keep their planned destination because of its high cognitive load. We contributed to the literature by showing that lowering this burden could happen by decreasing psychological distances (via lowered spatial and cultural distances), as well as by providing relevant incentives (i.e. free rerouting). It is also found that interactive effects of spatial distance and NFC led to altering planned behavior via susceptibility to apps. When the ads are about near destinations, consumers who are interested in bypassing puzzles, have more positive reactions due to their enhanced susceptibility to being persuaded by the ads. Finally, this study could facilitate further inquiries into the field of recommender systems and targeted ads (Goldfarb and Tucker 2011).

Managerial implications

The level of satisfaction determines whether consumers would reuse a product or a service (Alalwan 2020; Park 2019; Shirdastian *et al.* 2019). As ridesharing apps (i.e. Uber) have access to big datasets of locations and destinations, they could utilize it to foster their user's journey experience. To this end, they need to lessen the amount of information the consumers should process to decide whether adopting this ad is what they would be interested in. For example, by conjoining *mechanical*, *thinking*, and *feeling AI* they can first collect individual-level data on usage and experience, emotions, social networks, consideration heuristics, locations, destinations, and past trips (Huang and Rust 2021). Then by optimizing AI algorithms, they can profile and cluster their users in order to simultaneously individualize their app design (i.e. cultural atmospherics) and offerings (both in terms of where to propose as the alternative destination and what kind of incentive to offer) to enhance the chance of adoption.

Our research suggests that these companies would be able to collaborate with advertising companies to offer destination-based ads. Even if the offer is far from the user destination, a relevant benefit, i.e. free rerouting, helps alter preplanned destinations. The findings also show that sending ads far from a user's destination would not decrease the intention to reuse the app, and so it does not negatively affect their existing business of ridesharing.

Limitations and future research

This study has some limitations that offer opportunities for future research. First, in Study 2, 3, and 4 we provided the option of choosing from both popular mobile device brands, iPhone and Samsung. This helped users perceive the device as their own. However, since there was no chance of personalizing the device, there might still be some effects of not using their own mobile phones. Therefore, their level of personal attachments to mobile devices may have an impact on their attitudes to engage in mobile marketing activities (Sultan *et al.* 2009). Future researchers could ask the participants to run the app on their own devices and conduct the experiment, *if* participants are assured that there would be no privacy concerns about installing the app. Second, for cases where people could use both desktops and mobile devices (i.e.

booking on Airbnb from a computer or on a mobile device), the role of planning ahead (vs. when on the go) is also worthy of research. Third, although we developed the mock apps, gathering data in a lab setting might have impacts in a real setting. The next step is conducting an in-field experiment to validate the findings.

PAPER 3: FROM USER GENERATED CONTENTS TO AUTOMATIC CONTENT GENERATION: HOW SALES DESCRIPTION CONCRETENESS IMPACTS SOCIAL COGNITION OF SERVICE PROVIDERS

ABSTRACT

Building on the speech act theory, we examined how sale descriptions have an impact on consumers' social cognitions of service providers and how they could be used to generate new content. Previous studies only focused on general linguistic characteristics. However, in this research, we added to the social cognition theory by examining the combination of warmth and competence in an in-field observation of user-generated content in the sharing economy context. The dataset contained 398,926 reviews for 11,678 properties, located in a North American city (scrapped in November 2021 by Inside Airbnb). In Study 1, we showed the role of linguistic concreteness, service provider type, and sentiment analysis on the perceived level of warmth and competence of the service providers. In Study 2, we used two text mining methods (frequency analysis and modified LDA) to better understand the differences between the four kinds of property descriptions and their hosts. In Study 3, we employed the findings of Study 1 and Study 2 to better train the Natural language generation (NLG) algorithm and showed how Long-Short-Term Memory (LSTM), as well as cleaned input, could help service providers generate more engaging content. The research not only contributes to more easily incorporate linguistic theories in marketing but also advances the literature in terms of understanding how textual communications influence user perceptions. It could also help develop more accurate NLG algorithms to communicate.

INTRODUCTION

Knowing the diversity of customers' attitudes about products and services and their providers helps marketers design or modify their product/service to address the needs of their customer profile. Inferring consumers' preferences, interests, and needs, disclosed through their online activities, has a broad kind of applications. By having access to the motivations and interests of the customers, service providers can either adapt their services or segment their market (Farhadloo *et al.* 2016). As reviews show a significant role in projecting sales volume (Woolley and Sharif 2021), there is a large body of literature concerning how to analyze and use reviews and other user-generated content to gain marketing insights (Shirdastian *et al.* 2017). However, still little is known about how sale descriptions have an impact on consumers' social cognitions of service providers and how they could be used to generate new content.

Previous studies only focused on certain general linguistic characteristics such as sentiment, precision, and concreteness without going into details of its impact on consumer perceptions and cognitions. However, in this research, we added to the social cognition theory by examining the combination of warmth and competence in an in-field observation of user-generated content in the sharing economy context. In Study 1, we were concerned with different types of service providers (amateurs vs. professionals) and how certain ways of communication and contents (i.e. whether abstract or concrete) led to different social cognition attitudes towards the service providers via the level of their communication sentiment. We discovered that in abstract descriptions, amateurs used a more positive sentiment description than professionals. Not only did the results suggest that higher positive listing sentiment led to more positive social cognition but also it fully meditated the effects of listing concreteness on social cognition. The findings showed the role of linguistic concreteness and sentiment analysis on both social cognition dimensions for service providers.

In Study 2, using two text mining methods (frequency analysis and modified LDA), we investigated the differences between the four kinds of property descriptions and their hosts by looking into the distinct tokens for each group. We found that, for example, abstract amateurs (novices) were mainly renting out their own spare spaces (i.e. basements), to-be-professionals understood the importance of using concrete language (i.e. fully equipped kitchen), juniors acted more similarly to novices in terms of abstractly communicating, while real-estate agents addressed more information about the units by utilizing words like fully equipped and living area. Moreover, the modified LDA extracted the relevant tokens to social cognition dimensions from property descriptions. Results showed that amateur hosts were more inclined to describe their competence via the features of their property while professionals highly emphasized their own hospitality capabilities. Surprisingly, we did not find a statistically significant difference in terms of property-related (vs. host-related tokens) for the warmth dimension.

While there have been few empirical tests to apply Natural language generation (NLG) in marketing contexts, in Study 3, we applied neural networks, i.e. Bidirectional Long-Short-Term-Memory (BLSTM), to train the model on the existing Airbnb descriptions in order to generate contents accordingly. More specifically, using the findings of Study 1 and Study 2, we compared the generated contents once it was trained with a more concrete dataset. This will help service providers and marketers adapt or generate their product/service to be relevant, appealing, and customer-tailored.

The findings of this research not only contribute to incorporating linguistic theories more easily in marketing research but also advance the marketing literature in terms of developing more accurate NLP and NLG algorithms for the purpose of persuasive marketing communications, copywriting, and advertising.

BACKGROUND

According to Searle's (1969), speech act theory, written or verbal language not only may explain or describe situations, but also may ask for some behaviors or actions by those who receive and understand the message. For instance, consider someone shares their evaluation of a home-sharing property by words like this:

> "*The host was a very helpful sweet lady, and the king bed was very comfortable but the apartment itself is jam-packed, cluttered falling apart, no shower towels at all, and nothing had been cleaned in a long time… I felt it was more like a college dorm feel, rather than a vacation rental for adults. Honestly, [I] would get a hotel next time.*"

While the reviewer uses nice words about the property's host and the comfort of its bed, she implicitly encourages those who are looking for a vacation rental to book a hotel room instead. In other words, her shared review (also known as *locutionary*) is intended to impede others from booking that house (*illocutionary*) and may result in skipping this property by other consumers in action (*perlocutionary*) (Austin 1975). Drawing on the appraisal theory (Scherer, Schorr, and Johnstone 2001), we argued that someone's estimate of others' language extracts emotions that cause specific reactions. Therefore, in peer-to-peer marketplaces, choosing the words by either side of the market has impacts on the way their ideas are publicly shared. Subsequently, the reactions and behaviors of those who receive the utterances, i.e. peers, competitors, consumers, and policymakers could be changed accordingly (Hovy *et al.* 2021). For example, service providers try to showcase their service in a manner that users get engaged and appreciate. More specifically, consumers engage with several information resources and look into different types of cues, either user-generated content or brand-generated ones, to overcome the inherent intangibility of the service industry, i.e. booking a property in a destination of interest.

To make those contents engaging, each of two aspects of language –semantic and pragmatic– should be considered. The first impression comes out of the explicit word meaning (language semantic). Readers interpret the language (i.e. attitudes, emotions, sentiments, and passions) through its set of semantic features (Johnson-Laird and Oatley 1989). Next, to grasp the communication better, readers explore the interaction of content and context as well as their relationship with the publisher (pragmatic) (Humphreys and Wang 2018). For example, Peng *et al.* (2021) showed that a crowdfunding success or failure relied on the types of words and phrases used for communicating the project. If the language reflects the integrity of the team, certainty, and its relevant social interactions, expecting its success is in line with the language expectancy theory (Burgoon *et al.* 2002).

In this research, we looked at concreteness and sentiment to uncover semantics. For the pragmatic portion of language, we investigated inferences about publishers' social cognition via its two dimensions of competence and warmth. Here, we first discuss the social cognition of the service providers and how their type could affect perceptions towards them. Then, we spell out the role of linguistic concreteness and sentiment of their service descriptions.

Competence and warmth

Stereotyping, although it may not be stable for a long time and may change by social pressures, could be helpful for studying interpersonal and intergroup relationships (Fiske *et al.* 2002). When people interact with each other, they use stereotypes to categorize their counterparts to better make inferences about their affairs. Social cognition is the information that people store, retrieve, and use in social contexts to explain and predict their own and others' behavior (Bulgarelli and Molina 2016). Past research introduced warmth and competence as the universal dimensions of social cognition. Warmth is associated with intention traits like friendliness, helpfulness, sincerity, trustworthiness, and morality. However, the competence dimension captures traits about perceived ability, including intelligence, skill, creativity, and efficacy (Fiske *et al.* 2007). Authors suggested warmth and competence characteristics not only for humans (i.e. salespeople and employees) (Wang *et al.* 2017) but also for non-humans (i.e. for-profits, non-profits, products, brands, robots, and countries) (Aaker *et al.* 2010; Bernritter *et al.* 2016; Halkias and Diamantopoulos 2020; Fournier and Alvarez 2012; Touré-Tillery and McGill 2015; Chen *et al.* 2014).

Sharing economy platforms have two categories. In the first category, the service provider is part of collaborative consumption in the whole process (i.e. ridesharing platforms like Uber). In the second category, service is delivered to the users by themselves, and the service provider does not actively participate in providing the service (i.e. carsharing platforms like Turo). In the present research, we focused on the second category. Although, once they rent a vehicle or book a property, the owner is not anymore part of the collaborative consumption, their role in making the service available for their consumers (a.k.a guests) starts long before the service request through addressing the questions, confirming the booking, providing necessary amenities, and similar tasks. So, for example, when someone wants to rent a car on a sharing platform like Turo or book a service, he may evaluate the responsiveness and trustworthiness of the service provider (provider-related warmth), their hospitality skills (provider-related competence) as well as the service features that make getting a good ride and staying well possible (service-related warmth: i.e. nice view; service-related competence: i.e. neighborhood). In the same way, service providers write descriptions and provide information that would be useful for the purpose of stereotyping.

Service provider type

We expect the type of service provider (amateur vs. professional) makes an impact on the level of social cognition that service providers receive from their consumers. Based on the conceptual model of service quality (Haywood-Farmer 1988), Haywood-Farmer and Stuart (1990) suggested a degree of professionalism, in which knowledge base, job autonomy, societal importance, and own superiority change the level of professionalism. Mobile technologies and Internet infrastructure have made it easier for individuals, with different degrees of professionalism, to share and trade their idle resources (Li *et al.* 2019). More specifically, research shows that professional and nonprofessional service providers (i.e. Airbnb hosts) perform substantially differently operationally and financially. For example, while controlling for property and market characteristics, Li *et al.* (2019) found that properties managed by professional hosts earn 16.9% more in daily revenue, have 15.5% higher occupancy rates, and are 13.6% less likely to exit the market, and thus, they suggested necessary assistance for amateurs in terms of pricing and property management. Gibbs *et al.* (2018) discovered that physical characteristics (i.e. room type, capacity, distance, parking, pool), location (i.e. neighbor and distance), and host characteristics (i.e. professional and super host) significantly impact pricing.

In terms of the service provider impact in the textual field, the authors found that it is common for amateurs to get more favorable reviews and ratings than more professional providers (Pitt *et al.* 2021) (see Appendix C for an example of a professional host on Airbnb). Analyzing the self-presentation of the Airbnb hosts, Tussyadiah and Park (2018) identified two types of hosts and discussed how they had impacts on inducing trust and accommodation booking. The first group included those who were patient and eager to meet new people. They promoted themselves as an empathetic host who understands and is sympathetic to guests by understanding the "ins and outs" of traveling. The second type consisted of individuals with certain professional skills. These hosts were establishing themselves as regular individuals by revealing a greater amount of self-disclosure through projecting themselves based on their profession and personal information. It is also recommended to use other kinds of cues, i.e. verified photos, to decrease the level of uncertainty due to the amateur status of most service providers (Ma *et al.* 2022). This would be helpful in terms of stereotyping of the hosts and how they are related together. For example, previous research suggested that if a surfer does not pay for the co-consumption of the shared property, host and surfer are related more emotionally, i.e. eating food together or having recreational activities, rather than professionally (Geiger *et al.* 2018).

Research showed that individuals responded more positively to a professional service provider when they were oriented towards being sensitive to others' problems. Moreover, a positive review of a warmly positioned service provider increased their willingness to pay. So, we expect that service provider type (professional vs. amateur) changes the social cognition of the service providers as described by their consumers. Guests who book a private entire place are still dependent on the service provider's warmth and competence while staying in the property, and so may relate the comfort of their bed to the friendliness of their host and may appreciate the opportunity of living in a very good location to their competence for listing such a good property. More specifically, we posit being an amateur (vs. professional) would lead to higher (lower) warmth but lower (higher) competence. As such:

H1: *Amateur (vs. professional) service providers receive higher (vs. lower) warmth but lower (vs. higher) competence.*

Linguistic concreteness

Concrete is defined as "*clear and certain, or real and existing in a form that can be seen or felt"* (Cambridge Advanced Learner's Dictionary and Thesaurus, 2022). Research in linguistic concreteness is concerned with the role of lexical concreteness on recipients. Packard and Berger (2021) find that using lexical concreteness (i.e. tangible, specific, or imaginable) signals how cautiously a salesperson is listening to their customers. In the textual context, for example, while being informed that an apartment is "*fully furnished"* provides kind of information, it might not still address a customer's concern that if the unit has *cable TV* and *wireless high-speed internet* or if *linens and towels* are provided. Relational learning theory suggests that consumers form implicit clusters of engineered attributes, leading to easier multidimensional comparisons (Wang *et al.* 2021). Then, these new higher-level attributes (for example *fully equipped with everything you need*) vs. concrete ones (*it has cable TV and wifi*), facilitate the information processing by attenuating the perceived cognitive load. However, this comes with the cost of losing being informative by using less concrete language. Therefore, we expect the more concrete the sales description is, the higher social cognition would result. More formally,

H2: *The concrete (vs. abstract) sales description results in higher (vs. lower) perceived warmth and competence.*

Sentiment analysis

As discussed in the first essay, sentiment is a complex form of textual expression of emotions, feelings, and opinions (Berger *et al.* 2020). Research shows that potential customers appreciate negative reviews conveying service failure information while they admire positive reviews when they find information on the core functionalities, technical aspects, and aesthetics (Ahmad and Laroche 2017). Furthermore, consumers evaluate the same customer experience message differently based on the medium which is used to post the review (Grewal *et al.* 2021). For example, while Twitter is replete with complaints (Shirdastian *et al.* 2019), most peer-to-peer websites hold a very small number of negative sentiments. Therefore, one could interpret trivial objections on those platforms as a red flag. We expect the impact of service-provider type on their perceived social cognition transmits via the sales description sentiment. More formally,

H3*: The impact of concreteness on social cognition is mediated by the level of sale description sentiment. Moreover, being a professional host (vs. an amateur) increases the positive effect of concreteness on description sentiment.*

RESEARCH DESIGN

According to Statista (2021), roommates Brian Chesky and Joe Gebbia launched AirBed and Breakfast on the living room floor of their San Francisco apartment in 2007. Nathan Blecharczyk joined AirBed and Breakfast in 2008, and Airbedandbreakfast.com was launched (shortly after, it was shortened to Airbnb.com). Basically, Airbnb is an online accommodation rental marketplace where hosts can list their properties and rooms for rent. Rather than owning the properties, it offers for rent on its website, Airbnb generates revenue through the service fees it charges hosts and guests. Airbnb was valued at 113 billion U.S. dollars in 2021, a significant increase from 75 billion the previous year. According to Airbnb, the number of nights and experiences booked worldwide increased year over year between 2017 and 2019 but decreased in 2020 due to travel restrictions (Statista 2021).

Study 1

The research dataset contained 398,926 reviews (from August 2010 to November 2021) for 15,156 properties, located in Toronto, ON (scrapped in November 2021 and got via Inside Airbnb). We removed 3478 listings due to not having reviews, leading to 11,678.

Field experiment design

The design was a 2 (listing concreteness: low vs. high) × 2 (host type: amateur vs. professional) Here, we explain the design in more detail.

Linguistic concreteness: low vs. high

We used the dictionary of concreteness ratings for 39,954 generally known English word lemmas ($M_{concreteness\ rating} = 3.04$, SD = 1.04) prepared by Brysbaert *et al.* (2014) to measure concreteness score (CS). Using the Natural language processing (NLP) method of lemmas tokenization (via NLTK library in python), the number of tokens in common between our dataset and the concreteness dictionary was counted, and its sum of ratings was calculated ($M_{listings'\ in\text{-}common\ tokens} = 137.41$, SD = 61.62 and $M_{listings'\ CS} = 7.77$, SD = 24.22). Then, we normalized the concreteness score by dividing the number of in-common tokens for each listing description and grouped the listings to low (vs. high) concreteness based on its average ($M_{normalized\ listings'\ CS} = .403$, SD = .07).

Service provider type: amateur vs. professional

From the 11,678 properties, we assigned properties that are managed by hosts who have four or more than four unique listings (vs. less than four listings) as being hosted by professionals (vs. amateurs). In previous studies, the number of listings managed by professionals was arbitrarily set at two or three (Li *et al.* 2016). However, even a two-bedroom unit could make three listings (two bedrooms separately as well as the entire unit) and so the number of listings itself is not a reliable tool. As professional hosts are more likely renting out a property that they are not primarily living in (i.e. guest suites and second properties), they would be open to book in advance and make the property available for more days per year.

Therefore, we considered the number of days that the listing is available over the next 365 days as well: if the listing was available to book for more (less) than 90 days, we coded the unit as highly (slightly) available. We found a significant difference between the number of listings based on their availability ($M_{slightly\ available} = 4.26$, SD = 1.66 vs. $M_{highly\ available} = 7.77$ (SD = 24.22); $F(1,11676) = 93.73$; $p < .001$). So we chose the four listings based on the less available units as the threshold for the purpose of host type grouping ($M_{slightly\ available} = 4.26$, CI 95% [3.91, 4.62]) rather than the simple average of the units listed by a host (M = 5.89, SD =19.62) or an arbitrary one.

Perceived warmth and competence

We measured perceived social cognition in the reviews via its two dimensions of warmth (friendly, kind, likable, and nice) and competence (capable, competent, efficient, and skillful) (Halkias and Diamantopoulos 2020). Two individual coders classified 1343 distinct review tokens based on the proposed items for the two dimensions of social cognition (initial intercoder agreements = 78%, conflicts resolved by mutual consent). Then Valence Aware Dictionary and Sentiment Reasoner (VADER), which is a lexicon and rule-based sentiment analysis tool geared toward identifying and analyzing sentiment expressed in social media (Hutto and Gilbert 2014), was applied to measure the polarity of the keywords for each dimension to make the calculation of the respective social cognition possible.

Previously, we expected that professionals are less likely to get high social cognition, specifically for the warmth dimension (H1). Moreover, in H2, we posited that if the listing's textual is more concrete, guests' perception of social cognition increases. We examined the perceived social cognition for the two types of the hosts amateur (less than four listings) and professional (four and more listings) conditions. Running an ANOVA with host type (amateur vs. professional) and listing concreteness (low vs high) as independent variables, and warmth, competence, and the social cognition as the dependent variables, showed a significant negative main effect of host type on perceived warmth ($M_{amateur}$ = 2.27, SD = 1.14 vs. $M_{professional}$ = 2.04, SD = 1.12; $F(1,11674) = 95.30$; $p < .001$; $\eta_p^2 = .008$), supporting H1. Furthermore, results revealed a significant positive main effect of listing concreteness (M_{low} = 2.16, SD = 1.16 vs. M_{high} = 2.24, SD = 1.13; $F(1,11674) = 9.82$; $p = .002$; $\eta_p^2 = .001$) on perceived warmth as well. We found the same pattern for the competence dimension, surprisingly a significant negative main effect of host type ($M_{amateur}$ = .52, SD = .51 vs. $M_{professional}$ = .45, SD = .51; $F(1,11674) = 40.03$; $p < .001$; $\eta_p^2 = .003$) and as expected, a positive main effect of listing concreteness (M_{low} = .48, SD = .51 vs. M_{high} = .51, SD = .52; $F(1,11674) = 6.82$; $p = .009$; $\eta_p^2 = .001$) on perceived competence. We obtained the same significance directions and levels for social cognition (see Table 3.1). The interaction effects between host type and listing concreteness were not significant in any condition ($p > .52$).

Listing's compound sentiment

We also checked the effects of host type and listing concreteness on the listing's compound sentiment. We previously predicted that amateur hosts would build a property description with more positive sentiment. Also, using more concrete language would increase positive sentiment of the listing. An ANOVA with host type (amateur vs. professional) and listing concreteness (low vs. high) as independent variables, and listing's compound sentiment as the dependent variable revealed a negative significant main effect of host type ($M_{amateur}$ = .83, SD = .28 vs. $M_{professional}$ = .81, SD = .30; $F(1,11674) = 9.49$; $p = .002$; $\eta_p^2 = .001$). We also found a significant positive main effect of listing concreteness (M_{low} = .75, SD = .34 vs. M_{high} = .86, SD = .24) on listing sentiment ($F(1,11674) = 354.74$; $p = .001$; $\eta_p^2 = .029$). The negative main effect of host type and positive effect of listing concreteness were qualified by a significant interaction effect ($F(1,11674) = 8.89$; $p = .003$; $\eta_p^2 = .001$). In low concrete condition, hosts who manage less than four properties used more positive textual description than professionals ($M_{amateur}$ = .77, SD = .38 vs. $M_{professional}$ = .73, SD = .36; $F(1,4313) = 9.98$; $p = .002$; $\eta_p^2 = .002$. However, this was not significant for listings with high concreteness (M_{low} = .86, SD = .23 vs. M_{high} = .86, SD = .25; $F(1,7361) < 1$; $p > .9$).

More importantly, we posited that the impact of concreteness on social cognition is mediated by listing's polarity while that level of sentiment is moderated by the host type. A moderation-meditation analysis (PROCESS model 7, 5000 bootstrap samples; Hayes, 2013) by including listing concreteness as the independent variable, host type as the moderator, listing's compound sentiment as the mediator, and social cognition as the dependent variable supported the predicted effects. More specifically, we found that being a professional host (vs. an amateur) heightened the positive effect of concreteness on properties' sentiment ($\beta_{amateur}$ = .095, SE = .006, $p < .001$ vs. $\beta_{professional}$ = .131, SE = .010, $p < .001$). Furthermore, not only higher positive listing's sentiment led to more positive perceived social cognition (β = .277, SE = .023, $p < .001$) but also it fully meditated the effects of listing concreteness on perceived social cognition (β_{direct} = .021, SE = .014, $p = .120$). As captured by the index of moderation-mediation (index = .010, SE = .004, 95% CI[.003, .018]), the indirect effects were significantly different between amateur hosts and professionals ($\beta_{amateur}$ = .026, SE = .003, 95% CI[.021, .032] vs. $\beta_{professional}$ = .036, SE = .005, 95% CI[.028, .046]).

Table 3.1: The main effect and interaction effect of host type and concreteness on social cognition and listing sentiment.

		Social cognition	Listing's compound sentiment	n
Main effects				
Host type	Amateur	1.40 *(.70)*	.82 *(.28)*	**8509**
	Professional	1.24 *(.69)*	.81 *(.30)*	**3169**
Listing concreteness	Low	1.32 *(.70)*	.75 *(.34)*	**4315**
	High	1.37 *(.70)*	.86 *(.24)*	**7363**

Interactions	**Low listing concreteness**		**High listing concreteness**		**Low listing concreteness**		**High listing concreteness**	
Host type	M *(SD)*	*n*	M *(SD)*	*n*	M *(SD)*	*n*	M *(SD)*	***n***
Amateur	1.37 *(.70)*	3129	1.41 *(.70)*	5380	.76 *(.33)*	3129	.86 *(.23)*	**5380**
Professional	**1.21 *(.66)***	**1186**	**1.27 *(.70)***	**1983**	**.73 *(.88)***	**1186**	**.86 *(.25)***	**1983**

Discussion

Past research suggested that due to having listed services by the amateurs on the peer-to-peer platforms, loss of service quality prevents potential users to get a service from these platforms, and so Del Chiappa *et al.* (2021) suggested promoting the listing via textual information. In Study 1, we illustrated the impact of the host types and their listing textual concreteness on guests' perception of the host's warmth and competence. The results supported that guests shared different views towards amateur and professional hosts and the level of concreteness of the descriptions changed the social cognition of the hosts. It also depicted that, while this image in guests' minds was not directly the result of their language concreteness, it changed their perceptions via enhanced listing sentiment.

Study 2

Study 1 showed the role of linguistic concreteness and service provider type via sale description on the perceived level of warmth and competence of the service providers. We also observed the way amateur (vs. professional) service providers are communicating their listings and receiving recognitions from their users. While we expected to find that being professional (vs. amateur) is more associated with competency (vs. warmth), the findings supported no significant difference. This encouraged us to investigate in more detail the four kinds of properties (amateur and abstract, professional and abstract, amateur and concrete, and professional and concrete) and how their service provider was promoting its listing. This could also be helpful to understand better different cues guests could use to get an impression of warmth and competence about the unit. For example, those who book an entire unit and not a private or shared room are not dependent on the service provider for living in the property. However, they still could feel and evaluate the comfort of their bed, the accuracy of the descriptions, the location, and other characteristics, and would associate it to either warmth or competence of the host or the property itself. Here, we start by explaining two textual analysis methods that we used to get a better sense of the four property types, and then we provide more detail on host-related vs. property-related social cognition.

Building on the results of Study 1, in Study 2, first we investigated the top one hundred tokens in each condition. To do so, we used *Ngram* (i.e. fully, fully_equipped, and fully equipped kitchen) and *phrases* libraries of *the Gensim* package to both provide a more broad spectrum of tokens but to drop the Ngrams that are not real phrases in English.

Among the top one hundred tokens in each condition, Table 3.2 shows those tokens that are not in common for all the four conditions (yellow: unique for one condition, blue: observed in two out of four conditions, and green: observed in three conditions). While the results highlighted the close usage of keywords among most of the properties (out of four hundred top tokens, one-hundred thirty-six distinct tokens, and only thirty-two tokens not seen in all four conditions - yellow ones in Table 3.2), the rank in each condition provided further support for the classification of host type and level of concreteness. For example, while in condition 1 (amateur and abstract) we observed general terms like *includes* and *including*, in conditions 3 and 4, the hosts used more concrete words like *full kitchen*, *fully equipped kitchen,* and *gym.* Moreover, while a group of professionals abstractly described their listing as *nice*, the other group solidly mentioned the unit feature or amenity like having a *closet*. As another example, the results showed that both amateurs and professionals considered the importance of mentioning their propriety's *location*; however, their respective ranking was considerably different (55 out of 100 for professionals vs. 78 and 90 out of 100 for amateurs).

Next, we looked more closely at each of the four conditions to find appropriate and relevant names for them. First, those amateur hosts who have not used concrete descriptions tried in several ways to highlight that they were mostly renting their spare property (i.e. *basement apartment*) with *separate/private entrance*. As the *furnished* token was not raised among the top 100 tokens, we interpret that the units were either not *furnished*, or the hosts were not able to effectively communicate their amenities. However, this was not the case with concrete amateurs. They provided more information about their unit amenities and characteristics (i.e. *gym*, *fully equipped kitchen*, *queen size bed*, and *luxury*). This showed that their approach in promoting their few listings (less than four properties) was more similar to concrete professionals than their other fellow amateurs. Finally, the abstract professionals utilized common positive tokens like *big*, *nice*, and *smart* to describe different attributes of their place, but they missed addressing in more detail what they really meant by those modifiers (i.e. *furnished* vs. *fully furnished* or *fully equipped kitchen*). Therefore, we named the first group of amateurs as *novices*. Those amateurs who used concrete words are recognized as *to-be-professionals*. But as the professionals who have not used concrete words were more similar to amateurs', they were still in the *junior* stage of their career, and the real professionals with high potential to effectively communicate indeed are *real estate agents*.

Table 3.2: The distinct tokens in each of the four property groups.

	Amateur and Abstract: ***Novice***	**Rank**	**Professional and Abstract:** ***Junior***	**Rank**	**Amateur and Concrete** ***To-Be-Professional***	**Rank**	**Professional and Concrete:** ***Real Estate Agent***	**Rank**
Observed in one condition	separate entrance	61	big	66	full kitchen	79	fully equipped	78
	basement apartment	62	away	68	gym	91	closet	80
	separate	70	time	75	everything	92	Netflix	82
	private entrance	76	nice	78	loft	97	studio	90
	beach	77	door	79	one bedroom	65	full bathroom	93
	many	83	entertainment district	81	fully furnished	69	den	94
	shopping	95	guest access	85	family	73	centre	96
	transit	98	smart	87	fully equipped kitchen	76	living area	97
Observed in two conditions	one bedroom	71	double bed	92	lot	84	sleep	98
	station	73	free parking	93	luxury	89	window	99
	lot	78	subway station	51	queen size bed	93	fully furnished	40
	grocery store	84	station	67	open concept	95	luxury	44
	subway station	85	free	73	included	96	fully equipped kitchen	45
	main floor	86	near	76	full	41	furnished	52
	near	88	tower	83	includes	56	queen size bed	75
	free	97	furnished	86	living	63	included	86

Observed in three conditions	step away	67	grocery store	89	view	68	tower	87
	includes	79	main floor	90	including	71	open concept	88
	min walk	80	family	99	small	74	full	24
	full	87	queen bed	53	location	78	view	46
	come	89	step away	70	Wi-Fi	80	living	53
	location	90	min walk	71	queen bed	82	location	55
	Wi-Fi	91	small	84	come	83	queen bed	62
	brand-new	93	come	91	brand-new	86	includes	71
	washer dryer	94	brand-new	95	min walk	88	washer dryer	84
	including	96	living	96	step away	90	Wi-Fi	85
	small	99	view	97	washer dryer	98	including	92

Warmth and competence: host-related vs. property-related

Next, we looked into the emerged tokens considering warmth and competence in listings' descriptions. We expected the number of tokens related to property-associated tokens of social cognition dimensions to be different than the host-associated ones based on host type and linguistic concreteness.

We used modified LDA, a topic modeling approach, to extract common tokens for each of the four groups. The logic of topic modeling is that each document has a set of topics associated with a collection of a fixed vocabulary of terms, and each document has these topics with different proportions (Blei and Lafferty, 2009). So, if an algorithm can find these latent associations between the words in a text, it can find the underlined topics of a document. There are different methods for topic modeling. Latent Dirichlet allocation (LDA) is a powerful method in extracting topics from a document. A three-level hierarchical Bayesian model is used to model collections of discrete data, i.e. text corpora, using a generative probabilistic model in which each item is modeled by a finite mixture over an underlying set of topics. As a result, each topic is modeled as a set of infinite combinations of associated topic probabilities (Blei *et al.* 2003). Calculating the probability, rather than a simple emerged topic, provides a useful tool for the purpose of prioritizing (Blei and Lafferty, 2007).

As we had two topics of warmth and competence in mind, by inputting warm, kind, and generous for the first topic (warmth) as well as competent, efficient, and effective for the second one (competence), LDA via *tomotopy* Python package helped in categorizing common words that were similar to the inputs for each of the two topics. To get more specific tokens and avoid the generic ones (i.e. *place* and *bedroom*), we removed the top one hundred words. Then, two coders independently coded the emerged tokens as either host-related or property-related (initial intercoder agreements = 81%, conflicts resolved by mutual consent). Those tokens that made them in place were directly the result of planning, behaving, or acting of the host, i.e. *fully_equipped_kitchen*, *comfortable_suite_everything_need*, and *private_bedroom_lock*, were coded as host-related; however, the tokens that had them available were directly the result of the property itself, i.e. being *minute_away*, *quiet_street*, and *floor_ceiling_window* were coded as property-related. Next, we counted the number of tokens in each condition as the dependent variable in a one-way ANOVA. Considering type of the hosts (amateur vs. professional) as the independent variable, we found more host-related tokens for professionals (juniors and real-estate agents) than amateurs (novices and to-be-professionals) in the competence dimension ($M_{amateur} = 24$, SD = 8.48 vs. $M_{professional} = 63.5$, SD = .71; $F(1,2) = 43.04$; $p = .022$). However, for the property-related tokens, we found higher tokens for amateurs than professionals ($M_{amateur} = 76$, SD = 8.48 vs. $M_{professional} = 36.5$, SD = .71; $F(1,2) = 43.04$; $p = .022$). Surprisingly, while the results showed not the same pattern for the warmth dimension, we found the differences were insignificant ($M_{amateur} = 48.5$, SD = 4.95 vs. $M_{professional} = 43.5$, SD = .71; $F(1,2) = 43.04$; $p = .022$ as well as $M_{amateur} = 51.5$, SD = 4.95 vs. $M_{professional} = 56.5.5$, SD = .71; $F(1,2) = 43.04$; $p = .022$). Table 3.3 shows the number of tokens in each condition.

Table 3.3: Number of tokens for warmth and competence for each of the four types of properties based on host-related and property-related tokens.

	Novices		Juniors		To-be-professionals		Real-estate agents	
	Competence	**Warmth**	**Competence**	**Warmth**	**Competence**	**Warmth**	**Competence**	**Warmth**
Host-related tokens	18	52	63	43	30	45	64	44
Property-related tokens	82	48	37	57	70	55	36	56

Discussion

Study 2 provided further support for the impact of service provider type as well as the concreteness of the textual content they used to describe and promote their service. We used two text mining methods (frequency analysis and topic modeling) to better understand the differences among the four kinds of property descriptions that we observed on Airbnb. To this end, we excluded the common tokens among these four categories. We discovered that amateur hosts who used abstract description (*novices*) were mainly renting out their own spare spaces (i.e. basements). However, they either missed to showcase how competent they were in hospitality, or they were actually not well-trained in this regard. We also recognized *to-be-professionals* as a group of amateurs who understood the importance of using concrete language (i.e. *fully equipped kitchen*).

Among those who listed several units, we detected a totally different approach between *juniors* and *real-estate agents*. While juniors acted more similarly to novices in terms of communicating abstractly, i.e. using words like *big* and *nice*, real-estate agents addressed more information about the units by utilizing words like *fully equipped* and *living area.* The difference was also in place when we grouped host-related tokens and property-related tokens for each of the warmth and competence dimensions among the four property management kinds. Amateur hosts were more inclined to describe their competence via the features of their property (i.e. *within_walking_distance*, *entrance*, *amazing_view*); however, professionals highly emphasized their own hospitality capabilities (i.e. *brand_new_renovated*, *pillow_towel_complimentary*, *perfect_savvy_budget_traveler*). Unexpectedly, we did not find a statistically significant difference in terms of warmth dimension. A possible explanation could be the way service providers see sharing economy platforms and their endeavor to depict a welcoming and caring image in their users' minds (Ranchordás 2015).

Study 2 shed light on differences between the services based on their provider and the way they communicate both their own and that of their service's warmth and competence. In Study 3, we employ the findings of Study 1 and Study 2 to better train the Natural language generation (NLG) algorithm. Below, we introduce NLG and how using Long-Short-Term Memory (LSTM) as well as cleaned dataset could help service providers generate more engaging content.

Study 3

Natural language generation (NLG)

Natural language generation (NLG) is the process in which a machine generates text that is accurate, useful, and easy to comprehend (Reiter, 2019). According to Reiter and Dale (1997), NLG systems have three interrelated tasks: text planning, sentence aggregation, and linguistic realization. For the first task, the content determination and discourse planning subtasks are combined because it could be difficult to separate these two activities in real-life situations. The second task combines aggregation, lexicalization, and referring expressions (i.e. pronominalization), while in the last task a realizer controls whether the output is consistent with grammatical rules in syntactic (i.e. discourse markers), morphological (i.e. singular vs. plural), and orthographic (i.e. spelling) (Reiter and Dale 1997). Table 3.4 provides more explanations about each of the three tasks and their respective subtasks. While NLG shows an acceptable outcome level in the relatively short texts, i.e. captions, poems, lyrics, the challenge is how to increase the perceived quality in long formats. Research suggests Long-Short-Term Memory (LSTM) as an appropriate method for keeping long-term dependencies.

Table 3.4: NLG tasks, subtasks, and definitions.

NLG tasks	Subtasks	Definition
Text planning	Content determination	Select which information to convey in the text.
	Discourse planning	Creating structure and order over a collection of messages.
Sentence aggregation	Aggregation	Putting together the messages to form sentences.
	Lexicalization	Make specific word and phrase choices to convey the domain concepts and relations that are presented in the messages.
	Referring expressions	Identifying domain entities using words or phrases.
Linguistic realization	Syntactic	For example, adding function words like *from* and *to*
	Morphological	For example, changing a singular word to plural
	Orthographic	For example, capitalizing the first word and adding a full stop as the end

Bidirectional Long-Short-Term Memory (BLSTM)

Long-Short-Term Memory (LSTM) keeps some useful information in mind as long as it is necessary, whether the period is short or long (Hochreiter and Schmidhuber 1997). The method is capable of different kinds of equational data, i.e. text time series, music rhythms, and so on. Since it is a kind of recurrent neural network (RNNs), in which layers are connected to each other in a way that each outcome gets affected by layers before, the final outcome (i.e. next several words prediction in the context of NLG) represents the accumulated knowledge. Olah (2015) explained the LSTM algorithm by focusing on its cell state (the horizontal line from *t-1* to *t+1* in Figure 3.1) which carries the useful information all through the network. LSTM begins by deciding what information should be discarded from the cell state using a sigmoid function at the *forget gate layer*. Comparing h_{t-1} and x_t, the first sigmoid function gives a number in the range of 0 (totally forget) to 1 (completely keep), which will be later on multiplied by C_t. Next, it must decide what new information will be stored in the cell state. In the first step, it updates values with a sigmoid called the *input gate layer*. After that, the tanh layer (from -1 to 1) creates a vector of new candidate values that can be added to the cell state. In fact, the new cell state is composed of the forget value (f_t, which is between 0 and 1) multiplied by C_{t-1} (old state value) and the new candidate values multiplied by i_t (updating scaler). The last step is to calculate the output (h_t), which is an updated C_{t-1} based on the new input (x_t). Again, a sigmoid gate is used to come up with the level of discarding/keeping of the combination of h_{t-1} and x_t. Then another tanh is used to make the cell state ready (values between -1 and 1) for being multiplied by the output of the sigmoid function to result in h_t. Figure 3.1 illustrates the algorithm.

If the same calculation happens from *t+1* to *t-1* direction as well, that is called Bidirectional LSTM (BLSTM). It is useful for inferences about elements that part of the information is provided before and the rest comes after. For example, consider the following property description: "*This is a comfortable and fully equipped condo in a trendy neighborhood. I am available to help out or can leave you alone to enjoy your stay. Great stainless-steel appliances, Washer and Dryer in the suite for your convenience. Nicely furnished with a leather couch, dining table, and queen bed. Cable included. All linens, towels, dishes, and necessities are available. Transit outside your door"*. Based on this paragraph, if the machine wants to predict the next word in "*This is a comfortable and fully equipped condo …*" there is no useful information before the unknown word; however, if it uses explanations that come afterward as well such as "*great stainless-steel appliances, washer and dryer, leather couch, dining table, queen bed, Cable, linens, towels, dishes, and necessities*", it is safe to predict that *fully equipped condo* should be followed by such amenities.

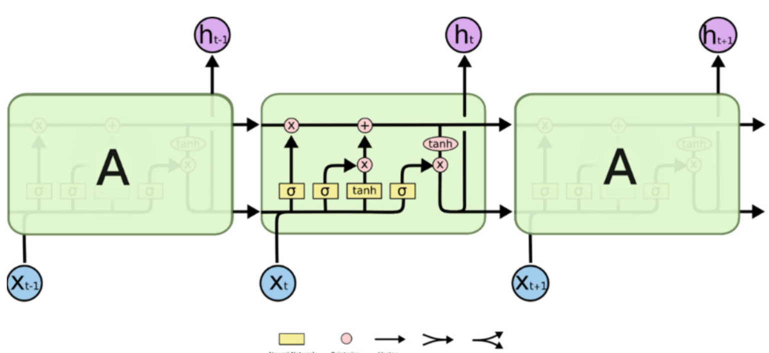


Figure 3.1: LSTM networks (Olah, 2015)

Research methodology

We expected that training the NLG algorithm over the descriptions posted by real-estate agents (professionals with high linguistic concreteness) would lead to higher readability and naturalness of the generated description. To test it, we used the Keras Sequential model, which is defined as "*a plain stack of layers where each layer has exactly one input tensor and one output tensor*" (Keras documentation, 2022). Then, we added *the Embedding* layer, in which categorical features are converted into fixed-length dense numerical vectors of the features (i.e. words) (Saxena, 2020). The input size of the layer is equal to vocabulary size while for the output dimension, it is suggested to increase the size for a better representation of the input if the vocabulary is not composed of sparse documents (Shusterman, 2020). We started with vocab size as the output but we did not get meaningful results and high accuracy, and so decreased it to fifty. The next step is to add the *Bidirectional LSTM* (BLSTM). As discussed before, it carries over the useful information in the neural network, consisting of the chain of repeating modules, via its cell state. The last layer that we added was the *Dense* layer which each of its neurons receives input from all of the neurons in the previous layer. Matrix-vector multiplication occurs in the dense layer. Among different activation functions for the Dense layer, we chose *Softmax* which converts a vector of values to a probability distribution in the range of (0, 1) and sum to 1. As the optimizer, we used adaptive moment estimation (Adam) which is a gradient-based approach to stochastic objective functions (Kingma and Ba 2014).

Results and discussion

To train the model, we first used the whole dataset, without considering the four types that we found in Study 1 and Study 2. The *Google Colab Pro+* (RAM: 51.01 GB, and GPU Tesla V100-SXM2-16GB) was used for coding and running the algorithm. By inputting a keyword, the neural network predicted the next probable keywords. For example, by requesting to generate content for keywords "*come*" and "*fully*", we got the following meaningless next forty keywords for "*come:*

> *"Come and cook we'll help you started is a quiet and comfortable and comfortable place for a memorable*
>
> *the space*
>
> *this is a quiet residential street in the heart of toronto in the heart of toronto in the heart of toronto"*

and here for "fully":

> *"fully equipped kitchen and a gourmet kitchen and a gourmet kitchen and a gourmet kitchen and a gourmet kitchen and a gourmet kitchen and a gourmet kitchen and a gourmet kitchen and a gourmet kitchen and a gourmet kitchen and a"*

We found the same output problem with other keywords like "*basement*" and "*bright"*. In the next step, we checked whether the results get improved by training the model on any of the four groups that we identified in Study 1 and Study 2. The results supported this notion, and the output significantly refined once trained on descriptions written by group 4, the real estate agents. As the sample size decreased, the machine was also able to afford to be trained on uppercase characters as well as periods and semicolons. We started with total trainable parameters: 240,475,592 once the output size was equal to vocabulary size. As we did not get a meaningful outcome, and as was suggested in the literature for sparse documents, we decreased it to fifty, and so the total parameters decreased to 5,859,006. Table 3.5 shows the model summary. For example, with the keyword *"equipped'*, we got the next forty words as:

> *"Equipped with pots, pans, etc.*
>
> *The space*
>
> *This is a private room with a queen size bed and a small closet and a comfortable mattress and pillows.*
>
> *Guest access*
>
> *Guests have access to the backyard and a small balcony and the"*

And for two-hundred words after "*bright*":

> *"Bright and spacious furnished room in a quiet area situated in the heart of Toronto's downtown core. It is the top tier of the downtown core, and Little Italy, Kensignton, Art Gallery of Ontario and Royal Ontario Museum for an immersive historical experience!*
>
> *The space*
>
> *The private room consists of a full bed, workspace, and a closet. The room is furnished with hardwood floors and a high ceiling, located near major Toronto attractions that are walking distance!*
>
> *Guest access*
>
> *There will be a shared kitchen, washroom, and laundry room Samsung washer and dryer available to use. bedroom is located in the main floor of the two story in the heart of downtown located in the city across from the CN Tower and the University of Toronto and the spanning lakeshore. and other excellent you can be easy access to all the amenities in the home in the city. The apartment is a well-maintained apartment with a large updated bathroom in the living room and a large deck with a large window. with a view to ceiling windows in the living room with a large sofa bed and a big bathroom that home to be made into a secure or a vocalist."*

Table 3.5: Neural Network model specifications.

Layer type	Output Shape	Number of parameters
Embedding	(None, 182, 50)	385300
Bidirectional	(None, 600)	842400
Dense	(None, 7706)	4631306
Total parameters	5,859,006	
Trainable parameters	5,859,006	
Non-trainable parameters	0	

While the content's readability and naturalness were increased, there were still a few flaws in the text which required a human evaluator to modify the text. This was in line with previous research findings. More importantly, although after 20 epochs, the training accuracy reached 85%, testing the validity over 20% of data was 39%. This suggested overfitting; however, the results would be still useful for new hosts or existing hosts if they want to copycat their fellow hosts. We tried several ways to overcome this overfitting, which were incrementally helpful. We have some suggestions accordingly that we discuss in the future research section.

GENERAL DISCUSSION

Building on the speech act theory, we examined how sale descriptions have an impact on service providers' social cognitions in the eyes of consumers, and how they could be used to generate new content. Previous studies only focused on general linguistic characteristics. However, in this research, we added to the social cognition theory by examining the combination of warmth and competence in an in-field observation of user-generated content in the sharing economy context. We used Airbnb, which is an online peer-to-peer marketplace and hospitality service dataset. Through this platform, hosts and guests communicate publicly with each other through the property's description by the host and reviews from the guests.

In Study 1, we measured how different types of service providers (amateurs vs. professionals) and certain concreteness levels (abstract vs. concrete) change the perceived warmth and competence of the service providers. Applying Natural Language Processing methods, we analyzed the guest reviews and the property descriptions and found lower perceived warmth and competence for amateur hosts as well as higher social cognition from more concrete descriptions. Moreover, we discovered that in abstract descriptions, amateurs used a more positive sentiment description than professionals. Not only did the results suggest that higher positive listing sentiment led to more positive social cognition but it also fully meditated the effects of listing concreteness on social cognition. The findings showed the role of linguistic concreteness and sentiment analysis on both social cognition dimensions for service providers. In summary, Study 1 showed the role of linguistic concreteness, service provider type, and sentiment analysis on the perceived level of warmth and competence of the service providers.

In Study 2, we used two text mining methods (frequency analysis and modified LDA) to better understand the differences among the four kinds of property descriptions and their hosts. We found that, for example, abstract amateurs (*novices*) were mainly renting out their own spare spaces (i.e. basements), *to-be-professionals* understood the importance of using concrete language (i.e. fully equipped kitchen), *juniors* acted more like novices in terms of abstractly communicating, while *real-estate agents* addressed more information about the units by utilizing words like fully equipped and living area. Moreover, the modified LDA extracted the relevant tokens to social cognition dimensions from property descriptions. Results showed that amateur hosts were more inclined to describe their competence via the features of their property, while professionals highly emphasized their own hospitality capabilities. Surprisingly, we did not find a statistically significant difference in terms of property-related (vs. host-related tokens) for the warmth dimension.

In Study 3, we employed the findings of Study 1 and Study 2 to better train the Natural language generation (NLG) algorithm and showed how Long-Short-Term Memory (LSTM), as well as a cleaned input, could help service providers generate more engaging content. The results showed following professionals and using more concrete language led to getting more persuasive sales descriptions. Table 3.6 summarizes the finding in the three studies.

Table 3.6: Summary of the studies

Study	Purpose	Key findings
Study 1	o Professionals are less likely to get high social cognition, specifically for the warmth dimension o If the listing textual is more concrete, guests' perception of social cognition increases. o The effects of host type and listing concreteness on the listing's compound sentiment o The impact of concreteness on social cognition is mediated by listing's polarity while that level of sentiment is moderated by the host type	✓ A significant negative main effect of host type on perceived warmth ✓ A significant positive main effect of listing concreteness on warmth ✓ A significant negative main effect of host type on competence ✓ The positive main effect of listing concreteness on competence ✓ The negative significant main effect of host type on listing sentiment ✓ The positive main effect of listing concreteness on sentiment ✓ In low concrete condition, hosts who manage less than four properties used more positive textual descriptions than professionals ✓ Being a professional host (vs. an amateur) heightened the positive effect of concreteness on social cognition ✓ Not only did higher positive listing's sentiment lead to more positive perceived social cognition but also it fully meditated the effects of listing concreteness on perceived social cognition
Study 2	o Providing further support for the impact of service provider type as well as the concreteness of the textual content they used to describe and promote their service o Using two text mining methods (frequency analysis and topic modeling) to understand better the differences between the four kinds of property descriptions	✓ Amateur hosts who used abstract description (*novices*) were mainly renting out their own spare spaces (i.e. basements). We also recognized *to-be-professionals* as a group of amateurs who understood the importance of using concrete language (i.e. *fully equipped kitchen*). ✓ While juniors acted more similarly to novices in terms of communicating abstractly, real-estate agents addressed more information about the units by utilizing words like *fully equipped* and *living area.* ✓ Amateur hosts were more inclined to describe their competence via the features of their property ✓ However, professionals highly emphasized their own hospitality capabilities. ✓ Unexpectedly, we did not find a statistically significant difference in terms of warmth dimension.
Study 3	o Employing the findings of Study 1 and Study 2 to better train the Natural language generation (NLG) algorithm.	✓ Showing how Long-Short-Term Memory (LSTM), as well as cleaned input, could help service providers generate more engaging content. ✓ The results showed following professionals and using more concrete language led to getting more persuasive sales descriptions.

Theoretical contributions

We added to the textual analysis literature by linking sales descriptions to consumers' perceptions which are shared via their comments and reviews. This also contributed to previous studies that looked into the persuasiveness of marketing messages. Drawing from linguistic theories and appraisal theory, we proposed that the way service providers use some words (concrete vs. abstract) has substantial and significant power in changing their social cognition. The findings also supported results from past research that in sharing economy platforms, there could be two types of motivations; namely “to earn cash” and “to share beauty” and “to meet people” (Chung *et al.* 2021) based on the service provider type (amateur vs. professional).

Methodologically, the proposed way of content generation and the research findings on opinion marketing could address the existing gap in the literature of topic modeling and NLG. Recent research suggests combining AI and Human capabilities to both get benefits of AI’s hard data skills and human’s personal and interpersonal skills (Luo *et al.* 2021). This was among the very first studies to use modified LDA as a strategic tool to identify the common words related to a specific topic (Tirunillai and Tellis 2014). More importantly, we trained and tested an NLG neural network that generated meaningful sales descriptions.

Practical implications

Traditional techniques are being rendered obsolete by the sheer size and diversity of big data (text, numbers, emoji, or video) (Sheth 2021). Data quality is a fundamental concern when it comes to applying machine learning and artificial intelligence as their applications are only as good as the training data they are based on (Hair and Sarstedt 2021). Practically, the findings help evaluate customers' perceptions about product/service, measure the match between product/service with customer expectations and needs, and generate target-specific messages, offers, and recommendations.

We observed four groups among the service providers; namely: novice, junior, to-be-professional, and real estate agent, and so we have specific recommendations for each of them. First, the novice service providers should extensively work on their offerings (i.e. welcome packages and equipped kitchen) and if they already provide them, they need to clearly mention and promote them. For juniors, while they are listing several units, they are not enjoying the vast opportunities due to their low-level descriptions or their service. For the last two groups, we believe that to-be professionals are on the right track to get the benefits of sharing economy; but similar to real estate agents, they need to not lose their focus on the warmth (either service provider-related or service-related).

Limitations and future research

This research has some limitations that offer opportunities for future research. First, while message persuasion is the main goal of marketing and communications, marketers need to be watchful of preternaturally persuasive messages (Donath 2021). We suggest future research to investigate the description accuracy and honesty of the service provider. This could be done by incorporating image processing techniques (i.e. dense captioning and visual question answering) to verify textual description and image illustrations. Second, we introduced a tool that service providers could use to generate more concrete descriptions. However, we did not study if, for any reason, they would or would not be interested in using it. For example, past research suggested that consumer reactions to new technologies might be influenced by fear, technophobia, and speciesism (Schmitt 2019). It would be worthwhile to research how service providers use such a tool in a lab or in an in-field study. Third, although we used a big dataset for this research, it contained just one North American city. A cross-cultural comparison could be helpful for the purpose of testing the generalizability of the research findings or finding boundary conditions. Fourth, as it was an in-field study without prompting reviewers, many attributes could be missed (Chakraborty *et al.* 2021). Fifth, future research could chunk the data based on neighborhood, property type (room, entire place, etc.) to increase accuracy for the testing dataset. Therefore, future research could verify the findings in more controlled research settings by manipulating the respondents in terms of social cognition. Finally, we did not have access to property-level data on host income. A dataset from *AirDNA* could be helpful not only to measure property equity using text analysis but also to estimate the impact of perceived warmth and competence on novices, juniors, to-be-professionals, and real estate agents.

CONCLUDING REMARKS

Drawing the research questions from everyday life observations, the thesis investigated how analyzing the immense amount of real-time and retrospective data can contribute to marketing theories and strategies and how society at large gets influenced by big data analytics. More specifically, the three essays explored big data studies in branding, advertising, and communication.

The first essay examined how brand sentiment gets shared and viral on social media and how machine learning algorithms can be used to more precisely predict them for the purpose of brand management. It proposed a new algorithm to analyze sentiments towards brands on social media. Our study shed light on how brands can quantify sentiments towards them in terms of perceived brand authenticity. We demonstrated its effectiveness via both qualitative and quantitative studies. As part of the qualitative study, we discussed several tweets from our dataset that our coders classified under quality commitment, heritage, uniqueness, or symbolism. A few of the tweets we found were about brand authenticity, but they were not connected to any of our four categories, and they were placed in the none of them category. There were also examples of tweets that were irrelevant to brand authenticity or did not contain enough information to judge brand authenticity. In each category, the common words were extracted using LSA. Finally, both the brand authenticity dimension and its sentiment strength were predicted with high accuracy by SVM analysis.

In the second essay, we studied how apps installed on geo-enabled devices can get benefit from the immense amount of locational data to alter one's destination by providing relevant destination-based advertising. Four studies were designed to examine the effects of spatial distance, NFC, cultural distance, and marketing incentives, the underlying mechanism of susceptibility, and how advertising based on destination helps alter planned behavior. Providing relevant incentives (i.e., free rerouting) as well as decreasing psychological distances can lower the cognitive load. Practically, having access to big datasets of locations and destinations, ridesharing apps (i.e. Uber) could use them to enrich users' journey experiences. Therefore, they should reduce the amount of information that consumers have to process in order to determine whether this ad is something they might be interested in adopting. Different AI-related methods, from data collection to analysis, would be helpful in this regard.

The third essay looked into how service providers could use AI to communicate their service descriptions considering user-generated content, i.e. reviews or other service providers' descriptions, to enhance their social cognition, i.e. higher perceived warmth and competence for Airbnb listing. More specifically, we contributed to the textual analysis literature by linking sales descriptions to consumers' perceptions, which we gathered from their comments and reviews. Results supported the impact of linguistic concreteness on social cognition via listings' sentiment. Hosts were divided into four groups (based on their concreteness and professionalism: novices, juniors, aspiring professionals, and real estate agents, and we made specific recommendations for each.

In addition to research on the positive impacts of big data analytics in marketing, more research is necessary to look at an important but usually neglected field of research, which is big data ethics. Big data's dark-side increasingly shows high potential to disrupt users' lives, companies' brand image, and society's well-fare in so many ways. That includes behavioral surveillance, informed consent and data privacy, algorithms dominance, deep fake, fake reviews and so on. Understanding the dynamic nature of the challenge and its effects on consumer behavior provides novel knowledge for both theory and practice and helps firms implement appropriate marketing strategies.

The findings could contribute to further research in terms of both developing novel methodologies in the field of machine learning and deep learning algorithms, and implicating in social media marketing, brand sentiment, sales, advertising, product harm crises, and information systems. Growing the well-being of society, brands, and consumers in this new digital era requires conducting more research at the intersection of marketing and computer science to explore big data implications for marketing theory and practice.

Bibliography

Aaker, J., Vohs, K. D., & Mogilner, C. (2010). Nonprofits are seen as warm and for-profits as competent: Firm stereotypes matter. *Journal of Consumer Research*, *37*(2), 224-237.

Ahmad, S. N., & Laroche, M. (2017). Analyzing electronic word of mouth: A social commerce construct. *International Journal of Information Management,* (June), 37(3), 202-213..

Ailawadi, K. L., Lehmann, D. R., & Neslin, S. A. (2003). Revenue premium as an outcome measure of brand equity. *Journal of Marketing, 6*7(4), 1-17.

Ajzen, I. (1985). *From intentions to actions: A theory of planned behavior* (pp. 11-39). Springer Berlin Heidelberg.

Alalwan, A. A. (2020). Mobile food ordering apps: An empirical study of the factors affecting customer e-satisfaction and continued intention to reuse. *International Journal of Information Management*, *50,* 28-44.

Appel, G., Grewal, L., Hadi, R., & Stephen, A. T. (2020). The future of social media in marketing. *Journal of the Academy of Marketing Science*, *48*(1), 79-95.

Assiouras, I., Liapati, G., Kouletsis, G., & Koniordos, M. (2015). The impact of brand authenticity on brand attachment in the food industry. *British Food Journal, 117*(2), 538-552.

Austin, J. L. (1975). *How to do things with words*. Oxford university press.

Barbier, G., & Liu, H. (2011). Data mining in social media. *Social network data analytics* (pp. 327-352). Springer.

Bart, Y., Stephen, A. T., & Sarvary, M. (2014). Which products are best suited to mobile advertising? A field study of mobile display advertising effects on consumer attitudes and intentions. *Journal of Marketing Research, 51*(3) 270-285.

Bartikowski, B., & Singh, N. (2014). Should all firms adapt websites to international audiences? *Journal of Business Research*, *67*(3), 246-252.

Bartikowski, B., Taieb, B., & Chandon, J. L. (2016). Targeting without alienating on the Internet: Ethnic minority and majority consumers. *Journal of Business Research*, *69*(3) 1082-1089.

Bearden, W. O., Netemeyer, R. G., & Teel, J. E. (1989). Measurement of consumer susceptibility to interpersonal influence. *Journal of Consumer Research*, *15*(4), 473-481.

Berger, J. Humphreys, A. Ludwig, S. Moe, W. W. Netzer, O. and Schweidel, D. A. (2020). Uniting the tribes: Using text for marketing insight. *Journal of Marketing*, *84*(1), 1-25.

Bernritter, S. F. Verlegh, P. W. and Smit, E. G. (2016). Why nonprofits are easier to endorse on social media: The roles of warmth and brand symbolism. *Journal of Interactive Marketing*, *33*, 27-42.

Beverland, M. B. (2005). Crafting brand authenticity: The case of luxury wines. *Journal of Management Studies, 42*(5), 1003-1029.

Beverland, M. B., & Farrelly, F. J. (2010). The quest for authenticity in consumption: Consumers' purposive choice of authentic cues to shape experienced outcomes. *Journal of Consumer Research, 36*(5), 838-856.

Bifet, A., & Frank, E. (2010, October). Sentiment knowledge discovery in twitter streaming data. In: *International Conference on Discovery Science* (pp. 1-15). Springer Berlin Heidelberg.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological methods & research, 17*(3), 303-316.

Bolton, R. N., Gustafsson, A., Tarasi, C. O., & Witell, L. (2021). Designing satisfying service encounters: website versus store touchpoints. *Journal of the Academy of Marketing Science*, 1-23.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662-679.

Brown, S., & Stayman, D. M. (1992). Antecedents and consequences of attitude toward the ad: A meta-analysis. *Journal of Consumer Research* 1*9*(1), 34-51.

Brown, S., Kozinets, R. V., & Sherry Jr, J. F. (2003). Teaching old brands new tricks: Retro branding and the revival of brand meaning. *Journal of Marketing*, *67*(3), 19-33.

Brumbaugh, A. M. (2002). Source and nonsource cues in advertising and their effects on the activation of cultural and subcultural knowledge on the route to persuasion. *Journal of Consumer Research* 2*9*(2) 258-269.

Bruner, G. C., & Kumar, A. (2007). Attitude toward location-based advertising. *Journal of Interactive Advertising, 7*(2), 3-15.

Bulgarelli, D. and Molina, P. (2016). Social cognition in preschoolers: Effects of early experience and individual differences. *Frontiers in Psychology*, *7*, 1762.

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition*. Data Mining and Knowledge Discovery, 2*(2), 121-167.

Burke, R. R., & Srull, T. K. (1988). Competitive interference and consumer memory for advertising. *Journal of Consumer Research*, *15*(1), 55-68.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116.

Chakraborty, I. Kim, M. and Sudhir, K. (2021). EXPRESS: Attribute Sentiment Scoring with Online Text Reviews: Accounting for Language Structure and Missing Attributes. *Journal of Marketing Research*, 00222437211052500.

Chan, K. W., Li, S. Y., & Zhu, J. J. (2015). Fostering customer ideation in crowdsourcing community: The role of peer-to-peer and peer-to-firm interactions. *Journal of Interactive Marketing*, *31*, 42-62.

Chen, C. P., & Zhang, C. (2014). Data-intensive applications, challenges techniques and technologies: A survey on big data. *Information Sciences, 275*, 314-347.

Chen, C. Y. Mathur, P. and Maheswaran, D. (2014). The effects of country-related affect on product evaluations. *Journal of Consumer Research*, *41*(4), 1033-1046.

Chen, Q., Clifford, S. J., & Wells, W. D. (2002). Attitude toward the site II: New information. *Journal of Advertising Research*, *42*(2), 33-45.

Cho, C., & Cheon, H. J. (2004). Why do people avoid advertising on the internet*? Journal of Advertising, 33*(4), 89-97.

Choi, H., Ko, E., Kim, E. Y., & Mattila, P. (2015). The role of fashion brand authenticity in product management: A holistic marketing approach. *Journal of Product Innovation Management, 32*(2), 233-242.

Chu, S., & Kim, Y. (2011). Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites. *International Journal of Advertising, 3*0(1), 47-75.

Chung, J. Johar, G. V. Li, Y. Netzer, O. and Pearson, M. (2021). Mining Consumer Minds: Downstream Consequences of Host Motivations for Home Sharing Platforms. *Journal of Consumer Research*, *48*(5), 817-838.

Cleveland, M., Rojas-Méndez, J. I., Laroche, M., & Papadopoulos, N. (2016). Identity, culture, dispositions and behavior: A cross-national examination of globalization and culture change. *Journal of Business Research*, *69*(3) 1090-1102.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

Danziger, S., Montal, R., & Barkan, R. (2012). Idealistic advice and pragmatic choice: A psychological distance account. *Journal of Personality and Social Psychology*, 102(6), 1105.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 4*1(6), 391.

Del Chiappa, G. Pung, J. M. Atzeni, M., & Sini, L. (2021). What prevents consumers that are aware of Airbnb from using the platform? A mixed methods approach. *International Journal of Hospitality Management*, *93*, 102775.

Donath, J. (2021). Commentary: The Ethical Use of Powerful Words and Persuasive Machines. *Journal of Marketing*, *85*(1), 160-162.

Driscoll, K., & Walker, S. (2014). Big data, big questions working within a black box: Transparency in the collection and production of big twitter data. *International Journal of Communication, 8*, 20.

Drolet, A., & Frances Luce, M. (2004). The rationalizing effects of cognitive load on emotion-based trade-off avoidance. *Journal of Consumer Research*, *31*(1), 63-77.

Ducoffe, R. H. (1995). How consumers assess the value of advertising. *Journal of Current Issues and Research in Advertising 17*(1) 1-18.

Eggers, F., O'Dwyer, M., Kraus, S., Vallaster, C., & Güldenberg, S. (2013). The impact of brand authenticity on brand trust and SME growth: A CEO perspective. *Journal of World Business, 48*(3), 340-348.

Eisend, M. (2008). Explaining the impact of scarcity appeals in advertising: The mediating role of perceptions of susceptibility. *Journal of Advertising*, *37*(3), 33-40.

Fiske, S. T. Cuddy, A. J. and Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77-83.

Fournier, S. and Alvarez, C. (2012). Brands as relationship partners: Warmth, competence, and in-between. *Journal of Consumer Psychology*, *22*(2), 177-185.

Fournier, S., & Avery, J. (2011). The uninvited brand. *Business Horizons, 54*(3), 193-207.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137-144.

Gaspar, R., Pedro, C., Panagiotopoulos, P., & Seibt, B. (2016). Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events. *Computers in Human Behavior, 5*6, 179-191.

Geiger, A. Horbel, C. and Germelmann, C. C. (2018). “Give and take”: how notions of sharing and context determine free peer-to-peer accommodation decisions. *Journal of Travel and Tourism Marketing*, *35*(1), 5-15.

Gensler, S., Völckner, F., Liu-Thompkins, Y., & Wiertz, C. (2013). Managing brands in the social media environment. *Journal of Interactive Marketing, 27*(4), 242-256.

Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with Applications, 40*(16), 6266-6282.

Ghiassi, M., Zimbra, D., & Lee, S. (2016). Targeted Twitter Sentiment Analysis for Brands Using Supervised Feature Engineering and the Dynamic Architecture for Artificial Neural Networks. *Journal of Management Information Systems*, *33*(4), 1034-1058.

Ghose, A., Kwon, H. E., Lee, D., & Oh, W. (2019). Seizing the commuting moment: Contextual targeting based on mobile transportation apps. *Information Systems Research*, *30*(1) 154-174.

Ghose, A., Li, B., & Liu, S. (2019). Mobile targeting using customer trajectory patterns. *Management Science*, *65*(11), 5027-5049.

Gibbs, C. Guttentag, D. Gretzel, U. Morton, J. and Goodwill, A. (2018). Pricing in the sharing economy: A hedonic pricing model applied to Airbnb listings. *Journal of Travel and Tourism Marketing*, *35*(1), 46-56.

Gilmore, J. H., & Pine, B. J. (2007). *Authenticity: What consumers really want?* Harvard Business Press.

Goldfarb, A., & Tucker, C. (2011). Online display advertising: Targeting and obtrusiveness. *Marketing Science*, *30*(3), 389-404.

Gopaldas, A. (2014). Marketplace sentiments. *Journal of Consumer Research, 41*(4), 995-1014.

Gopinath, M., & Nyer, P. U. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science* 27(2) 184-206.

Gordon, B. R., Zettelmeyer, F., Bhargava, N., & Chapsky, D. (2019). A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook. *Marketing Science*, *38*(2) 193-225.

Grewal, D., Bart, Y., Spann, M., & Zubcsek, P. P. (2016). Mobile advertising: A framework and research agenda. *Journal of Interactive Marketing*, *34*, 3-14.

Grewal, R. Gupta, S. and Hamilton, R. (2021). Marketing Insights from Multimedia Data: Text, Image, Audio, and Video. *Journal of Marketing Research*, *58*(6), 1025-1033.

Habibi, M. R., Laroche, M., & Richard, M. (2014). Brand communities based in social media: How unique are they? Evidence from two exemplary brand communities. *International Journal of Information Management, 34*(2), 123-132.

Hair Jr, J. F. and Sarstedt, M. (2021). Data, measurement, and causal inferences in machine learning: opportunities and challenges for marketing. *Journal of Marketing Theory and Practice*, *29*(1), 65-77.

Halkias, G. and Diamantopoulos, A. (2020). Universal dimensions of individuals' perception: Revisiting the operationalization of warmth and competence with a mixed-method approach. *International Journal of Research in Marketing*, *37*(4), 714-736.

Harris, J., & Blair, E. A. (2006). Consumer preference for product bundles: The role of reduced search costs. *Journal of the Academy of Marketing Science*, *34*(4), 506-513.

Haugtvedt, C. P., Petty, R. E., & Cacioppo, J. T. (1992). Need for cognition and advertising: Understanding the role of personality variables in consumer behavior. *Journal of Consumer Psychology*, *1*(3), 239-260.

He, W., Wu, H., Yan, G., Akula, V., & Shen, J. (2015). A novel social media competitive analytics framework with sentiment benchmarks. *Information & Management, 52*(7), 801-812.

Herrmann, L. K., & Kim, J. (2017). The fitness of apps: a theory-based examination of mobile fitness app usage over 5 months. *Mhealth*, *3*(1), 1-9.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*(1), 104-121.

Hofacker, C. F., Malthouse, E. C., & Sultan, F. (2016). Big data and consumer behavior: Imminent opportunities. *Journal of Consumer Marketing, 3*3(2), 89-97.

Hovy, D. Melumad, S. and Inman, J. J. (2021). Wordify: A Tool for Discovering and Differentiating Consumer Vocabularies. *Journal of Consumer Research*, *48*(3), 394-414.

Hu, N., Koh, N. S., & Reddy, S. K. (2014). Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decision Support Systems, 5*7, 42-53.

Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, *49*(1), 30-50.

Huberty, M. (2015). Can we vote with our tweet? On the perennial difficulty of election forecasting with social media. *International Journal of Forecasting, 31*(3), 992–1007.

Humphreys, A. and Wang, R. J. H. (2018). Automated text analysis for consumer research. *Journal of Consumer Research*, *44*(6), 1274-1306.

Hutto, C. and Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, 8, 1, 216-225).

Internet live stats (2015). Twitter statistics. Retrieved July 14, 2015, from http://www.internetlivestats.com/twitter-statistics/

Izquierdo-Yusta, A., Olarte-Pascual, C., & Reinares-Lara, E. (2015). Attitudes toward mobile advertising among users versus non-users of the mobile internet. *Telematics and Informatics, 32*(2), 355-366.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology, 60*(11), 2169-2188.

Johnson, A. R., Thomson, M., & Jeffrey, J. (2015). What does brand authenticity mean? Causes and consequences of consumer scrutiny toward a brand narrative. *Brand Meaning Management, 12*, 1-27.

Johnson-Laird, P. N., & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. *Cognition and Emotion*, *3*(2), 81-123.

Kadirov, D. (2015). Private labels ain't bona fide! Perceived authenticity and willingness to pay a price premium for national brands over private labels. *Journal of Marketing Management, 31*(17-18), 1773-1798.

Kanemoto, E., & Dai, Z. (2015). Who cares about their names: Case study in Starbucks. Ray Browne Conference on Cultural and Critical Studies, 3

Kang, H., Hahn, M., Fortin, D. R., Hyun, Y. J., & Eom, Y. (2006). Effects of perceived behavioral control on the consumer usage intention of e-coupons. *Psychology & Marketing, 23*(10), 841-864.

Kannan, P. K. (2017). Digital marketing: A framework, review and research agenda. *International Journal of Research in Marketing*, *34*(1) 22-45.

Kim, B., & Sullivan, M. W. (1998). The effect of parent brand experience online extension trial and repeat purchase. *Marketing Letters,* 9(2), 181-193.

Kim, C., Li, W., & Kim, D. J. (2015). An empirical analysis of factors influencing M-shopping use. *International Journal of Human-Computer Interaction, 31*(12), 974-994.

Kim, K. J. (2014). Can smartphones be specialists? Effects of specialization in mobile advertising. *Telematics and Informatics, 31*(4), 640-647.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Klein, J. G., Smith, N. C., & John, A. (2004). Why we boycott: Consumer motivations for boycott participation. *Journal of Marketing, 6*8(3), 92-109.

Ko, D., Seo, Y., & Jung, S. U. (2015). Examining the effect of cultural congruence, processing fluency, and uncertainty avoidance in online purchase decisions in the US and Korea. *Marketing Letters* 2*6*(3), 377-390.

Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. (2016). From social to sale: The effects of firm-generated content in social media on customer behavior. *Journal of Marketing, 8*0(1), 7-25.

Laroche, M., Habibi, M. R., & Richard, M. (2013). To be or not to be in social media: How brand loyalty is affected by social media? *International Journal of Information Management, 33*(1), 76-82.

Lee, H., Han, J., & Suh, Y. (2014). Gift or threat? An examination of voice of the customer: The case of MyStarbucksIdea.com. *Electronic Commerce Research and Applications, 13*(3), 205-219.

Lee, J., & Hong, I. B. (2016). Predicting positive user responses to social media advertising: The roles of emotional appeal, informativeness, and creativity. *International Journal of Information Management, 36*(3), 360-373.

Lee, M. (2009). Factors influencing the adoption of internet banking: An integration of TAM and TPB with perceived risk and perceived benefit. *Electronic Commerce Research and Applications,* 8(3) 130-141.

Lee, S., Kim, K. J., & Sundar, S. S. (2015). Customization in location-based advertising: Effects of tailoring source, locational congruity, and product involvement on ad attitudes. *Computers in Human Behavior, 51,* Part A, 336-343.

Lee, T. Y., & Bradlow, E. T. (2011). Automated marketing research using online customer reviews. *Journal of Marketing Research, 48*(5), 881-894.

Li, H., Edwards, S. M., & Lee, J. (2002). Measuring the intrusiveness of advertisements: Scale development and validation. *Journal of Advertising, 31*(2), 37-47.

Li, J., Moreno, A., & Zhang, D. J. (2019). Agent pricing in the sharing economy: Evidence from Airbnb. In *Sharing economy* (pp. 485-503). Springer, Cham.

Liberman, N., & Trope, Y. (2014). Traversing psychological distance. *Trends in cognitive sciences*, 18(7), 364-369.

Lichtenstein, D. R., Netemeyer, R. G., & Burton, S. (1995). Assessing the domain specificity of deal proneness: A field study. *Journal of Consumer Research 22*(3), 314-326.

Lin, T. T., Paragas, F., Goh, D., & Bautista, J. R. (2016). Developing location-based mobile advertising in Singapore: A socio-technical perspective. *Technological Forecasting and Social Change 103*, 334-349.

Liu, C., Sinkovics, R. R., Pezderka, N., & Haghirian, P. (2012). Determinants of consumer perceptions toward mobile advertising: A comparison between Japan and Austria. *Journal of Interactive Marketing 26*(1) 21-32.

Lu, H., & Yu-Jen Su, P. (2009). Factors affecting purchase intention on mobile shopping web sites. *Internet Research 19*(4), 442-458.

Luo, X. Qin, M. S. Fang, Z. and Qu, Z. (2021). Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions. *Journal of Marketing*, *85*(2), 14-32.

Luo, X., Andrews, M., Fang, Z., & Phang, C. W. (2014). Mobile targeting. *Management Science*, *60*(7) 1738-1756.

Ma, S. Cui, X. Xiao, X. and Zhao, X. (2022). The impact of photo verification service on sales performance in the peer-to-peer economy: Moderating role of customer uncertainty. *Journal of Business Research*, *142*, 45-55.

MacKenzie, S. B., & Lutz, R. J. (1989). An empirical examination of the structural antecedents of attitude toward the ad in an advertising pretesting context. *Journal of Marketing, 53*(2), 48-65.

Mahrt, M., & Scharkow, M. (2013). The value of big data in digital media research. *Journal of Broadcasting & Electronic Media, 57*(1), 20-33.

Marafino, B. J., Davies, J. M., Bardach, N. S., Dean, M. L., & Dudley, R. A. (2014). N-gram support vector machines for scalable procedure and diagnosis classification, with applications to clinical free text data from the intensive care unit. *Journal of the American Medical Informatics Association, 21*(5), 871-875.

McGuire, W. (1968). Personality and susceptibility to social influence. In E. F. Borgatta, & W. W. Lambert (Eds.), *Handbook of personality theory and research* (pp. 1130–1187). Chicago: Rand McNally.

Merisavo, M., Kajalo, S., Karjaluoto, H., Virtanen, V., Salmenkivi, S., Raulas, M., & Leppäniemi, M. (2007). An empirical study of the drivers of consumer acceptance of mobile advertising. *Journal of Interactive Advertising,* 7(2), 41-50.

Miller, F. M. (2015). Ad authenticity: An alternative explanation of advertising's effect on established brand attitudes. *Journal of Current Issues and Research in Advertising, 36*(2) 177-194.

Milstein, S., Lorica, B., Magoulas, R., Hochmuth, G., Chowdhury, A., & O'Reilly, T. (2008). *Twitter and the micro-messaging revolution: Communication, connections, and immediacy--140 characters at a time*. O'Reilly Media, Incorporated.

Moody, G. D., Galletta, D. F., & Dunn, B. K. (2017). Which phish get caught? An exploratory study of individuals′ susceptibility to phishing. *European Journal of Information Systems*, *26*(6), 564-584.

Morhart, F., Malär, L., Guèvremont, A., Girardin, F., & Grohmann, B. (2015). Brand authenticity: An integrative framework and measurement scale. *Journal of Consumer Psychology, 25*(2), 200-218.

Mostafa, M. M. (2013). More than words: Social networks' text mining for consumer brand sentiments. *Expert Systems with Applications, 40*(10), 4241-4251.

Muehling, D. D. (1987). Comparative advertising: The influence of attitude-toward-the-ad on brand evaluation. *Journal of Advertising 16*(4), 43-49.

Napoli, J., Dickinson, S. J., Beverland, M. B., & Farrelly, F. (2014). Measuring consumer-based brand authenticity. *Journal of Business Research, 67*(6), 1090-1098.

Naylor, R. W., Lamberton, C. P., & West, P. M. (2012). Beyond the "like" button: The impact of mere virtual presence on brand evaluations and purchase intentions in social media settings. *Journal of Marketing, 76*(6), 105-120.

Netemeyer, R. G., Krishnan, B., Pullig, C., Wang, G., Yagci, M., Dean, D., ... & Wirth, F. (2004). Developing and validating measures of facets of customer-based brand equity. *Journal of Business Research*, *57*(2), 209-224.

Noguera, J. M., Barranco, M. J., Segura, R. J., & Martínez, L. (2012). A mobile 3D-GIS hybrid recommender system for tourism. *Information Sciences 215*, 37-52.

Nysveen, H., Pedersen, P. E., &Thorbjørnsen, H. (2005). Intentions to use mobile services: Antecedents and cross-service comparisons. *Journal of the Academy of Marketing Science, 33*(3), 330-346.

Olah, C. (2015). Understanding LSTM Networks. Colah's blog, Retrieved February 28, 2022, from http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior, 31*, 527-541.

Packard, G. and Berger, J. (2021). How Concrete Language Shapes Customer Satisfaction. *Journal of Consumer Research*, *47*(5), 787-806.

Park, E. (2019). The role of satisfaction on customer reuse to airline services: An application of Big Data approaches. *Journal of Retailing and Consumer Services*, *47*, 370-374.

Peng, L., Cui, G., Bao, Z., & Liu, S. (2021). Speaking the same language: the power of words in crowdfunding success and failure. *Marketing Letters*, 1-13.

Pine, B. J., Peppers, D., & Rogers, M. (1995). *Do you want to keep your customers forever?* Harvard Business Press.

Pitt, C., Plangger, K., & Eriksson, T. (2021). Accommodation eWOM in the sharing economy: automated text comparisons from a large sample. *Journal of Hospitality Marketing and Management*, *30*(2), 258-275.

Plangger, K., & Montecchi, M. (2020). Thinking beyond privacy calculus: Investigating reactions to customer surveillance. *Journal of Interactive Marketing*, *50*, 32-44.

Ranchordás, S. (2015). Does sharing mean caring: Regulating innovation in the sharing economy. *Minnesota Journal of Law Science & Technology. 16*, 413.

Richard, M. O. (2005). Modeling the impact of internet atmospherics on surfer behavior. *Journal of Business Research*, *58*(12) 1632-1642.

Richard, M. O., & Habibi, M. R. (2016). Advanced modeling of online consumer behavior: The moderating roles of hedonism and culture. *Journal of Business Research*, *69*(3) 1103-1119.

Rogers, T., & Bazerman, M. H. (2008). Future lock-in: Future implementation increases selection of 'should' choices. *Organizational Behavior and Human Decision Processes*, *106*(1), 1-20.

San Martín, H., & Herrero, Á. (2012). Influence of the user's psychological factors on the online purchase intention in rural tourism: Integrating innovativeness to the UTAUT framework. *Tourism Management*, *33*(2), 341-350.

Saxena, S. (2020). Retrieved February 28, 2022, from https://medium.com/analytics-vidhya/understanding-embedding-layer-in-keras-bbe3ff1327ce

Schallehn, M., Burmann, C., & Riley, N. (2014). Brand authenticity: Model development and empirical testing. *Journal of Product & Brand Management, 23*(3), 192-199.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.

Schmitt, B. (2019). From atoms to bits and back: A research curation on digital technology and agenda for future research. *Journal of Consumer Research*, *46*(4), 825-832.

Sheth, J. (2021). New areas of research in marketing strategy, consumer behavior, and marketing analytics: the future is bright. *Journal of Marketing Theory and Practice*, *29*(1), 3-12.

Shi, Z., Rui, H., & Whinston, A. B. (2013). Content sharing in a social broadcasting environment: evidence from twitter. *MIS Quarterly*, *38*(1), 123-142.

Shirdastian, H., Laroche, M., & Richard, M. O. (2019). Using big data analytics to study brand authenticity sentiments: The case of Starbucks on Twitter. *International Journal of Information Management*, *48* 291-307.

Shobeiri, S., Mazaheri, E., & Laroche, M. (2018). Creating the right customer experience online: The influence of culture. *Journal of Marketing Communications* 2*4*(3) 270-290.

Shumanov, M. Cooper, H. and Ewing, M. (2021). Using AI predicted personality to enhance advertising effectiveness. *European Journal of Marketing*, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/EJM-12-2019-0941

Shusterman, A. (2020). Retrieved February 28, 2022, from https://datascience.stackexchange.com/questions/83984/how-to-choose-dimension-of-keras-embedding-layer

Sidorov, G., Velasquez, F., Stamatatos, E., Gelbukh, A., & Chanona-Hernández, L. (2014). Syntactic N-grams as machine learning features for natural language processing. *Expert Systems with Applications, 41*(3), 853-860.

Sidorov, G., Velasquez, F., Stamatatos, E., Gelbukh, A., & Chanona-Hernández, L. (2014). Syntactic n-grams as machine learning features for natural language processing. *Expert Systems with Applications*, *41*(3), 853-860.

Smith, A. N., Fischer, E., & Yongjian, C. (2012). How does brand-related user-generated content differ across YouTube, Facebook, and Twitter? *Journal of Interactive Marketing, 26*(2), 102-113.

Social bakers (2015), Starbucks statistics. Retrieved July 14, 2015, from http://www.socialbakers.com/statistics/twitter/profiles/detail/30973-starbucks

Song, R., Moon, S., Chen, H. A., & Houston, M. B. (2018). When marketing strategy meets culture: the role of culture in product evaluations. *Journal of the Academy of Marketing Science*, *46*(3), 384-402.

Spiggle, S., Nguyen, H. T., & Caravella, M. (2012). More than fit: brand extension authenticity. *Journal of Marketing Research*, *49*(6), 967-983.

Steenkamp, J. B. E., & Maydeu-Olivares, A. (2015). Stability and change in consumer traits: evidence from a 12-year longitudinal study 2002–2013. *Journal of Marketing Research*, *52*(3) 287-308.

Steinhoff, L., Arli, D., Weaven, S., & Kozlenkova, I. V. (2019). Online relationship marketing. *Journal of the Academy of Marketing Science*, *47*(3), 369-393.

Stryker, S., & Burke, P. J. (2000). The past, present, and future of an identity theory. *Social Psychology Quarterly, 63*(4), 284-297.

Sultan, F., Rohm, A. J., & Gao, T. (2009). Factors influencing consumer acceptance of mobile marketing: A two-country study of youth markets. *Journal of Interactive Marketing 23*(4), 308-320.

Tellis, G. J., Chandy, R. K., MacInnis, D., & Thaivanich, P. (2005). Modeling the microeffects of television advertising: Which ad works, when, where, for how long, and why? *Marketing Science*, 359-366.

Thompson, C. J., Rindfleisch, A., & Arsel, Z. (2006). Emotional branding and the strategic value of the doppelgänger brand image. *Journal of Marketing*, *70*(1), 50-64.

Touré-Tillery, M. and McGill, A. L. (2015). Who or what to believe: Trust and the differential persuasiveness of human and anthropomorphized messengers. *Journal of Marketing*, *79*(4), 94-110.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review 117*(2), 440.

Tussyadiah, I. P. and Park, S. (2018). When guests trust hosts for their words: Host description and trust in sharing economy. *Tourism Management*, *67*, 261-272.

Twitter counter (2015), Starbucks. Retrieved July 14, 2015, from http://twittercounter.com/Starbucks

Vasisht, D., Kumar, S., & Katabi, D. (2016). Decimeter-level localization with a single wifi access point. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)* (pp. 165-178).

Walsh, M. G. (2013). Protecting your brand against the heartbreak of genericide. *Business Horizons, 56*(2), 159-166.

Wang, X., He, J., Curry, D. J., & Ryoo, J. H. (2021). Attribute Embedding: Learning Hierarchical Representations of Product Attributes from Consumer Reviews. *Journal of Marketing*, 00222429211047822.

Wang, Z. Mao, H. Li, Y. J. and Liu, F. (2017). Smile big or not? Effects of smile intensity on perceptions of warmth and competence. *Journal of Consumer Research*, *43*(5), 787-805.

Weinstein, N. D. (1984). Why it won't happen to me: perceptions of risk factors and susceptibility. *Health psychology*, *3*(5), 431.

Wikipedia (2015). Starbucks. Retrieved July 14, 2015, from https://en.wikipedia.org/wiki/Starbucks

Wilson, J., Crisp, C. B., & Mortensen, M. (2013). Extending construal-level theory to distributed groups: Understanding the effects of virtuality. *Organization Science* 2*4*(2), 629-644.

Wolin, L. D., Korgaonkar, P., & Lund, D. (2002). Beliefs, attitudes and behavior towards web advertising. *International Journal of Advertising 21*(1), 87-113.

Woolley, K. and Sharif, M. A. (2021). Incentives Increase Relative Positivity of Review Content and Enjoyment of Review Writing. *Journal of Marketing Research*, *58*(3), 539-558.

Wright, S., & Calof, J. L. (2006). The quest for competitive, business and marketing intelligence: A country comparison of current practices. *European Journal of Marketing, 40*(5/6), 453-465.

Xu, D. J., Liao, S. S., & Li, Q. (2008). Combining empirical experimentation and modeling techniques: A design research approach for personalized mobile advertising applications. *Decision Support Systems, 44*(3), 710-724.

Xu, H., Teo, H., Tan, B. C., & Agarwal, R. (2009). The role of push-pull technology in privacy calculus: The case of location-based services. *Journal of Management Information Systems 26*(3) 135-174.

Xue, A. Y., Zhang, R., Zheng, Y., Xie, X., Huang, J., & Xu, Z. (2013, April). Destination prediction by sub-trajectory synthesis and privacy protection against such prediction. In *2013 IEEE 29th international conference on data engineering (ICDE)* (pp. 254-265). IEEE.

Yang, K. (2010). Determinants of US consumer mobile shopping services adoption: Implications for designing mobile shopping services. *Journal of Consumer Marketing 27*(3) 262-270.

Yang, K. (2012). Consumer technology traits in determining mobile shopping adoption: An application of the extended theory of planned behavior. *Journal of Retailing and Consumer Services 19*(5), 484-491.

Yun, T., Sim, K., & Kim, H. (2000). Support vector machine-based inspection of solder joints using circular illumination. *Electronics Letters, 3*6(11), 949-951.

Zhang, Y. (1996). Responses to humorous advertising: The moderating effect of need for cognition. *Journal of Advertising*, *25*(1), 15-32.

Zhang, Y., & Buda, R. (1999). Moderating effects of need for cognition on responses to positively versus negatively framed advertising messages. *Journal of Advertising*, *28*(2), 1-15.

Zhou, X., & Tuck, D. P. (2007). MSVM-RFE: Extensions of SVM-RFE for multiclass gene selection on DNA microarray data. *Bioinformatics (Oxford, England), 2*3(9), 1106-1114.

Appendices

Table A summarizes the measurement constructs, items, and Cronbach's alpha for Paper 2.

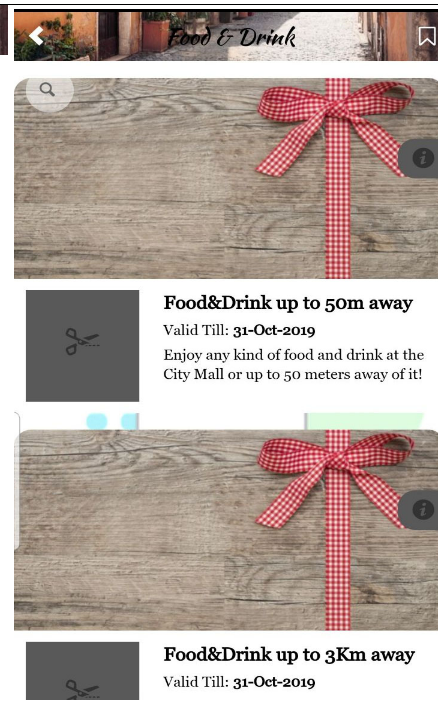
Table A- Measurement constructs, items, and Cronbach's alpha

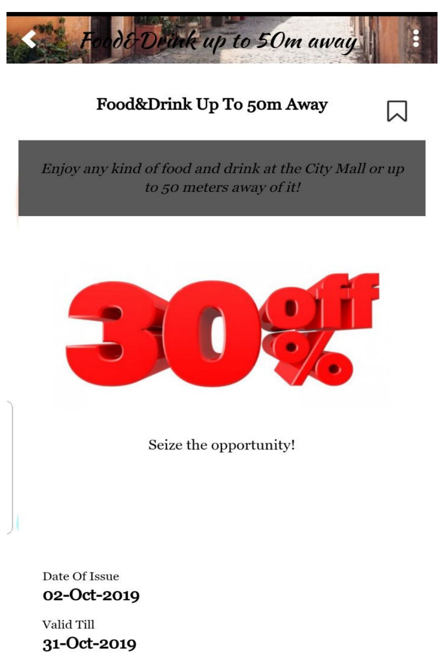
Constructs	**Items**	**Cronbach's alpha**	**Adapted from**
Attitudes toward DBA	I do not like destination-based advertising in general. Overall, I like destination-based advertising.	.899	
Need for cognition	I like to have the responsibility of handling a situation that requires a lot of thinking. I really enjoy a task that involves coming up with new solutions to problems. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought. I would prefer complex to simple problems.	.924	**Cacioppo and Petty (1982)**
Susceptibility	Since I have little experience with the app, I have to ask my friends about it. I often consult other people to help choose the best alternative available from a range of apps. I frequently gather information from friends or family about an app before I use.	.901	**Bearden *et al.* (1989)**
Purchase intention	If I received a destination-based advertisement, I intend to purchase the advertised product/service. Given that I got the destination-based advertisement, I predict that I would purchase the advertised product/service.	.882	**Lu and Yu-Jen Su (2009)**
App reuse intention	I am intended to use this app in the future … Unlikely/Likely Improbable/Probable Impossible/Possible No chance/Certain	.975	
Coupon redemption	Are you going to redeem the coupon? Unlikely/Likely No chance/Certain	.947	
Cultural distance (manipulation check)	This app reflects typical Canadian (international) aspects. The images, colors, and symbols on this app remind me of Canada (an international app) This app is designed to target Canadian (international) consumers in Canada.	.731	**Singh *et al.* (2006)**
Deal proneness (manipulation check)	Buying products with cents-off deals or straight discounts makes me feel good. Beyond the money I save, buying products that offer a rebate gives me a sense of joy.	.736	**Lichtenstein *et al.* (1995**
Self-ethnicity identification	I identify myself as a Canadian.		

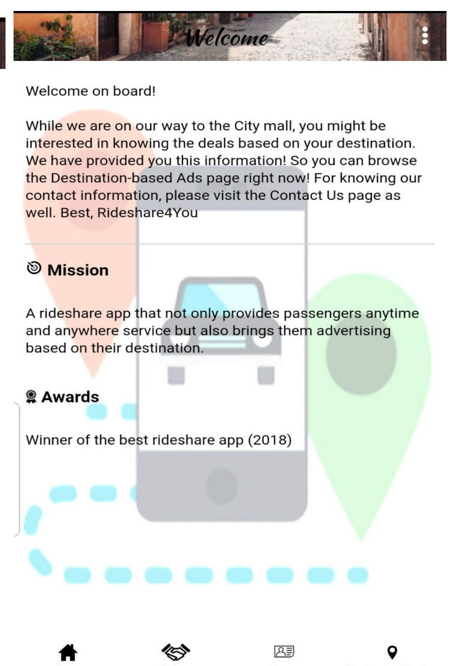
Table B shows the research conditions and stimuli for Paper 2.

Table B- Research conditions and stimuli

<table>
<tr><th></th><th>Stimuli</th></tr>
<tr><td>High cultural distance</td><td>

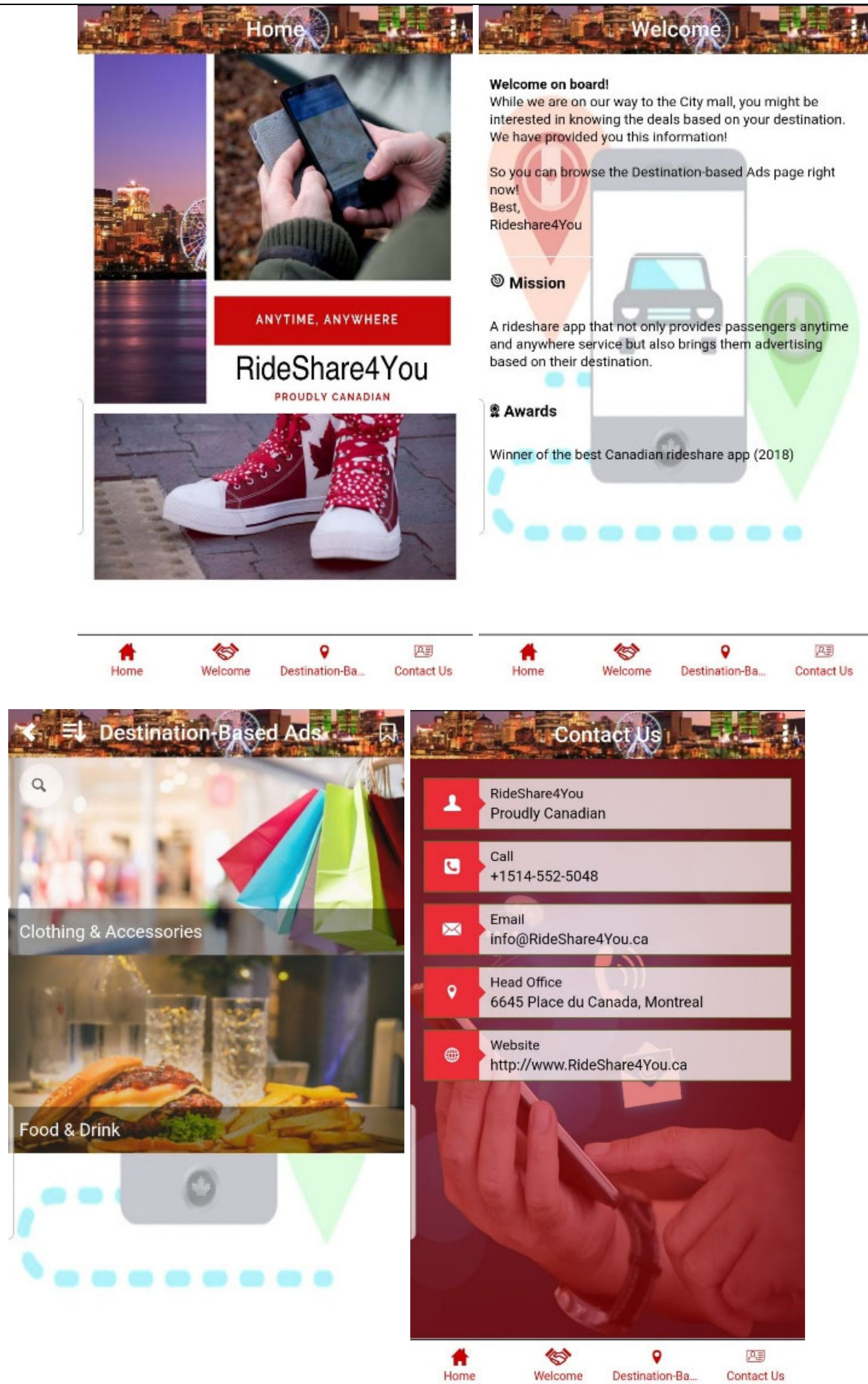



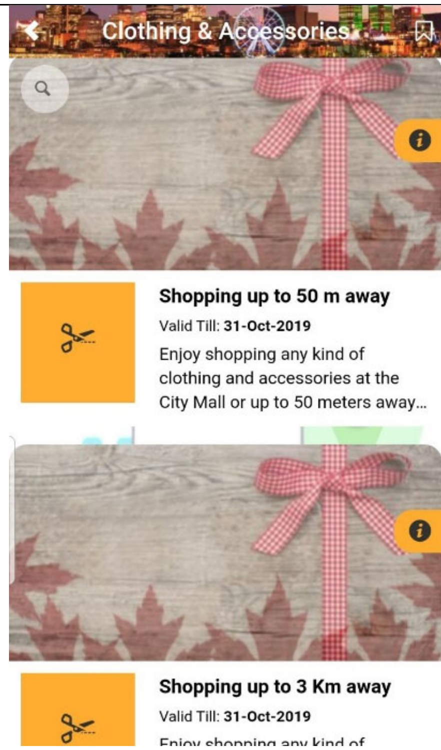


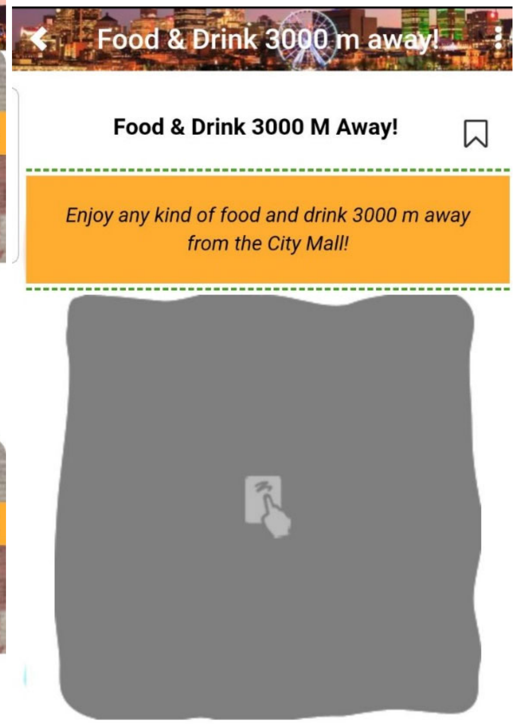


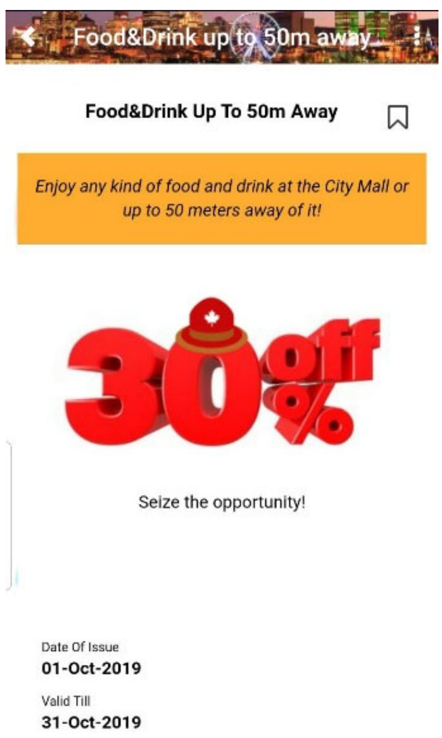
</td></tr>
</table>

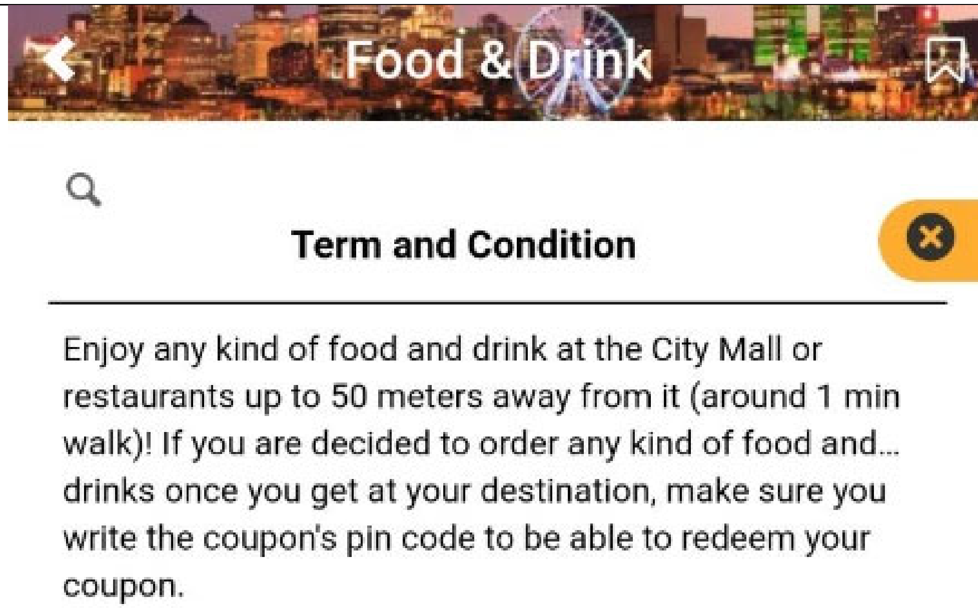
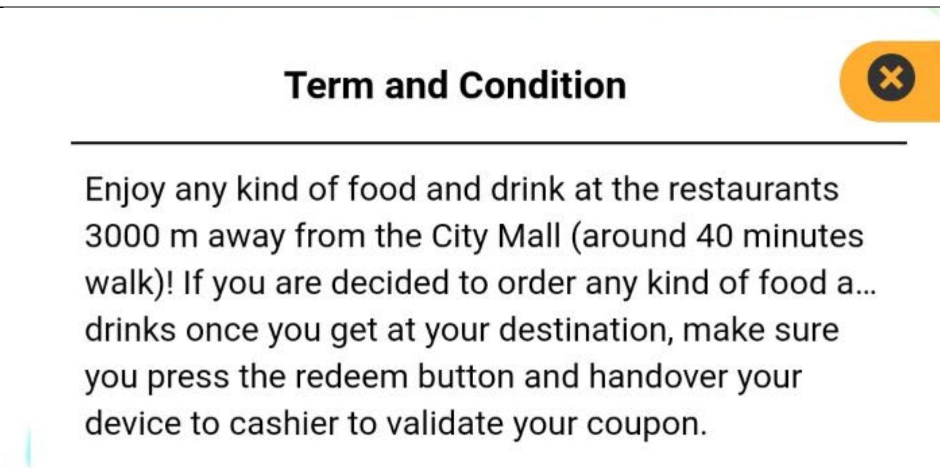
Low cultural distance




Clothing & Accessories
Shopping up to 50 m away
Valid Till: 31-Oct-2019
Enjoy shopping any kind of clothing and accessories at the City Mall or up to 50 meters away...
Shopping up to 3 Km away
Valid Till: 31-Oct-2019


Food & Drink 3000 m away!
Food & Drink 3000 M Away!
Enjoy any kind of food and drink 3000 m away from the City Mall!


Food&Drink up to 50m away
Food&Drink Up To 50m Away
Enjoy any kind of food and drink at the City Mall or up to 50 meters away of it!
30% off
Seize the opportunity!
Date Of Issue
01-Oct-2019
Valid Till
31-Oct-2019

<table>
<tr><td>Low spatial distance</td><td>
</td></tr>
<tr><td>High spatial distance</td><td>
</td></tr>
</table>

High marketing incentives	**Term and Condition** Enjoy purchasing any kind of clothes and accessories 3000 meters away from the City mall! Usually, rerouting costs for our users but using current ride's... free rerouting, you can go to the new destination for free. If you are decided to purchase any kind of clothes and accessories once you get at your new destination, make sure you press the redeem button and handover your device to cashier to validate your coupon.
Low marketing incentives	**Term and Condition** Enjoy purchasing any kind of clothes and accessories 3000 meters away from the City mall! You can go to the new destination by paying an extra fee added to... the top of current rideshare fee. If you are decided to go and purchase any kind of clothes and accessories once you get at your new destination, make sure you press the redeem button and handover your device to cashier to validate your coupon.

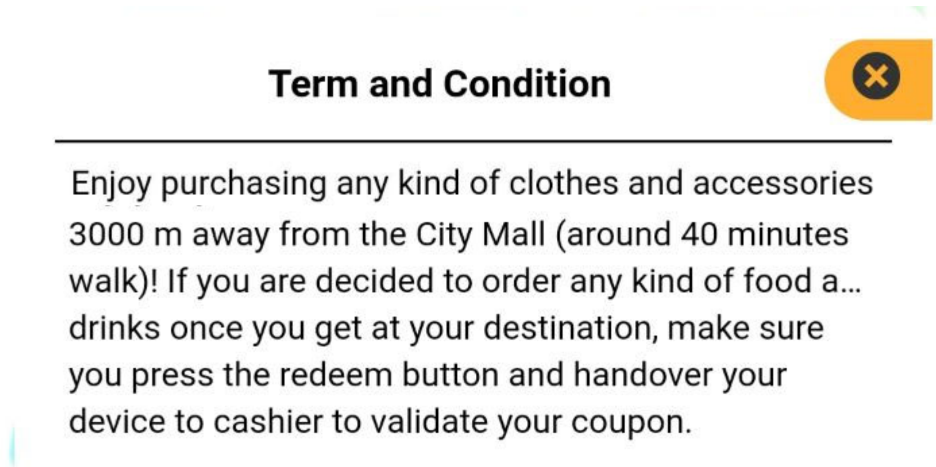
Control	Term and Condition Enjoy purchasing any kind of clothes and accessories 3000 m away from the City Mall (around 40 minutes walk)! If you are decided to order any kind of food a... drinks once you get at your destination, make sure you press the redeem button and handover your device to cashier to validate your coupon.

Figure C in paper 3 shows a professional host profile with 212 listings on the Airbnb website.

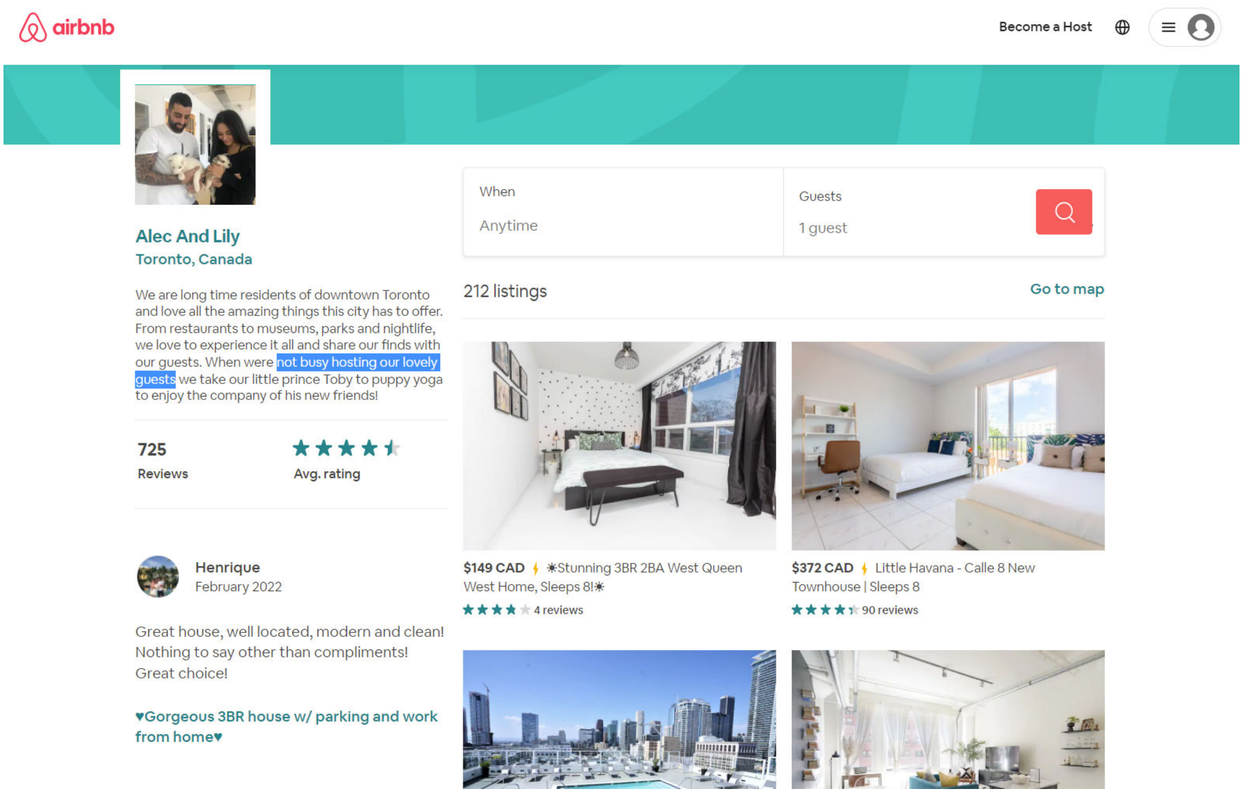


Figure C: A professional host profile on Airbnb