International Journal of Research in Marketing 34 (2017) 265–285

Contents lists available at ScienceDirect

IJRM

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

A picture is worth a thousand words: Translating product reviews into a product positioning map

Sangkil Moon [a], Wagner A. Kamakura [b]

[a] The University of North Carolina at Charlotte, Charlotte, NC 28223, United States
[b] Jones Graduate School of Business, Rice University, 6100 Main Street, Houston, TX 77005, United States

ARTICLE INFO

Article history:
First received on June 1, 2015 and was under review for 10 months
Available online 9 June 2016

Area Editor: Oded Netzer

Keywords:
Online product reviews
Product positioning map
Text mining
Psychometrics
Experience products
Consumer segmentation

ABSTRACT

Product reviews are becoming ubiquitous on the Web, representing a rich source of consumer information on a wide range of product categories (e.g., wines and hotels). Importantly, a product review reflects not only the perception and preference for a product, but also the acuity, bias, and writing style of the reviewer. This *reviewer* aspect has been overlooked in past studies that have drawn inferences about brands from online product reviews. Our framework combines ontology learning-based text mining and psychometric techniques to translate online product reviews into a product positioning map, while accounting for the idiosyncratic responses and writing styles of individual reviewers or a manageable number of reviewer groups (i.e., meta-reviewers). Our empirical illustrations using wine and hotel reviews demonstrate that a product review reveals information about the reviewer (for the wine example with a small number of expert reviewers) or the meta-reviewer (for the hotel example with an enormous number of reviewers reduced to a manageable number of meta-reviewers), as well as about the product under review. From a managerial perspective, product managers can use our framework focusing on meta-reviewers (e.g., traveler types and hotel reservation websites in our hotel example) as a way to obtain insights into their consumer segmentation strategy.

1. Introduction

Multiple online reviews on the same specific product contrast individual consumers' perceptions, preferences, and writing styles. Although we may find common product features evaluated by multiple reviewers, each reviewer tends to select a different set of product features that each thinks are important, based on his or her own consumption experiences. Even when they write about the same product feature, their descriptions and assessments of the feature diverge. Overall, they differ in the way they express their opinions on the same product, resulting in distinctively different words among multiple reviews. Thus, each review reflects not only the perceptions and preferences for product features but also the acuity, bias, and writing style of the reviewer. In brief, each review contains critical information about the product, as well as the reviewer. With such individual differences in mind, we propose a framework to translate online product reviews into a product positioning map that parses out the perceptions and preferences for competing brands from the individual characteristics (acuity, biases, and writing styles) embedded in these reviews. Instead of aggregating the product reviews into a brand-by-attribute table, as is typically done in the literature (Aggarwal, Vaidyanathan, & Venkatesh, 2009; Lee & Bradlow, 2011; Netzer, Feldman, Goldenberg, & Fresko, 2012), we look into each review directly as the basic unit for our analysis, first generating a lexical taxonomy for the product category using ontology

E-mail addresses: smoon13@uncc.edu (S. Moon), kamakura@rice.edu (W.A. Kamakura).

http://dx.doi.org/10.1016/j.ijresmar.2016.05.007
0167-8116/

learning-based text mining, and then analyzing the subsequent coding of each product review using psychometric mapping techniques.

Nowadays, consumers are being overwhelmed by product recommendations in the form of online reviews written by product experts (e.g., Wine Spectator for wines) or by sophisticated consumers taking the role of opinion leaders (e.g., "top contributor badge" on TripAdvisor). These online reviews, covering a broad range of products, are becoming more prevalent and influential with the recent explosion of social media (Kostyra, Reiner, Natter, & Klapper, 2016). The information contained in product reviews reveals idiosyncratic perceptions and preferences of individual reviewers, when the researcher reads "the voice of the consumer" by translating the reviews into a product positioning map portraying the relative brand locations in an attribute space (Lee & Bradlow, 2011; Netzer et al., 2012; Tirunillai & Tellis, 2014). Existing studies that generate product positioning maps from product reviews have overlooked or downplayed the critical role of such individual differences. To fill this gap, we propose a framework that produces a product positioning map directly from product reviews, revealing brand locations in a latent product attribute space while simultaneously adjusting for reviewer differences. Therefore, our framework helps managers obtain additional insights into what consumers tend to focus on when praising or complaining about their products, in such a way that corrects for perceptions and communication differences among consumers. For example, in the case of hotels, there are different traveler types (e.g., personal travelers vs. business travelers) and different online hotel booking sites (e.g., Expedia vs. TripAdvisor). Because our framework allows us to analyze reviews from specific types of travelers based on the combination of these two key hotel business aspects (e.g., Expedia business travelers and TripAdvisor personal travelers) separately from the other reviews, we can understand what hotel service features (e.g., room amenities and free eating) a targeted consumer segment (e.g., TripAdvisor pleasure travelers) notices and appreciates. In obtaining consumers' inputs to generate such positioning maps, marketing researchers in industry and academia alike have traditionally used the survey method, asking consumers to rate products on a pre-determined set of attributes (Carroll & Green, 1997). On the other hand, the recent emergence of the Internet has made social media and online product reviews popular as sources of brand information that are not affected by biases induced by researchers or managers when selecting product features to be evaluated. For example, Aggarwal et al. (2009) built an aggregate index from the co-occurrences of brand–attribute pairs on Google searches to produce unfolding maps of brands and attributes. These authors aggregated the extracted brand-feature data before producing a map, resulting in a loss of information about individual reviewers and reviewer–brand interactions.

Importantly, our proposed framework differs from recent text-based product positioning mapping approaches in two notable ways. First, we apply our proposed framework to experience products, wines and hotels, whereas most previous studies translating product reviews into product/brand maps have used search products such as digital cameras and mobile phones, where there is a more clearly defined and widely accepted terminology for product attributes. Even though most of the product positioning mapping approaches based on product reviews have focused on search products, product reviews are even more important for experience products because for these types of products, quality can only be assessed a posteriori after the consumption experience (Kamakura, Basuroy, & Boatwright, 2006). When there is a significant amount of individual differences in product attribute selections and evaluations, it is challenging to identify important product attributes a priori and to evaluate heterogeneous consumer preferences with a limited number of respondents (Moon, Park, & Kim, 2014). In particular, for sensorial products (e.g., wines, perfume, coffee, cigars, spas, and massages), which are unique types of experience products, quality assessment is highly subjective because attributes determining quality are highly nuanced and are therefore not easily determined (Peck & Childers, 2003). Due to this complex nature of sensory perception and the difficulty in communicating quality, consumers and marketers often rely on experts who are highly sensitive to and knowledgeable about flavors and aromas (Stone & Sidel, 2004). For example, wine quality grading requires the development of references for color, aroma, taste, mouth feel, and appearance by wine experts because most consumers may be unable to verbalize why they like or dislike the product and tend to rely on reviews written by sophisticated consumers or experts.

Second, rather than relying on aggregate brand-by-attribute counts, we use the raw text from each product review as the unit of analysis to develop a product-specific lexicon extracted from the original language used by the reviewers and generate a review-based product positioning map that can locate individual reviewers. In this procedure, by using individual reviews as the unit of analysis, we account for the fact that each review reflects both the product and the reviewer, which allows us to isolate language differences across reviewers from perceptual differences across products. This method also generates both product and reviewer positions on the same map, which allows us to capture the relationships of products and reviewers effectively. Rather than specifying a priori the topics and attributes to be used, we utilize the verbal content generated by consumers or experts to define the topics and attributes determining the positioning map a posteriori via ontology learning (Cimiano, 2006; Feldman & Sanger, 2007). Ontology is an explicit, formal specification of a shared conceptualization within a domain of interest, where "formal" suggests that the ontology should be machine-readable and shared with the general domain community (Buitelaar, Cimiano, & Magnini, 2005). That is, ontology refers to all of the core concepts (including their terms, attributes, values, and relationships) that belong to a specific knowledge domain, which is specifically the wine and hotel businesses in our two empirical applications. Automatic support in ontology development is called ontology learning. Historically, ontology learning is connected to the semantic web, which builds on logic formalism restricted to decidable fragments of description logics (Staab & Studer, 2004). Thus, the domain models to be learned are restricted in their complexity and expressivity. Ontology learning is conducted using input data from which to learn the concepts relevant to the given domain, their definitions as well as the relations connecting them (Cimiano, 2006). In doing so, we maximize the number and relevance of identified product features from the reviews, in spite of the data sparseness problem stemming from individual reviewers' usage of a different set of self-selected features (Lawrence, Almasi, & Rushmeier, 1999).

From a managerial perspective, although marketing practitioners are not interested in individual consumers, they develop specific marketing strategies toward distinct consumer segments. For example, hotel managers utilize different strategies in serving business travelers and family travelers in response to their differential preferences. Family travelers tend to spend more leisure time in and around the hotel, whereas business travelers will be tied up with their work. Therefore, reviews from different groups of consumers tend to be noticeably different (e.g., “main tourist area” from personal travelers vs. “conference room” from business travelers). Similarly, travel websites tend to attract different types of customers due to their own characteristics. For example, TripAdvisor, a travel guide website, contains rich travel information from travelers highlighting travel attractions, whereas other hotel reservation websites are more business-driven. Our approach emphasizing reviewers illuminates the differences between travel websites and provides useful managerial insights. Our approach of highlighting such distinct reviewer groups (i.e., meta-reviewers in terms of traveler type and review source) on the product positioning map can help managers obtain additional insights into how different groups of consumers tend to focus on different attributes in their product reviews.

Because we propose a framework for product positioning mapping using online product reviews, we provide a literature review focusing on this specific subject in the next section. After this review, we develop our proposed framework combining ontology learning with a novel product–reviewer dual space-reduction model for wine reviews by wine experts. We then extend this framework to hotel reviews from a large number of ordinary consumers, thereby demonstrating how our procedure can be applied to reviews by both experts and average consumers, irrespective of the number of reviews from individual reviewers.

2. Prior studies eliciting product features from online product reviews

Marketing scholars have actively examined various aspects of online product reviews, primarily focusing on their numeric summary rather than their semantic text; for the most part, they have ignored a wealth of textual information contained in these reviews, despite evidence that consumers read text reviews rather than relying solely on summary numeric ratings (Chevalier & Mayzlin, 2006). Table 1 summarizes studies that explored textual product reviews to better understand consumer perceptions and preferences for competing products. These studies can be evaluated on three major criteria: (1) how product features are elicited (text input); (2) how reviewers and reviews are analyzed in association with the elicited product features (text analysis); and (3) what outputs are generated from the combination of the text input and text analysis used (text output).

Focusing on the first criterion (text input), we find different product feature elicitation approaches in the literature. Aggarwal et al. (2009) identified only a single attribute descriptor from each brand's positioning statement, and they combined all of the identified descriptors from all of the brands (e.g., Gentle from Ivory Snow, Tough from Tide) to determine a set of product attributes for the product category (P&G detergent brands) prior to their lexical text analysis. Because a single representative attribute is not enough to cover the broad spectrum of features in each product, other studies elicited multiple features for each product. Some of these studies focused on a limited number of features because of data sparsity (Archak, Ghose, & Ipeirotis, 2011) or analysis efficiency (Ghose, Ipeirotis, & Li, 2012). Other studies intended to be comprehensive in capturing product features (Decker & Trusov, 2010; Lee & Bradlow, 2011), but were based on aggregate product “pros and cons” counts by brand instead of full product reviews. To make the most of common product reviews, our proposed study takes the full text in these reviews to identify and taxonomize the product features (Netzer et al., 2012; Tirunillai & Tellis, 2014). We also apply a multiple-layer hierarchy of the features with

Table 1
Comparison of studies eliciting product features from online product reviews.

Study	Product feature elicitation	Reviewer–review analysis	Product perception map	Techniques	Analyzed products
Aggarwal et al. (2009)	Single attribute from product positioning statement	Reviews aggregated at the brand level	Yes, but product locations only	Lexical text analysis, computational linguistics	Detergent
Archak et al. (2011)	Only the most popular 20 features from reviews to alleviate data sparsity	Decompose reviews into segments describing different product features	No	Sentiment analysis, consumer choice model	Digital cameras and camcorders
Decker and Trusov (2010)	Multi-features from product pro/con summary data	Review clustering, but not reviewer analysis	No	Natural language processing, latent class Poisson regression	Mobile phones
Ghose et al. (2012)	Only the most popular 5 features from reviews	Estimate reviewer-specific random coefficients	No	Sentiment analysis, random coefficient hybrid structural model	Hotel reservations
Lee and Bradlow (2011)	Hierarchical multi-features elicited from product pro/con summary phrases	Reviews aggregated at the brand level	Yes, but product locations only	*K*-means text clustering, correspondence analysis	Digital cameras
Netzer et al. (2012)	Co-occurrence terms including product features with brands identified	Reviews aggregated at the brand level	Yes, but product locations only	Rule-based text mining, semantic network analysis	Sedan cars and diabetes drugs
Tirunillai and Tellis (2014)	Pre-processed reviews and seed words for valence	Reviews aggregated at the brand level	Yes, but product locations only	Part-of-speech tagging, latent Dirichlet allocation	Mobile phones, computers, toys, footwear, and data storage
This Study	Complete hierarchical multi-features elicited from expert reviews	Each reviewer's idiosyncratic linguistic propensities considered	Yes, with product, reviewer, and product attributes	Ontology learning, dual reduced-spaces	Wines and hotels

descriptive terms nested into grand topics (e.g., "vegetative" in the wine category) and topics (e.g., "fresh herb," "canned vegetable," and "dried herb" nested under "vegetative"), by applying ontology learning-based text mining (Feldman & Sanger, 2007).

Regarding the second and third criteria (text analysis and text output), past studies have aggregated product features extracted from product reviews at the brand level (Aggarwal et al., 2009; Lee & Bradlow, 2011; Netzer et al., 2012; Tirunillai & Tellis, 2014), ignoring individual reviewers' idiosyncratic focus on different product features and distinctive writing styles (text analysis). Capturing individual reviewers' use of particular product attributes and unique vocabularies is even more important for experience products, given that evaluating these products is more subjective and, accordingly, tends to draw more heterogeneity among consumers (Liebermann & Flint-Goor, 1996). Furthermore, individual reviewers tend to write about only a limited set of product features, which leads to data sparseness in capturing individual reviewers' propensities. To overcome the combination of data sparseness and reviewer heterogeneity inherent in experience products, we use ontology learning to produce an enriched list of product features and detailed terms describing each revealed feature directly from consumer reviews.

In particular, our first empirical application takes wines as an exemplary *sensorial experience product* category, where the consumption experience is not only subjective, but also delicately nuanced. For these reasons, consumers rely on experts to verbalize their deep and rich consumption experiences. In our second application, we extend our framework to another experience product category: hotel services. Without the sensorial nature of wines, consumers can verbalize a wide range of experiential features of hotel service consumption (e.g., hotel atmospheres, room amenities, staff services) without experts' assistance. Therefore, compared to wines, where one finds a limited number of reviews full of complex, technical, and nuanced terms (such as medium body, pungent, tannin, from a small number of wine experts), hotels generate a much larger number of reviews consisting of easy-to-understand terms (e.g., awesome room, complementary Wi-Fi, friendly doorman) from a large number of consumers. We demonstrate that our approach is fundamental and flexible enough to cover a variety of products ranging from wines to hotels. Specifically, as long as there are at least two reviews per reviewer, our approach can locate reviewers, products, and product attributes on the same map, irrespective of the product category.

To conclude, our research combines four desirable properties for translating product reviews into a product positioning map, which distinguishes our work from previous research. First, we apply ontology learning directly to raw full online product reviews (as opposed to aggregate summary texts) to generate enriched product descriptor terms, which results in a more extensive and deeper domain taxonomy for the product positioning map. Second, we use each product review as our basic unit of analysis, directly considering the dyadic relationship between each product and each reviewer. Third, we generate a multi-faceted product positioning map that illustrates the relationships among products, reviewers, and product features. It should be noted that our study is the first to consider review positions in this task of translating product reviews into a product positioning map.

3. Developing a framework to translate wine reviews into a product positioning map

The framework we propose to translate product reviews into a perceptual map involves two distinct processes. In the first process, we mine the text contained in product reviews to tokenize the text and elicit relevant part-of-speech (POS) terms to categorize these elements into a lexicon of meaningful product-specific topics and associated descriptive terms, which are utilized to code all of the reviews. In the second process, we analyze the topics and terms utilized by each reviewer in describing each product to produce two reduced spaces, one representing the position of the products and topics in a common space, and the other representing the "dialects" distinguishing the writing style and perceptual acuity of each reviewer. Because the proposed framework involves two distinct processes, we describe both processes successively in the context of one concrete application to the wine product category instead of covering each process separately in the model development and empirical application. Then, in the next section, we extend our framework to another category, hotel services, to show the wide applicability of the framework in terms of the nature and structure of the review text.

The first application we choose to describe and illustrate our proposed framework for both text mining and product positioning involves wine reviews written by seven best-known wine experts. We choose wine reviews because of the complexity of the aromas and flavors involved, which most consumers can perceive (and use these perceptions to form their preferences), but are often unable to verbalize precisely. This aspect makes experts' opinions important and valuable to both consumers and marketers. While the focus of our first illustration is on the seven experts' reviews on red wines as a representative *sensorial* experience product, the applicability of the proposed framework in other contexts should be clear, as long as consumers review competing products and services, as we will demonstrate in our second application using hotel reviews in the next section.

The 640 online wine reviews we use in our illustration were published by seven wine experts evaluating 249 different red wines from North America, Australia, Europe, Africa, and South America. The prices for the reviewed wines ranged from $7.30 to $75.90 (mean = $23.40; standard deviation = $13.80). Each review contains one or more paragraphs of text (mean = 306 characters; standard deviation = 168 characters) describing the wine in freewriting style, along with a numeric overall quality rating that varies in range from expert to expert. These quality ratings are used to relate the perceptual map to expert preferences. These reviews and quality ratings were obtained from public sources such as WineEnthusiast.com, as well as syndicated services such as JancisRobinson.com and winecompanion.com.au. While the amount of text collected across the 640 reviews seems substantial, in reality, the data are very sparse in covering a variety of product features for each of the 249 red wines, with an average of 2.7 reviews per wine and between 54 and 181 reviews written by each critic. This data sparsity problem is common on popular websites reviewing various product categories. As we will see later, our dual space reduction method resolves this data sparsity problem by producing a more parsimonious fit to the data, through ontology learning that generates a rich domain taxonomy.

3.1. Developing a product-specific taxonomy from product reviews by ontology learning

The first process in our proposed framework is to develop a lexical taxonomy of domain topics extracted from the wine reviews. We use text mining to elicit structured taxonomical information from unstructured textual product reviews. Product reviews in different product categories, such as hotels, restaurants, home repair contractors, and real estate agents, can lead to category-specific feature taxonomies with their own unique consumer vocabularies. When basic source information comes directly from experts, such a product taxonomy may present a richer product usage vocabulary, which can be effective in explaining a variety of consumers' delicate perceptions and subtle preferences (Alba & Hutchinson, 1987). This would be particularly relevant for sensorial products, where the typical consumer may not be able to verbalize his or her perceptions accurately. In the wine category, ordinary consumers may enjoy different wine flavors and aromas, but may not be able to effectively communicate the delicate and complex tastes they experience. Furthermore, ordinary consumers often seek recommendations from professional experts, and even expert consumers to find good-quality wines that would not spoil their consumption experiences (Camacho, De Jong, & Stremersch, 2014). As our initial "dictionary," we begin with the Wine Aroma Wheel, a wine attribute taxonomy developed by a wine expert, Ann Noble (winearomawheel.com); such an initial dictionary is not required in our framework, as shown in our second application to develop a hotel taxonomy. The Wine Aroma Wheel summarizes the range of flavors and aromas commonly found in wines. In fact, the Wheel is a three-level hierarchical taxonomy of wine attributes composed of 12 grand topics (e.g., pungent, vegetative, fruity), 29 topics (e.g., cool, dried, citrus), and 94 terms (e.g., menthol, tobacco, lemon) that serve as descriptors for the 29 topics. Because this original Wheel may not fully reflect what other experts perceive and does not include other topics beyond flavor and aroma (such as body and depth), we use ontology learning iteratively to expand, deepen, and revise this original taxonomy and our own following taxonomy until there is little room for improvement. In other words, we expand and enrich relevant wine terms (e.g., licorice, vanillin, medium finish) to establish a hierarchical taxonomy composed of the three layers of grand topic, topic, and terms, as summarized in Table 2. Fig. A1 in the web appendices demonstrates the validity of this approach by comparing the domain taxonomy from ontology learning and the original taxonomy provided by Ann Noble, given that there are no clear objective measures to compare the validity of multiple ontologies (Maedche, 2002). In summary, whereas the original taxonomy generated an average of 2.89 counts of relevant terms per review, our ontology learning approach

Table 2
Summary of the modified wine taxonomy.

Grand topic	Topic	Representative descriptive terms (the total number of terms in each topic)
Floral	Floral	Blossom, floral, floral-scented, geranium, violet (15)
Spicy	Clove	Cinnamon, clove, cola, nutmeg (4)
	Pepper	Pepper, peppery, spice, spicy (23)
	Licorice, anise	Anise, licorice, liquorice (8)
Fruity	Citrus	Citrus, citrusy, orange (3)
	Berry	Berry, black fruit, cassis, currant, mulberry, red-fruit (50)
	Tree fruit	Apple, apricot, cherry, kirsch, peach, yellow fruit (8)
	Dried fruit	Compote, cooked fruit, dried fruit, fig, fruitcake, jam (27)
	Other fruit	Fruit, fruity, juicy, yellow plum (16)
Vegetative	Fresh herb	Camphor, eucalyptus, grass, mint, minty, peppermint (7)
	Canned vegetable	Asparagus, beet, cooked vegetables, olive-accent (9)
	Dried herb	Briar, dried vegetative, fennel, leafiness, tealike (27)
Nutty	Nutty	Chestnut, nut, nutty, walnut (4)
Caramelized	Chocolate	Bean, chocolate, coffee, espresso, hoisin sauce (12)
Woody	Vanilla	Phenolic, vanilla, vanillin (5)
	Woody	Balsam, cedar, oak, oaky, resinous, sandalwood, woody (19)
	Smoked	Bacon, burned, charcoal, grilled, pain grille, toasty (31)
Earthy	Earthy	Chalky, dusty, earthiness, garrigue, loam, meaty, soil (24)
	Mineral	Graphite, lead, mineral, mineral-accent, minerality (18)
Chemical	Petroleum	Diesel, kerosene, leather, medicinal, tar, tar-scented (15)
	Sulfur	Cabbage, garlic, rubbery, skunk, wet dog, wet wool (6)
	Pungent	Acid, ethanol, pungent, vinegar (4)
	Metallic	Iron, metal, metallic (3)
	Other chemical	Fishy, fusel alcohol, soapy (3)
Alcohol	Alcohol	Alcohol, cool, menthol (5)
Microbiological	Lactic	Butyric acid, cheese, sauerkraut, sweaty (4)
	Other microbiological	Horsey, lees, leesy, mousey, musky, truffle (6)
Tannin	Tannin	Astringent, tannic, tannin (27)
General taste	Sweet	Bittersweet, candied, cane, pudding, sugar, sweet (18)
	Dry	Dried, dry, low sugar, no residual sugar (8)
	Crisp	Crisp, fresh acidity, moderate acid, taut (9)
Overall	Polished	Polished, soft, softness (8)
	Complex	Complex, complexity, elegant, layered (12)
	Balanced	Balanced, round, satiny, sleek, well-balanced (10)
	Medium body	Medium bodied, medium finish, medium framework (6)
	Full body	Concentrated, dense, full of flavor, full-bodied (22)

captured an average of 7.81 counts of relevant terms per review. In other words, compared to the original taxonomy, our approach generated 2.7 times the terms to be analyzed in perceptual mapping generation. This reduces the data sparsity problem inherent in the review-term frequency matrix. Concurrently, our ontology learning identified 7 new topics that the original taxonomy had missed. This is seen as a significant improvement because these new topics add significant insights into the complex relationships within the wine domain when the topics are visibly located in the following positioning map. To conclude, our ontology learning is used to reach a more complete and comprehensive domain taxonomy through enriched terms and expanded topics (Pazienza, Pennacchiotti, & Zanzotto, 2005).

Our ontology learning process is composed of five distinctive steps: (1) online review collection; (2) text parsing and filtering (to generate a relevant term list); (3) taxonomy building (to determine the grand topics, topics, and terms, as shown in Table 2); (4) sentiment analysis (to determine the valence of each topic and its associated terms into positive, negative, or neutral); and (5) text categorization (to count the frequency of each topic in each review). We describe each of the five steps briefly here, and provide more details in Web Appendix A.

In the first step (online review collection), either a manual or automatic approach can be used. For the wine review, we managed to collect 640 wine reviews manually. To collect a large number of reviews, which is typical in online product reviews, it is necessary to use web crawling techniques (Lau, Lee, Lam, & Ho, 2001). We apply this automated approach to our second application to extract close to half a million hotel reviews.

In the second step (text parsing and filtering), we process unstructured textual information into a structured list of terms that can be used as the foundation for taxonomy building in the following step. In doing so, we use various techniques for text parsing and text filtering. We use text parsing for information extraction from unstructured textual reviews. In our text parsing tasks, we use part-of-speech (POS) tagging to determine the word class of each term (e.g., adjective, noun) based on natural language processing. In this processing, we consider multiple-word terms that generate specific meanings relevant to the domain (e.g., medium bodied, full of flavor, low sugar). Using the uncovered word groups, we consider word classes that produce meaningful terms only, such as adjective/noun and adverb/verb pairs. Furthermore, we use a stemming function to group multiple words of the same root (e.g., clean, cleaner, cleaned, cleaners, cleans, cleaning, cleanest) into a single term (e.g., clean) because of their common meaning. On the other hand, text filtering allows us to select relevant terms by removing basic and common terms irrelevant to our domain (e.g., during, go, the) and by filtering out infrequent terms to raise the efficiency of our process. In particular, by removing such infrequent terms from a long list of terms, we can dramatically reduce our manual effort in determining the relevance of each term to the domain of interest. Web Appendix A provides details on this step.

In the third step (taxonomy building), we build a hierarchical domain-specific taxonomy comprising three layers of grand topics, topics, and descriptive terms, as summarized in Table 2. To optimize the taxonomy, we rely on three sources of knowledge: (1) our domain knowledge and learning based on the term list from text parsing and filtering in the previous step; (2) vocabularies and information adopted by the domain (e.g., Noble's original wine wheel); and (3) concept linking showing close terms' associations. In concept linking, we can use highly correlated terms to determine their topic associations (e.g., fruitcake and jam pertaining to the same topic of dried fruit).

In the fourth step (sentiment analysis), we determine the valence of each topic identified by taxonomy building via sentiment analysis (Pang & Lee, 2008). However, in the case of wine, it is neither straightforward nor clear how to determine the valence of individual topics such as clove, woody, earthy, and lactic, partially because wine experts use flowers, fruits, and other food-related terms as metaphors for flavors and aromas. To evaluate the valence of each topic more objectively instead of using our own subjective judgments, we use the overall quality rating and determine the valence of each topic as part of our model estimation below. In other words, we do not conduct sentiment analysis in the wine example a priori, because the valences of most wine topics cannot be determined before we actually estimate the impact of each topic on the overall wine rating. For the hotel reviews, where we can determine the valence of each topic more objectively, we use a more explicit sentiment analysis method, as explained in our hotel application later.

In the fifth and final step (text categorization), we count the frequency of each topic in each review. We use this review-by-topic frequency matrix (where a cell indicates the appearance frequency of relevant terms pertaining to the corresponding topic in the corresponding review) for our space-reduction mapping model in the next process below. To obtain the matrix, we use the text categorization method (Feldman & Sanger, 2007), which tags each review with a number of specified descriptive terms based on the topic definition and the term usage of the review (as in Table 2). We use this tagging information from text categorization to generate the matrix. After collecting the online product reviews we need, we repeat Steps 2 through 5 until we generate a stable and comprehensive taxonomy for the target product category.

The ontology learning-based text mining framework used in this research led to the coding of the 640 wine reviews into a wine taxonomy of 16 grand topics, 36 topics, and 476 terms. We removed three of the elicited 36 topics in the next quantitative mapping process of our proposed framework because the three removed topics (i.e., sulfur, other chemical, and lactic) were used very rarely (less than 0.5% of all reviews) when the seven experts described red wines. Importantly, these seven experts differ substantially regarding the language they used in their reviews. For the purpose of illustration, Fig. 1 compares two select critics' 33 topic usage proportions based on their aggregated reviews coded through our ontology learning process in the form of a Wordle chart (www.wordle.net). Wordle is a popular word cloud website. The word cloud is primarily used to visualize document content by a set of representative keywords. We use this popular visualization method to succinctly summarize wine reviews typically full of noise by irrelevant words (e.g., come, drink, though). In the chart, the size of each word is in proportion to its total frequency in all of the reviews used. However, the reader must beware that the relative positions of the words have no meanings. We use these word clouds only as a preliminary way to highlight different writing styles among wine reviewers.

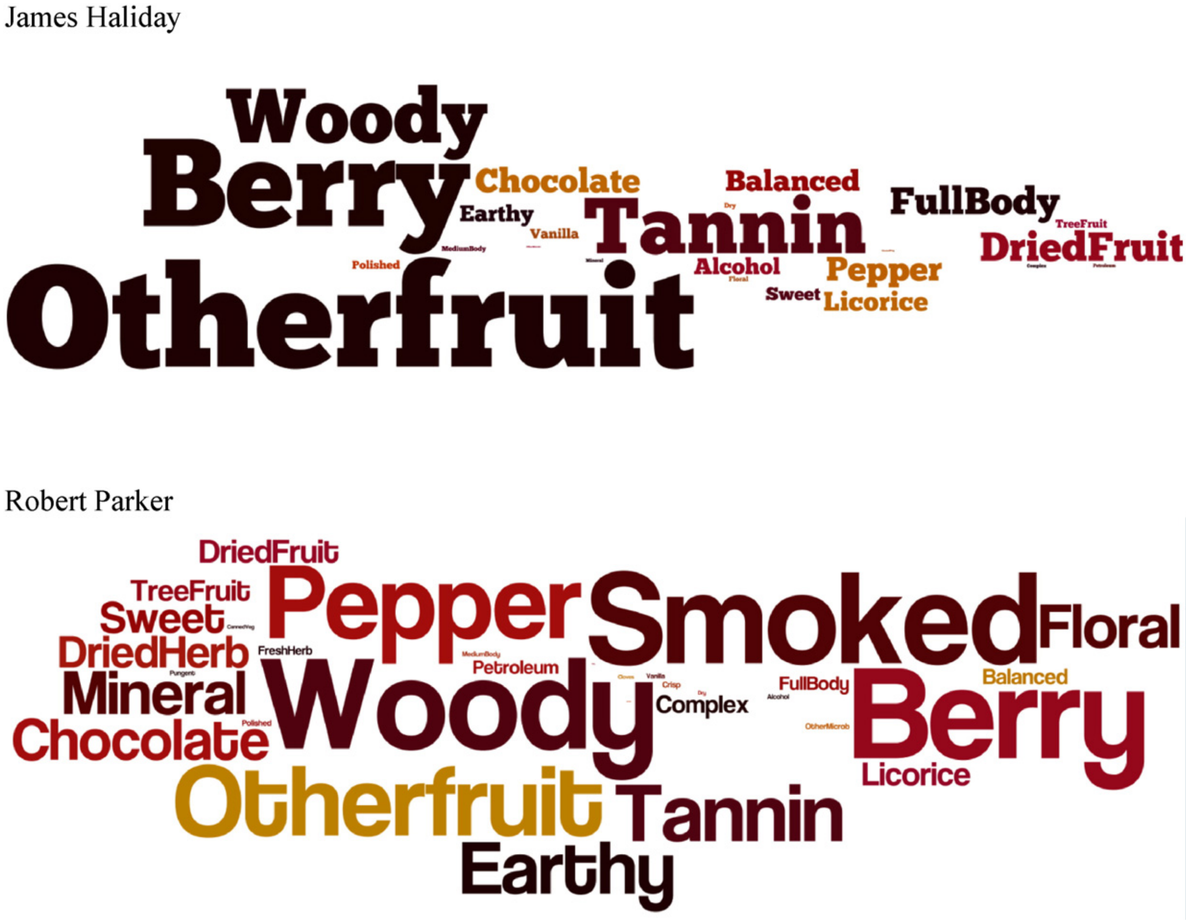

Fig. 1. Word clouds of topic counts in two wine experts' reviews.

Specifically, James Haliday's reviews (the top word cloud) appear dominated by the topics of "otherfruit," "berry," "tannin," and "woody," while Robert Parker's reviews (the bottom word cloud) have a more even distribution of topics. What is still unknown from these Wordle charts is whether these differences are due to the wines reviewed by these two experts or due to their idiosyncratic wine-related language. These two distinct components are parsed out in the model we present next.

3.2. Representing product reviews in a product positioning map

In our product positioning mapping model, we acknowledge that usage of a certain term by a reviewer to describe a certain product is a reflection of both the reviewer and the product, as suggested by Fig. 1. In other words, the topics and terms used by the seven wine critics to describe the 249 red wines contain information about the characteristics of the idiosyncratic language used by the individual wine critics as well as the wines. While the seven critics use the same lexicon of the 33 topics identified by our ontology learning, we allow individual critics to have their own "dialect." Then, the problem is to parse out the critics' propensity to use certain topics to describe red wines and the perceived intrinsic characteristics of the wines from the available reviews. To parse out the critic and wine characteristics from all of the topics observed across all reviews, we represent the propensity for wine critic i to use topic k in reviewing wine j by

$$u_{ijk} = u_{ik} + u_{jk} + \varepsilon_{ijk}, \quad (1)$$

where:

u_{ik} general propensity of reviewer (critic) i using topic k to describe any brand (wine),
u_{jk} general propensity for topic k to be used by any reviewer (critic) in describing brand (wine) j.
ε_{ijk} i.i.d. random error.

The simple formulation in Eq. (1) allows us to extract the intrinsic perceptions of the 249 wines along the 33 topics (u_{jk}), after correcting for the general propensity of each critic to use each topic when describing a specific red wine (u_{ik}). However, this simple formulation is intractable because it would involve the estimation of $33 \times (7 + 249) = 8448$ parameters. To make it more complex, as the number of reviewers increases, the number of parameters will increase by 33 with each additional reviewer. To obtain a more parsimonious and more managerially meaningful formulation in response to this issue, we rely on space reduction using two latent spaces: the reviewer–topic and brand–topic dyads. Accordingly, we replace the formulation in Eq. (1) by

$$u_{ijk} = u_k + \Gamma_k Z_i + \Lambda_k W_j + \varepsilon_{ijk}, \tag{2}$$

where:

u_k intercept representing the general tendency of any reviewer (critic) using topic k to describe any brand (wine), to be estimated.
Γ_k p-dimensional (p to be determined empirically) vector of factor loadings for topic k, to be estimated,
Z_i p-dimensional (q to be determined empirically) vector of independent, standard normal latent scores for reviewer (critic) i.
Λ_k q-dimensional vector of factor loadings for topic k, to be estimated,
W_j q-dimensional vector of independent, standard normal latent scores for brand j.

3.3. Implications of the proposed space-reduction model

The p-dimensional reduced space defined by (Γ_k,Z_i) provides valuable insights about the reviewers' linguistic idiosyncrasies (toward sensorial wine features), while adjusting for individual differences. This a posteriori adjustment for individual linguistic differences is akin to the a priori ipsatization of survey data to remove individual response biases. On the other hand, the q-dimensional reduced space defined by (Λ_k, W_j) is the perceptual map for all wines across all topics, adjusted for perceptual and language differences across reviewers (critics). This provides the analyst with a more parsimonious and meaningful depiction of the brands (wines) in terms of their intrinsic sensometric characteristics. We assume that the two reduced spaces are independent of each other because they represent two distinct constructs (language differences and wine perceptions). As most space-reduction models, there is invariance to rotation and, therefore, the usual constraints on the loadings (Γ_k and Λ_k) are imposed for identification (Kamakura & Wedel, 2000).

3.4. Error specification

The specification of the error term in Eq. (2) depends on the nature of the topic coding obtained from the ontology learning section of our framework. Our original intention was to count the number of times each topic was mentioned in each review, which would lead to a Poisson model for these counts. Therefore, we planned to use Eq. (2) as the log-hazard, assuming an exponential distribution for the error term. However, a detailed inspection of the data revealed that they were too sparse, where most topics were mentioned at most once or twice in a review. Across all topics, the distribution of counts was "0" = 82%, "1" = 15%, and "2 or more" = 3%, indicating that binary coding would be more appropriate than the Poisson model for the data we obtained from text mining. Accordingly, we assume an extreme value distribution for the error term in Eq. (2), leading to a multivariate binary logit model for the 33 topic codes as follows (an extension to Poisson counts or any other data format is fairly straightforward, requiring a different link function, as shown by Wedel and Kamakura (2001)).

$$g\left(y_{ijk} = 1 \middle| \Gamma_k, Z_i, \Lambda_k, W_j\right) = \left(\frac{\exp\left(u_k + \Gamma_k Z_i + \Lambda_k W_j\right)}{1 + \exp\left(u_k + \Gamma_k Z_i + \Lambda_k W_j\right)}\right). \tag{3}$$

Aside from the topics coded via ontology learning, our data also include wine quality ratings associated with each wine review, which we include in the positioning map to reflect the preferences expressed by the reviewers (wine critics). However, because each of these reviewers (wine critics) rates self-selected brands (wines) on an independent scale reflecting his/her own rating leniency/strictness, we estimate individual-specific parameters for each critic and his/her own ratings instead of relying on the general p-dimensional reduced space. Let q_{ij} represent the rating by reviewer (critic) i to brand (wine) j; we model the likelihood for this observation as

$$f\left(q_{ij} \middle| u_i, \sigma_i, \Lambda_i, W_j\right) = \phi^*\left\{\frac{q_{ij} - u_i - \Lambda_i W_j}{\sigma_i}\right\}, \tag{4}$$

where

$\phi^*\{\cdot\}$ standard normal probability density function,
u_i intercept representing the mean ratings for reviewer (critic) i,

σ_i standard deviation for reviewer (critic) i's ratings,
Λ_i q-dimensional vector of factor loadings for reviewer (critic) i,
W_j q-dimensional vector of independent, standard normal latent scores for brand j.

The space-reduction model described in Eqs. (2) through (4) is similar in form and purpose to the large family of factor-analytic (Wedel & Kamakura, 2001) models widely used by marketing researchers for perception and preference mapping. Similar to these models, our objective is to represent brands, attributes, and preferences in a joint reduced-space. However, our model produces two distinct spaces: (1) a space positioning brands (wines) in relation to reviewers' perceptions and preferences for topics, and (2) another reduced-space capturing the propensities by each reviewer in using each topic in his/her reviews, irrespective of the brand reviewed, thereby isolating brand perceptions from linguistic or acuity differences across reviewers.

3.5. Model estimation

Combining the multivariate binary logit model for the coded wine reviews in Eq. (3) with the likelihood in Eq. (4) for quality ratings, we obtain the following log-likelihood function for our model,

$$L\{Y,Q|M,\Gamma,\Lambda,\Sigma\} = \prod_i \int \prod_j \int \left\{ \prod_k \left| \begin{array}{c} f\left(q_{ij}|u_i,\sigma_i,\Lambda_i,W_j\right)* \\ g\left(y_{ijk}=1|\Gamma_k,Z_i,\Lambda_k,W_j\right)^{y_{ijk}}* \\ g\left(y_{ijk}=0|\Gamma_k,Z_i,\Lambda_k,W_j\right)^{1-y_{ijk}} \end{array} \right| \right\} \phi^*\left\{W_j\right\}dW_j\phi^*\{Z_i\}dZ_i \tag{5}$$

Estimation of the model parameters is implemented via the simulated maximum likelihood (Wedel & Kamakura, 2001). More details about our expectation–maximization (E-M) estimation method of this particular model and the computation of factor scores on the two (wine and critic) latent spaces are provided in Web Appendix B. In brief, the estimation process starts by randomly generating initial values for the $\{M,\Gamma,\Lambda,\Sigma\}$ parameters. Then, the estimation step computes the scores W_j for each brand and Z_i for each reviewer based on all of the available data, based on the assumption that the current parameter estimates $\{M,\Gamma,\Lambda,\Sigma\}$ are given. Then, new estimates of the $\{M,\Gamma,\Lambda,\Sigma\}$ parameters are obtained via maximum-likelihood in the maximization step, based on the assumption that the current factor scores are known data. These two (expectation and maximization) steps are iterated until convergence.

4. Empirical results on the wine positioning map

In order to determine the dimensions of the critic–topic and wine–topic spaces, we estimated the model by varying the critic–topic space from $p = 1$ to $p = 3$ dimensions and the wine–topic space from $q = 1$ to $q = 4$ dimensions. Based on the popular Bayesian information criterion (Schwarz, 1978), we selected a 2-dimensional space to represent the critics' general propensity to use the 33 topics in their reviews and a 3-dimensional space to represent their wine perceptions. The parameter estimates obtained for the selected model (the 2-dimensional representation of reviewers' language and the 3-dimensional positioning map) are reported in Table 3, where the estimates significant at the 0.05 level are highlighted in bold. Because we applied the model to binary coded reviews in terms of the 33 topics, assuming an extreme value error distribution, Eq. (2) produces the log-odds for topic k being used at least once by critic i to describe wine j. Accordingly, the term $\Gamma_k Z_i$ in Eq. (2) represents the shift in the log-odds due to critic i's general propensity to use topic k to describe any wine relative to other critics. This propensity space indicates how the reviews have been adjusted to account for idiosyncratic language and perceptual acuity for the purpose of producing a positioning map that better reflects the critics' perceptions and preferences for the 249 red wines. This reduced space is utilized as a device for capturing the idiosyncratic tendencies to use certain topics when describing wines, in a relatively parsimonious fashion, with 65 parameters rather than 231 required to capture all individual effects. As shown in Table 3, we observe substantial differences in the dual spaces across the critics and topics. In other words, the maps to be created by the results in the table would reveal topics most commonly used by individual critics, reflecting the "dialects" utilized by each critic.

Furthermore, these linguistic differences across the seven critics are summarized in Table 4, which shows the propensities measured in percentage points $\left(\frac{\exp(u_k+\Gamma_k Z_i)}{1+\exp(u_k+\Gamma_k Z_i)}\right)$. These propensities show that "berry" (e.g., black fruit, cassis, currant, and mulberry), "otherfruit" (e.g., fruity, juicy, and yellow plum), "woody" (e.g., balsam, cedar, oak, oaky, and sandalwood) and "tannin" (e.g., astringent, tannic, and tannin) are the most common topics found in the wine reviews, although there are clear differences in their usage by the seven critics. For example, "tannin" is the second most common topic, used by 44% of the reviews across the seven critics after "berry" (83%). However, Robinson's propensity to cover the topic is 62%, almost twice that of Haliday's (33%) or Parker's propensity (33%) to mention the same topic. Previously (Fig. 3), we compared James Haliday and Robert Parker on their usage of topics across all of their reviews. Using Table 4, we can now see how they differ by their propensity to use some other topics in their reviews, irrespective of the wines reviewed, because the wine effect has already been accounted for in the wine space. Specifically, Parker is more likely to use "floral," "pepper," "sweet," "mineral," and "fullbody" in his reviews than Haliday, while Haliday is more likely to mention "driedfruit" and "chocolate" than Parker. This indicates that differences in topic usage across

Table 3
Parameter estimates of the proposed model.

	Perceptual/preference space			Critic propensity space		
	β_i, Λ_k			Γ_k		μ_i, μ_k
Item	W1	W2	W3	Z1	Z2	Intercept
Haliday	**1.04**	**0.92**	**0.86**	0.00	0.00	**92.14**
Robinson	0.26	0.05	**0.40**	0.00	0.00	**16.41**
Parker	**1.12**	**0.82**	0.15	0.00	0.00	**91.46**
WRO	**0.90**	0.15	0.28	0.00	0.00	**89.73**
WS	**1.22**	**0.90**	0.29	0.00	0.00	**89.34**
Tanzer	**0.71**	**0.54**	−0.11	0.00	0.00	**90.95**
Enthusiast	**0.59**	−0.03	0.03	0.00	0.00	**91.05**
Floral	0.19	**0.81**	**0.62**	**0.94**	**1.11**	**−2.50**
Cloves	**1.46**	−0.21	0.18	**0.64**	−0.01	**−4.66**
Pepper	**−0.85**	**1.76**	**1.42**	**0.68**	**0.59**	**−1.13**
Licorice	**0.34**	−0.19	**0.33**	**0.54**	**0.33**	**−1.82**
Citrus	**−1.27**	0.76	**−1.46**	0.59	0.17	**−6.43**
Berry	**0.45**	0.17	**0.72**	0.14	**0.62**	**1.55**
TreeFruit	**−0.22**	**0.26**	−0.12	**0.33**	0.03	**−1.52**
DriedFruit	0.16	**0.53**	−0.17	0.17	−0.16	**−1.24**
Otherfruit	−0.12	0.05	**−0.17**	**−0.27**	−0.11	**−0.29**
FreshHerb	0.36	**−0.59**	**0.66**	0.32	0.06	**−3.28**
CannedVeg	**−0.98**	0.00	−0.07	−0.04	−0.21	**−3.88**
DriedHerb	−0.08	**−0.46**	**0.66**	**0.28**	−0.11	**−1.42**
Nutty	0.21	0.27	−0.86	0.93	0.49	**−5.48**
Chocolate	**0.55**	**0.84**	**−0.43**	−0.11	−0.16	**−1.35**
Vanilla	**1.09**	**0.47**	0.05	0.23	−0.16	**−3.25**
Woody	**0.39**	**−0.22**	**0.23**	**−0.24**	0.12	**−0.68**
Smoked	**0.21**	**0.41**	**0.31**	0.01	**0.50**	**−1.32**
Earthy	0.08	−0.12	−0.05	0.15	**0.22**	**−1.38**
Mineral	**0.40**	−0.09	**−0.30**	**0.51**	**0.50**	**−1.71**
Petroleum	0.06	0.03	**−0.46**	0.27	0.21	**−2.68**
Pungent	0.62	**−1.30**	0.27	0.70	**0.77**	**−5.41**
Metallic	−0.20	0.34	−0.40	0.17	−0.43	**−4.30**
Alcohol	−0.01	−0.01	0.08	**0.36**	−0.01	**−2.59**
OtherMicrob	0.04	−0.48	0.45	**0.72**	**0.82**	**−4.61**
Tannin	**0.31**	**−0.44**	**0.64**	**0.41**	−0.10	**−0.23**
Sweet	**0.35**	0.06	0.22	**1.12**	**0.56**	**−1.69**
Dry	**0.36**	−0.18	**1.07**	**0.48**	**−0.45**	**−2.52**
Crisp	**−0.41**	**−0.47**	0.02	0.18	−0.19	**−2.83**
Polished	−0.09	**0.37**	0.15	−0.08	**−0.37**	**−2.27**
Complex	**0.28**	**0.47**	−0.03	**−0.27**	**0.25**	**−1.33**
Balanced	−0.07	**0.33**	−0.13	**−0.29**	−0.02	**−1.56**
MediumBody	**−3.14**	**−1.28**	**1.14**	**−1.55**	0.34	**−6.51**
FullBody	**2.01**	**−0.35**	**0.32**	−0.22	**0.50**	**−1.42**

Notes: W1, W2, and W3 indicate the three dimensions of the wine–topic space. Z1 and Z2 indicate that the two dimensions of the critic–topic space. Boldfaced estimates are significant at the 0.05 level.

reviewers need to be accounted for in generating unbiased topic propensities in the wine product positioning map. From a managerial perspective, individual critics' language differences indicate that critics prefer different topics and different wines.

Reviewer-adjusted perceptions of the 249 wines are captured in the model by the term ($\Lambda_k W_j$), while preferences are captured by the term ($B_i W_j$), where W_j is a 3-dimensional array of factor scores for wine *j*. (β_i, Λ_k) are the coefficients listed in the second to fourth columns of Table 3. Aside from parsimoniously summarizing reviewer-adjusted perceptions reported in the 640 reviews, these components produce a product positioning map that is akin to those commonly utilized by marketing researchers and managers as an aid in brand management. These perceptions and preferences are depicted in the positioning map in Fig. 2, the first map, which displays the standardized (to the unit circle) factor loadings for wine critics (Λ_i) and topics (Λ_k) as standardized (unit-length) vectors in a 3-dimensional space. Like any other reduced space, this positioning map is invariant to rotation, and, therefore, these underlying dimensions (W1, W2, and W3) are only used as references to produce the plots. Furthermore, brand positions must be interpreted in relation to the directions defined by the vector for each product feature and reviewer preference. In other words, the preference vector for a reviewer points in the same direction of the product features that are relevant in forming the reviewer's preferences. This map suggests that Parker, Wine Spectator (WS), Tanzer, and Haliday have similar preferences and tend to give higher ratings to “complex” (complexity, elegant, layered) wines with a sensometric profile characterized by “vanilla” (vanilla, vanillin), “smoked” (charcoal, grilled, pain grille, toasty, burned), and “chocolate” (bean, chocolate, coffee, espresso). By contrast, Jancis Robinson, Wine Enthusiast, and Wine Review Online (WRO) favor “FullBody” (concentrated, dense, full of flavor) wines that evoke “cloves” (cinnamon, cola, nutmeg) and “berry” (black fruit, cassis, currant, mulberry). While the seven reviewers appear “clustered” in these two groups, as discussed here, the general consensus among them favors wines positioned in the direction of dimension W1 in the first map of Fig. 2, as their preference arrows point in the same direction as

Table 4
Critics' general propensities to use topics in their wine reviews (unit = %).

Term	Wine critic							Average
	Haliday	Robinson	Parker	WRO	WS	Tanzer	Enthusiast	
Floral	1	10	26	7	4	45	2	8
Clove	0	3	1	1	1	2	1	1
Pepper	9	33	38	22	19	58	12	24
Licorice, anise	6	21	17	13	12	31	8	14
Citrus	0	0	0	0	0	0	0	0
Berry	74	75	93	84	76	91	73	83
Tree fruit	12	26	15	17	18	26	15	18
Dried fruit	20	31	16	21	25	24	23	23
Other fruit	54	36	43	45	44	32	49	43
Fresh herb	2	6	3	3	4	6	3	4
Canned vegetable	2	1	2	2	2	2	2	2
Dried herb	15	30	14	18	21	24	18	20
Nutty	0	1	1	0	0	2	0	0
Chocolate	25	21	17	21	22	16	24	21
Vanilla	3	6	2	3	4	4	4	4
Woody	40	24	43	36	31	29	35	34
Smoked	17	14	41	23	16	30	15	21
Earthy	16	20	25	20	18	27	16	20
Mineral	7	19	25	14	12	36	8	15
Petroleum	4	8	8	6	6	11	5	6
Pungent	0	1	1	0	0	2	0	0
Metallic	1	3	1	1	2	1	2	1
Alcohol	4	12	5	6	7	11	6	7
Other microbiological	0	1	3	1	1	5	0	1
Tannin	33	62	33	41	47	56	40	44
Sweet	3	36	20	12	13	58	6	16
Dry	5	21	2	6	10	9	8	7
Crisp	5	9	4	5	6	6	6	6
Polished	12	12	5	9	12	6	13	9
Complex	25	12	34	23	18	19	20	21
Balanced	24	12	20	19	17	12	20	17
Medium body	1	0	1	0	0	0	0	0
Full body	20	9	43	22	14	22	16	20

W1. In this sense, W1 coincides with the general direction of preference for red wines. (Thus, our product positioning map shows the valence of each topic by the reviewer instead of determining the a priori valence of each topic using our own judgments.)

The second map of Fig. 2 displays the positions of the 249 red wines in the same 3-dimensional perceptual/preference space. As discussed earlier, the general preference across all of the seven experts points in the general direction of the first dimension of the latent space, W1. Therefore, wines #29, 106, and 220 would be highly rated by these experts, followed by wines #104, 182, and 209. These wines have strong associations with "vanilla," "cloves" and "fullbody" and project the highest in the direction of preferences for Wine Enthusiast and Wine Reviews Online (WRO), which are aligned with dimension W1. Tanzer, Parker, Wine Spectator (WS), and Haliday would prefer wines #237, 144, 225, and 29 strongly, which evoke "vanilla," "chocolate," "smoked," and "complex" and project high in the direction of their preference vectors. Jancis Robinson would prefer wines #106 and 220 strongly, which have stronger connotations of "cloves," "fullbody," "woody," and "licorice," more directly aligned in the direction of her preference vector. Thus, our maps uncover the direct relationships among critics, products, and topics. Using this information, marketers can promote their wines based on specific topics, mentioning certain critics who like their wines.

4.1. Predictive validity tests

In order to test the predictive validity of the second component of our framework (i.e., the positioning/propensity model) against a close competitor (Lee & Bradlow, 2011) and a simpler benchmark model, we randomly split our data into a calibration sample and a hold-out validation sample. This was done in such a manner that the calibration sample contained at least one review for each wine in order to properly measure the latent scores of each wine, as described earlier, resulting in the calibration and validation samples of 590 and 50 reviews, respectively. We apply the same model formulation discussed earlier (with $p = 2$ and $q = 3$) to the calibration data and use the estimated parameters, along with the latent scores for critics and wines, to compute the likelihood for the observed topics and predict the quality ratings in the holdout sample. In other words, we use the parameter estimates and the factor scores obtained from the calibration sample to predict what the critics would say about the wines in the 50 reviews in the hold-out sample, and how they would have liked these wines. We acknowledge that these predictions are far from perfect (particularly, for the topic predictions involving each review), because they involve the behavior of individual human beings on single events (reviewing a single wine) on a highly subjective task (unaided text writing). Notwithstanding, this predictive validity test can provide us with a general sense of how well our model behaves.

For an equitable assessment of the predictive validity of our framework, we use several benchmarks. First, we use a constrained version of our framework that ignores linguistic differences across reviewers, which produces a standard product positioning map. Second, we use a "naïve" benchmark model, a log-linear analysis estimating the log-odds for the 33 topics for each wine and critic, which we use to predict the probability that each topic would be mentioned in each of the 50 hold-out reviews, given the wine and critic involved. A similar process is applied for the quality ratings, using a linear regression model. As a third benchmark, we apply the procedure used by Lee and Bradlow (2011), our closest extant competitor, as shown in Table 1. For this critical benchmark model, we conduct correspondence analysis on the wine/critic by topic matrix to generate dimension coordinates to generate a product positioning map, and a regression of the average wine ratings (re-scaled to account for differences in scale across critics) on the wine coordinates. Our screen test indicates two dimensions as the optimal number of dimensions. Using the correspondence analysis results, we compute the binary probability for each of the 33 topics for each combination of wine and critic in the hold-out sample. Then, we use the average quality ratings (re-scaled to the original scales) for its quality prediction.

A detailed comparison of the four models is included in Table A1 of our web appendices, in terms of predictive performance measured by the prediction error (specifically, the mean absolute percent error) for the quality rating. Across the 50 wines held out for predictive validity testing, our proposed model produces a lower mean absolute percent error (2.3%) in predicting the

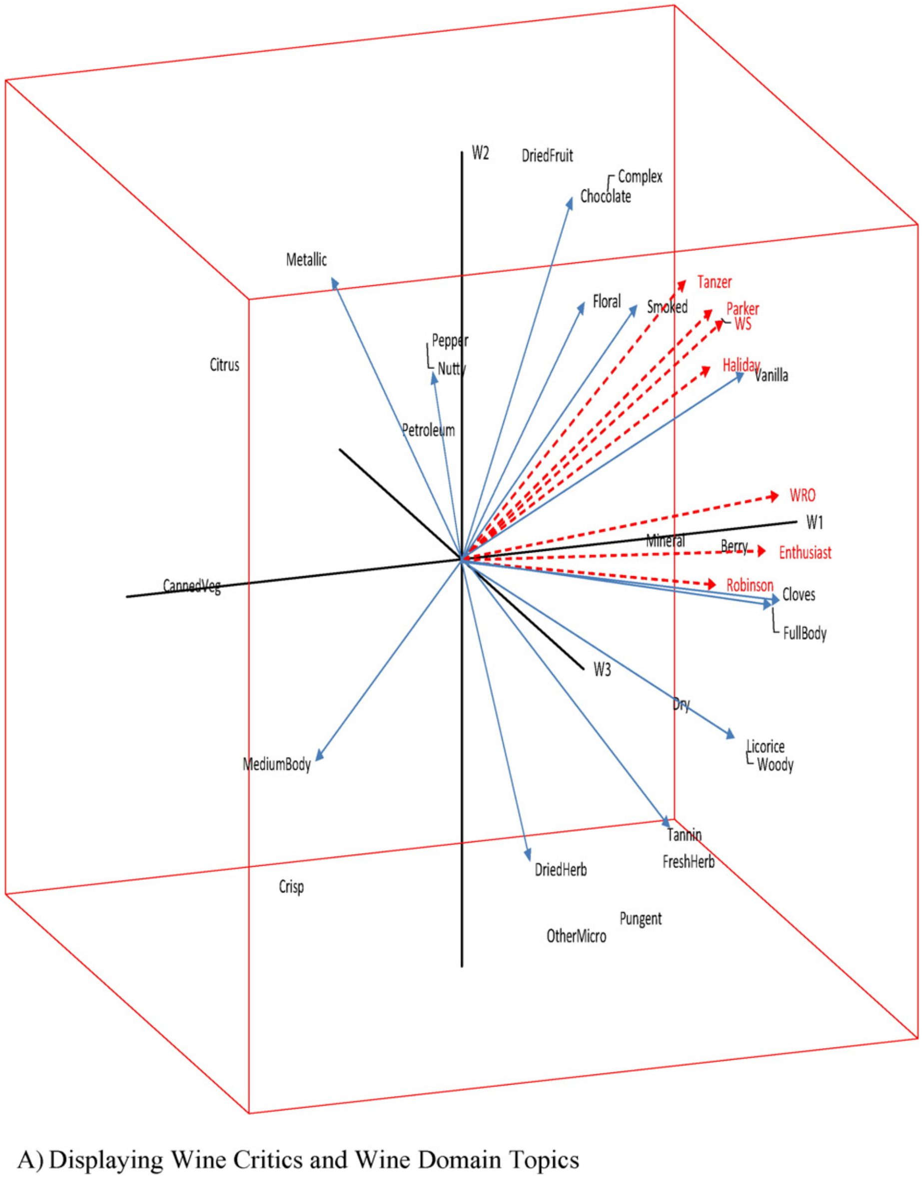

Fig. 2. Positioning map: perceptions and preferences for red wine attributes. Panel A. Displaying wine critics and wine domain topics. Panel B. displaying individual wine locations. Notes: The first map displays wine critics (Tanzer, Parker, WS, Haliday, WRO, Enthusiast, and Robinson) and wine domain topics on the three-dimensional map. The second map locates individual wines on the same map.

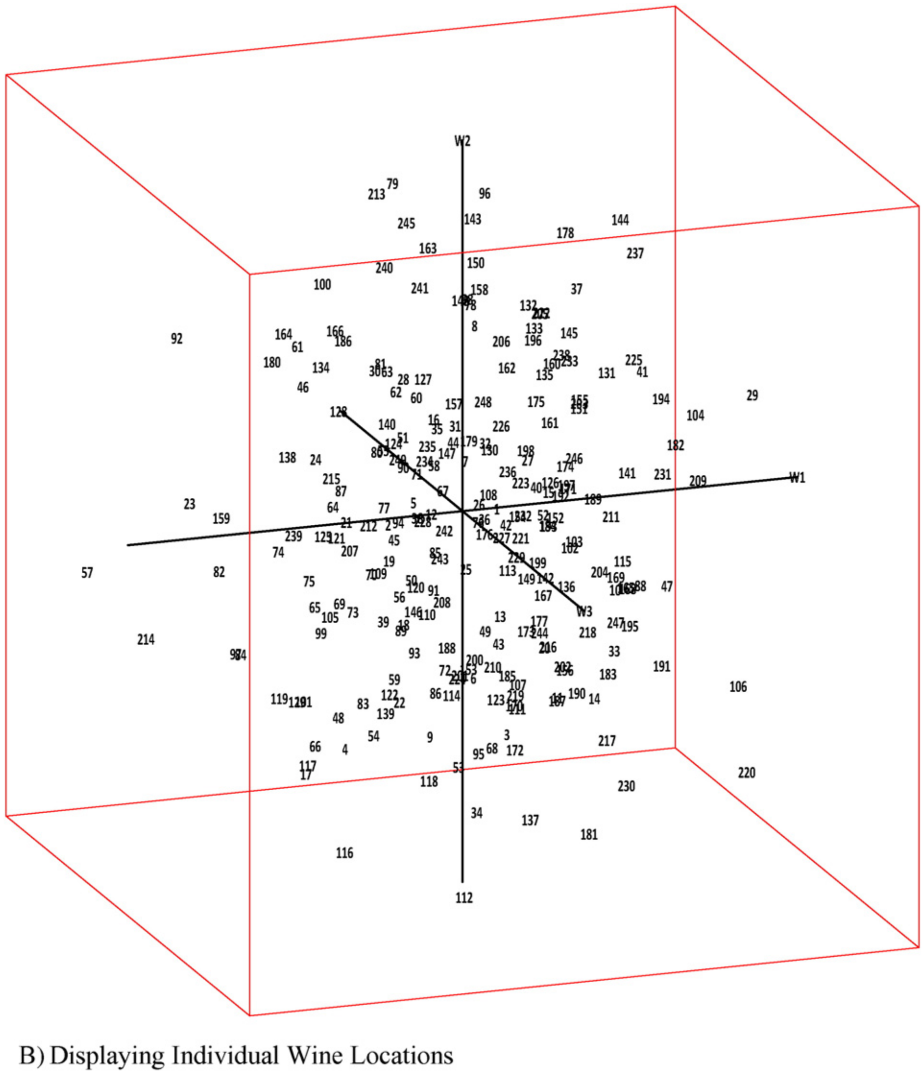

Fig. 2 (*continued*).

quality ratings than the benchmark models (2.4%, 2.4%, and 2.6%, respectively). The differences in predictive performance are small because these quality ratings tend to have much lower variances across wines than across critics. This happens for two reasons. First, as shown in Fig. 2, the seven wine critics are fairly convergent in their preferences, resulting in low variance in the quality ratings across critics. Therefore, most of the variance is accounted by the differences across wines. Second, the predictive performance of the naïve positioning model is already quite good (2.4% in mean absolute percent error = MAPE), resulting in a flooring effect for any improvements by more sophisticated ones, because MAPE has the lowest bound of zero. In summary, our hold-out tests indicate the improved performance of our proposed model because our model considers both the differences in reviewers' propensities to use certain topics in their reviews and the consensual perceptions of the product by multiple reviewers.

5. Integrating hotel reviews from multiple sources into a product positioning map

Some part of the translation of expert wine reviews into a product positioning map in the previous section may not be generalized to certain non-sensorial experience products primarily based on ordinary consumer reviews. As another illustration of how our proposed framework can be extended to typical product reviews written by ordinary consumers, we analyze 425,970 hotel reviews on five different types of travelers (business, couple, family, solo, and others) from five hotel review websites (Expedia, Hotels.com, Priceline, TripAdvisor, and Yelp), for 110 hotels in Manhattan, New York. We used all of the collected hotel reviews for our ontology learning to build a complete hotel taxonomy, using 345,826 reviews posted between December 2012 and November 2013. Similar to the wine reviews, in addition to the text of each hotel review, we gathered the summary rating given by the reviewer to the hotel, on the rating scale used by the website (a 5- or 10-point scale). This hotel review application serves to address three potential issues related to our earlier wine review application. First, the hotel application relies on reviews posted by ordinary consumers rather than experts. Although ordinary consumers tend to follow experts' opinions for sensory products, such as wines, there is some evidence indicating that consumers show different tastes from experts (Holbrook, Lacher, & LaTour, 2006; Liu, 2006). Second, it allows us to demonstrate how our framework can be adapted to a situation where each consumer posts only a small number of reviews. To address this data scarcity issue pertaining to individual reviewers, we adapt our framework to the types of travelers and the multiple websites, while allowing for differences in linguistic propensities and perceptions across the types of travelers and websites. Third, it demonstrates that our framework does not require a "seed" dictionary or existing expert knowledge to produce a category-specific domain taxonomy.

A growing number of consumers are relying on online consumer reviews from online travel sites such as Hotels.com and TripAdvisor.com when they search for a hotel. Without much effort, travelers can obtain valuable information in choosing the right hotel from like-minded consumers. Anderson (2012) found that the "guest experience factor" is the most important factor in hotel selection, even more important than the location factor. The importance of this guest experience factor is growing, as more travelers post and read hotel reviews. Accordingly, the impact of consumer reviews on hotel businesses is growing, and these reviews can often make or break individual hotels, particularly small ones with limited promotion resources.

Our goal in this application is to generate an unbiased product positioning map of 110 hotels in Manhattan based on the perceptions and preferences conveyed by different types of travelers on multiple websites. Because we integrate the reviews from different types of travelers across multiple websites, we produce more reliable perceptions and preferences that are less susceptible to outliers (e.g., fake or extreme reviews). We also account for potential biases in the review content of these websites, which may be caused by their different reviewing policies. Specifically, Yelp and TripAdvisor use an open review policy by allowing anyone to post a review. By contrast, the other three hotel sites (i.e., Hotels.com, Expedia, and Priceline) follow a closed review policy by allowing a review only after a verified stay. Therefore, our integrated hotel-positioning map allows consumers to find hotels that satisfy their idiosyncratic preferences without being swayed by fake or extreme reviews. Similarly, hoteliers can use our map to see where they stand against their competitor hotels based on clients' experience vocabulary. Their map positions would be based on the general reviews from their overall client base and would not be swayed by extreme consumers or manipulating rival hotels.

5.1. Developing a hotel taxonomy from online hotel reviews by ontology learning

As explained in Web Appendix A, we apply the iterative five-step ontology learning process to generate the hotel service taxonomy in this hotel review application. The main difference from the previous ontology learning is that, in the wine example, we had an initial lexicon to start with, while here we do not rely on any prior knowledge. As shown in Table 5, our final hotel taxonomy is composed of three layers: 6 grand topics, 27 topics, and 2690 descriptive terms. Compared to our wine review taxonomy building, we indicate two major technical differences in this hotel review taxonomy building. First, we use web crawling techniques to automatically search five targeted hotel review websites for vast amounts of hotel reviews in our target market (Manhattan, New York), whereas we manually collected a limited number of wine reviews in our wine application. Considering the vast number of reviews available from multiple websites, this automatic way of collecting a corpus of online information is an efficient and effective way for text gathering.

Second, we apply a basic level of sentiment analysis to classify the hotel topics. By contrast, we did not apply this analysis in the wine application because wine critics use flowers, fruits, and other food and non-food items as metaphors for flavors and aromas. The complex and nuanced nature of wine flavors and aromas (e.g., "flavors turn *metallic* on the finish," "with an *earthy* animal component") would increase errors in estimating the valence of each topic in each review, which is typically required in sentiment analysis. In the case of hotels, we can rely on the valence of the terms used for the topic in determining the valence of each topic, because it is mostly a straightforward process to determine the valences of most terms. For example, we use the valence of the adjective–noun pair (e.g., "unsafe neighborhood" and "cramped space" as negative terms) and the valence of the adverb of the adverb–verb pair (e.g., "highly recommend" and "definitely stay" as positive terms). As a result, out of the 27 topics, we classified 15 topics as positive topics (e.g., affordable stay, gorgeous view, pleasant experience) and 10 topics as negative topics (e.g., no recommendation, poor service, inconvenient surroundings). The remaining two topics related to travel types (personal travel and business travel) are classified as neutral because these travel types do not have a specific valence per se.

5.2. Representing hotel reviews by traveler type and review source in a product positioning map

This second application to hotel reviews can be seen as a variant of the first application to wine reviews, in that the hotel topic code counts were aggregated by traveler type (business, couple, family, solo, and others) and review source (Expedia, Hotels.com, Priceline, TripAdvisor, and Yelp). As a result, we end up with 18 combination cases of traveler type and review source, which we label as "meta-reviewers" (the number of the combinations of the five travel types and the five review sources is less than 25 because some sites did not have all of the five traveler type indicators. In particular, Yelp did not indicate the type of traveler in their reviews at all). From a managerial perspective, hotel managers would not be interested in individual reviewers' characteristics and preferences because they manage their hotels at the consumer segment level. On the other hand, each meta-reviewer type (e.g., Expedia business travelers, TripAdvisor family travelers) can be considered to be a distinct consumer segment that can be targeted by hotel managers. For each combination of the 18 meta-reviewers and 110 hotels, our dependent variables are the total count for each topic and the sum of the hotel ratings across all travelers in the specific "meta-reviewer" group. Therefore,

Table 5
Summary of the final hotel taxonomy.

Grand topic	Topic (abbreviation; valence)	Representative descriptive terms (the total number of terms in each topic)
Travel type	Personal travel (FunTrip; Neutral)	Birthday trip, family event, leisure traveler, vacation (122)
	Business travel (BusTrip; Neutral)	Business conference, conference room, job interview (37)
Hotel facilities	Affordable stay (Cost+; Positive)	Affordable, bargain price, conference rate, free room (93)
	Expensive stay (Cost-; Negative)	Additional charge, expensive breakfast, hidden charge, rip-off (77)
	Design and décor (Décor; Positive)	Beautiful old hotel, elegant lobby, great design, thoughtful touch, trendy décor (63)
	Interesting surroundings (Neighb+; Positive)	Adjacent restaurant, bistro next door, close to subway, main tourist area, nearby café (88)
	Inconvenient surroundings (Neighb-; Negative)	Major construction, no nearby restaurant, traffic jam, unsafe neighborhood, railway (39)
	Cool atmosphere (Ambnce+; Positive)	Beautiful boutique hotel, cool atmosphere, excellent standard, family-friendly hotel, upscale hotel (223)
	Poor atmosphere (Ambnce-; Negative)	Basic amenity, cramped space, no-frills hotel, outdated hotel, poor condition (84)
	Great facilities (Faclty+; Positive)	Accessible, amazing lobby, convenient laundromat, fast elevator, great facility (72)
	Broken facilities (Faclty-; Negative)	Elevator noise, maintenance issue, malfunction, narrow hallway, tiny elevator (54)
Bedroom	Awesome rooms (BdrmType+; Positive)	Awesome room, clean spacious room, great suite, superior room, wheelchair accessible room (130)
	Awful rooms (BdrmType-; Negative)	Awful room, college dorm room, room smell, thin wall, unrenovated room (66)
	Pleasant room amenities (BdrmFixtr+; Positive)	Awesome shower, big bed, clean linen, floor-to-ceiling window, large closet (233)
	Disappointing room amenities (BdrmFixtr-; Negative)	Brown water, cockroach, cold water, dirty bed, share bathroom (256)
	Modern bedroom appliances (Applnc+; Positive)	Big flat screen, complementary Wi-Fi, free high-speed, gratis Internet, large fridge (71)
	Outdated bedroom appliances (Applnc-; Negative)	Basic cable, broken phone, expensive internet, internet charge, slow internet speed (69)
Extra services	Convenient in-house eating (dining; positive)	Diner downstairs, in-house Starbucks, in-house restaurant, onsite restaurant, rooftop bar (137)
	Free eating (snacks; positive)	Complimentary breakfast, free water bottle, free cocktail (143)
	Extra activities (activity; positive)	Club lounge, good workout room, great gym, rooftop pool, tennis court (49)
	Gorgeous views (view; positive)	Gorgeous view, panoramic view, skyline view, unobstructed view (43)
Staff services	Excellent services (Srvc+; Positive)	Competent staff, excellent customer service, first class service, friendly doorman, welcome treat (118)
	Poor services (Srvc-; Negative)	Bad customer service, basic service, lengthy wait, major complaint, unhelpful staff (78)
Overall experience	Pleasant experiences (Exprnce+; Positive)	Amazing trip, delightful surprise, good weather, peaceful night, unexpected surprise (57)
	Unpleasant experiences (Exprnce-: Negative)	Awake all night, bad weather, false alarm, jet-lag, ongoing renovation (34)
	Back next time (Overall+; Positive)	Amazing hotel, awesome stay, back next time, best hotel, good overall experience (173)
	No recommendation (Overall-; Negative)	Bad hotel experience, mixed feelings, not recommend, unpleasant stay, wouldn't suggest (81)

instead of a binary logit link function, as in Eq. (3), the dual space-reduction model would require a Poisson formulation for the topic counts. However, given the vast amount of reviews available (the counts range from 0 to 1282 across hotels, sources, and topics for all reviews posted in the past year), a normal (identity link) formulation, as we did for the quality ratings in the wine application in Eq. (4), is more appropriate than the Poisson formulation. Moreover, there is a large variation in the total number of reviews posted by the different "meta-reviewers" for each of the 110 Manhattan, New York hotels, a factor that must be accounted for in the model formulation. This aspect required a minor extension of our original model, as follows:

$$u_{ijk} = u_k + \Gamma_k Z_i + \Lambda_k W_j + \beta_k R_{ij} + \varepsilon_{ijk}, \tag{6}$$

where all variables and parameters are as specified earlier, except for:

u_k intercept representing the general tendency of any reviewer (critic) using topic k to describe any brand (hotel), to be estimated.
Γ_k p-dimensional vector of factor loadings for topic k, to be estimated,
Z_i p-dimensional vector of independent, standard normal latent scores for reviewer (critic) i,
Λ_k q-dimensional vector of factor loadings for topic k, to be estimated,

W_j q-dimensional vector of independent, standard normal latent scores for brand j,
R_{ij} total number of reviews written by meta-reviewer i on hotel j,
β_k adjustment in topic k for the number of reviews by meta-reviewer i for hotel j,
ε_{ijk} i.i.d. random error.

5.3. Empirical results for the hotel reviews

Based on the BIC criterion, we concluded that the best solution for the model in Eq. (6) was the one with two dimensions in the positioning map ($q = 2$) and four dimensions for the linguistic propensity map ($p = 4$). The parameter estimates from this optimal model are shown in Table A2 in the web appendices because the interpretation of most parameters remains the same as before. More important for the purpose of product positioning is the interpretation of the perceptual/preference and propensity maps. Because our dependent variable in this second application is the total count for each hotel and meta-reviewer on each topic, the term $\Gamma_k Z_i$ in Eq. (6) shows how much more/less frequently than the mean a topic is mentioned by a meta-reviewer. Importantly, this estimation is implemented after adjusting for the total number of reviews posted, thereby showing the linguistic propensities by each meta-reviewer. These deviations are plotted in Fig. 3 for some select cases, for the purpose of illustration. In particular, it is noted that Priceline is not included because it yields a very low proportion (only 1.3%) of all hotel reviews. Because our main points are clearly shown in the figure without this hotel website, we exclude it to avoid overcrowding the figure. Looking at the top panel of Fig. 3, one can see that business travelers and couples from Expedia have a higher than average propensity to mention the topic BusTrip (e.g., business conference, business meeting, business stay, conference room, job interview) in their reviews, while other travelers have a low propensity to mention the BusTrip topic. On the other hand, as one would expect, business travelers have an extremely low propensity to mention FunTrip (e.g., birthday trip, family event, getaway with friends, leisure traveler, vacation) in their reviews. The middle panel of Fig. 3 shows the propensities of travelers posting their reviews at Hotels.com, which appear similar to those observed for Expedia travelers, except that couples have a lower propensity to mention the BusTrip topic and a higher propensity to mention the FunTrip topic on Hotels.com. The bottom panel of Fig. 3 displays the propensities of travelers on TripAdvisor. Specifically, this website shows that couples and families have an extremely high propensity to mention the FunTrip topic. These two meta-reviewer groups also mention the positive overall experience (overall+) topic much more often than business travelers because the primary purpose for most of these travelers is seeking travel pleasure. This pattern is consistent with the two traveler groups' frequent mentions of the positive neighborhood (Neighb+) and positive ambience (Ambnce+) topics. Next, TripAdvisor shows solo travelers as an explicit traveler type that does not appear on the other four websites. Solo travelers also tend to mention Business Trip more often than family and couple travelers, suggesting that many business travelers are solo travelers.

A notable difference among the hotel websites is that TripAdvisor reviews tend to contain more information about the hotels in terms of the topics identified in our lexicon, implying a unique nature of the website; TripAdvisor, a travel guide website, contains rich travel information from travelers highlighting travel attractions, whereas the other four hotel websites are more business-driven and, accordingly, post consumers' experiences with their hotel services without such varied travel information. Therefore, the environment of TripAdvisor attracts reviewers who are willing to post longer reviews. Furthermore, TripAdvisor (along with Yelp) has an open review policy and allows anyone to post his or her reviews, whereas the other three websites allow only actual hotel clients to post their reviews right after their stays. In the latter case, when clients are asked to post their reviews, they may often respond somewhat involuntarily and write relatively short reviews. Therefore, overall, websites with an open review policy tend to have longer reviews. Thus, our meta-reviewer analysis in Fig. 3 reveals many insights that may not have been captured otherwise. Most importantly, failure to account for such propensities across traveler types and review sources would have resulted in a biased assessment of consumer perceptions and preferences.

Fig. 4 displays the standardized (to the unit circle) factor loadings (Λ_k) for topics (the top panel representing perception) and meta-reviewers (the bottom panel representing preferences). In the perception map, one can expect that the desirable topics (with positive valence) point in the opposite direction (north) against the undesirable ones (with negative valence). In other words, hotels positioned in the northern region should be more desirable. This is confirmed by the preference vectors for the 18 meta-reviewers in the bottom panel. In the bottom panel, all vectors point north, indicating that travelers generally share similar preferences for hotel services. For example, irrespective of the traveler type and the review source, travelers would prefer affordable stay (against expensive stay), cool atmosphere (against poor atmosphere), excellent services (against poor services), and so on. However, one can see some differences in how much each meta-reviewer values specific topics, as shown in the preference map. For example, the four meta-reviewers on Priceline (Family, Couple, Solo, and Business) tend to have preferences slanting to the left. That is, preferences by travelers on Priceline tend to be more closely correlated with experience + (e.g., good weather, peaceful night), decor (e.g., elegant lobby, trendy décor), and facility + (e.g., accessible, fast elevator). The meta-reviewers on TripAdvisor have preferences pointing more directly north and, accordingly, more closely correlated with overall + (e.g., awesome stay, back next time), neighborhood + (e.g., close proximity, safe location), and dining (e.g., in-house Starbucks, rooftop bar). The meta-reviewers on Expedia and Hotels.com tend to have preferences leaning slightly to the right. Yelp, without any traveler type distinction, shows preferences close to the north-central area by averaging out the differences among traveler types. On the other hand, preferences by business travelers tend to point to the left, whereas preferences by personal travelers (families and couples) tend to point to the right. Importantly, overall, the 18 meta-reviewers classified by a

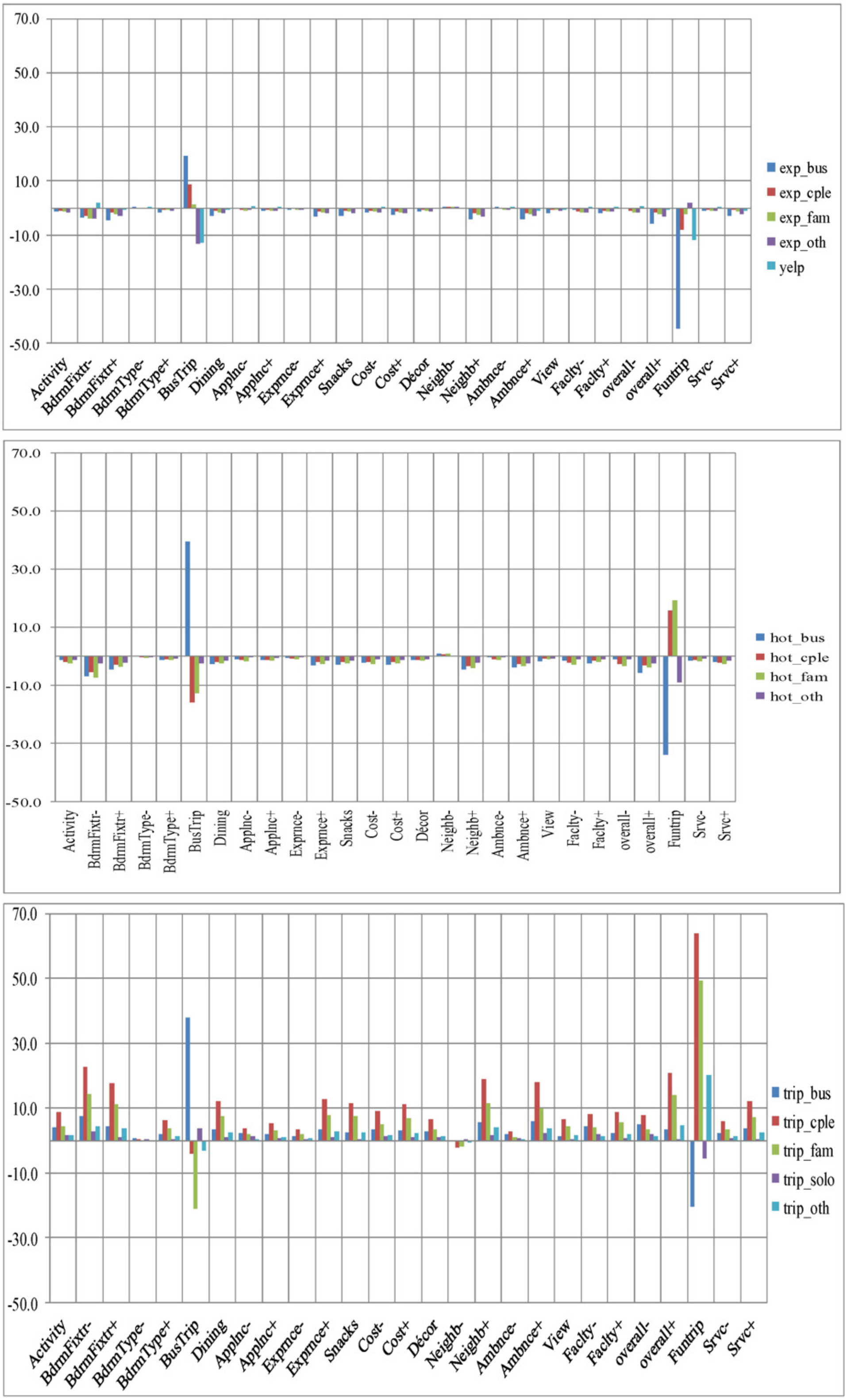

Fig. 3. Linguistic propensities for travelers posting reviews on various hotel websites. Notes: The right-hand side acronyms indicate hotel meta-reviewers. Each acronym is composed of two parts: hotel review website and traveler type. For the hotel review website, exp. = Expedia, yelp = Yelp, hot = Hotels.com, and trip = TripAdvisor. For the traveler type, bus = business, cple = couple, fam = family, solo = solo, and oth = others. The original term for each acronym at the bottom of each graph can be found in Table 5.

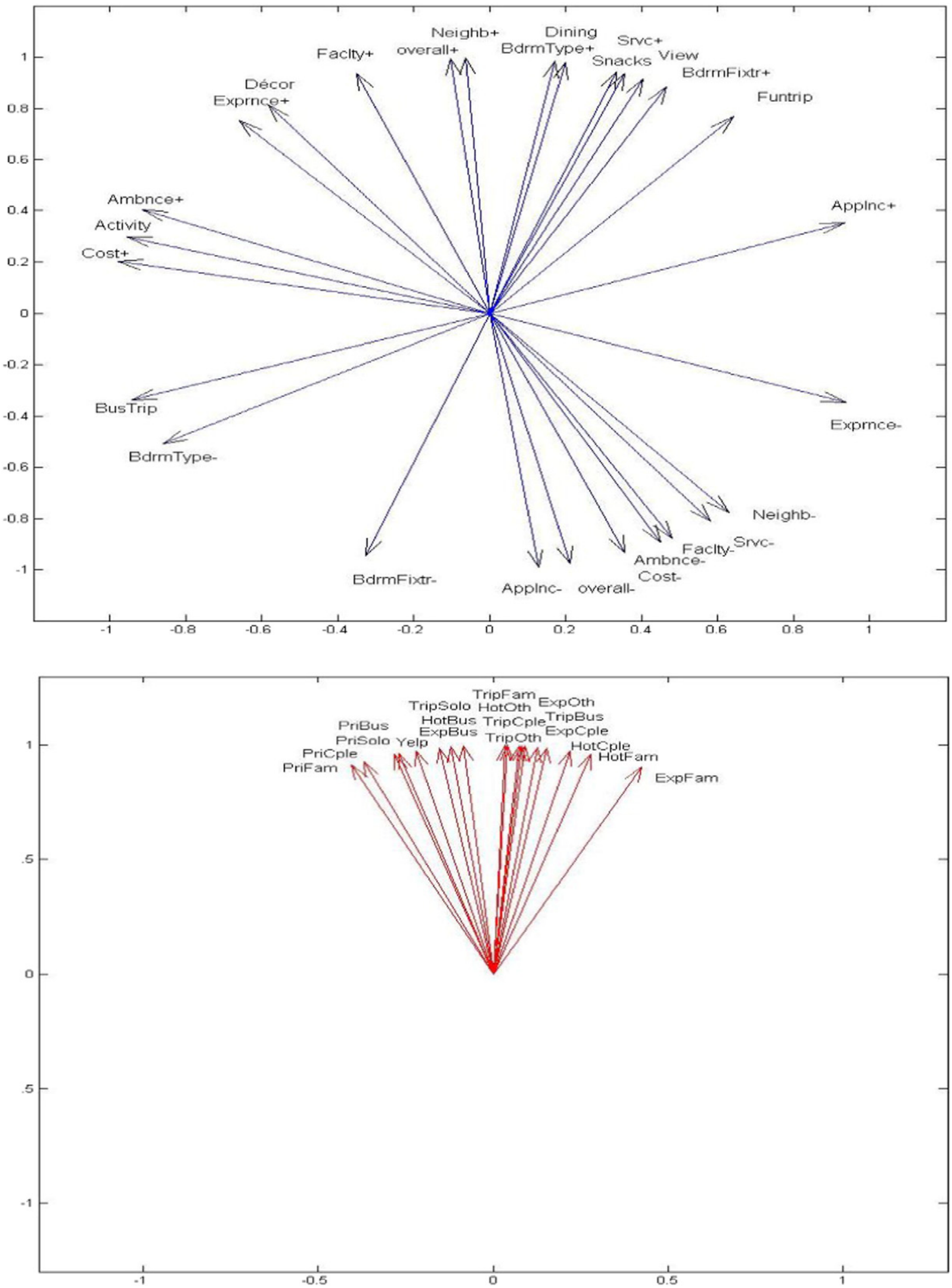

Fig. 4. Perceptual and preference maps for hotels in Manhattan.

combination of the two categories (traveler type and review source) reveal that each group has unique preferences, as well as idiosyncratic writing styles and perceptual acuities. Our framework demonstrates such subtle but critical preference differences among consumer segments (i.e., meta-reviewers) in connection with the perception/preference map (in Fig. 4) and the product positioning map (in Fig. 5).

Fig. 5 shows the product positioning map, placing the 110 hotels on the same perception/preference space. We observe that a big cluster of hotels at the center are typical hotels serving average hotel customers. In contrast, hotels far away from the cluster (e.g., Pod51 and ParkCtrl) would serve atypical customers with less competition. As in Fig. 4, overall, travelers tend to prefer the hotels in the northern region to those in the southern region. Accordingly, hotels in the northern region tend to receive higher customer satisfaction ratings. Besides, we observe that the origin has a high hotel concentration, with more hotels in the desirable northern side than in the southern side. This implies that some hotels near the origin could re-position themselves in the northern direction by improving their business in certain features, while strengthening their appeal to specific traveler segments by moving slightly to the northeast or the northwest. Lastly, because the two panels of Figs. 4 and 5 are mapped in the same latent space, it would be useful for practitioners to overlap them so as to gain further insights into the relationships of topics, meta-reviewers, and hotels.

6. Discussion and conclusions

6.1. Theoretical and managerial implications

In this study, we proposed a framework that combines ontology learning-based text mining of online product reviews with psychometric mapping to translate the extracted text into a product positioning map, while accounting for differences in perceptual acuity and writing style across reviewers. At first glance, the perceptual and preference maps we developed appear similar to those provided by many commercial marketing researchers. However, under closer scrutiny, our maps are distinct from the traditional positioning maps one finds in marketing research reports, and from the more recent positioning maps produced from aggregate counts of online reviews. To produce traditional positioning maps, researchers determine a priori the features or attributes that will define the dimensions of the map, and, then ask a sample of consumers to rate competing brands on these attributes (Carroll & Green, 1997; Faure & Natter, 2010). Even when using text from online product reviews, researchers often rely on summary counts from a pre-defined list of words or product features (Aggarwal et al., 2009). Thus, positioning mapping based on consumer surveys

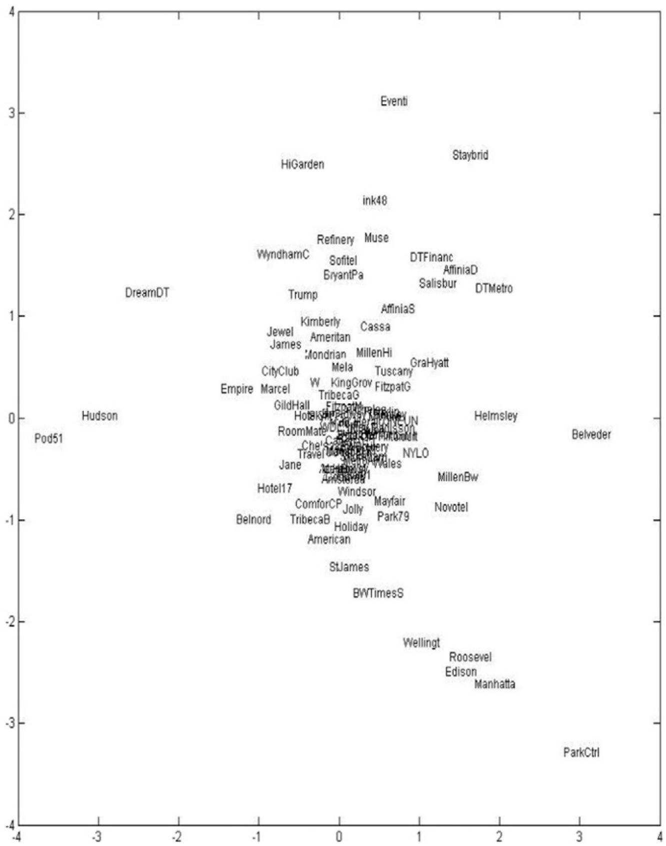

Fig. 5. Brand mapping for hotels in Manhattan.

forces consumers to view the competing brands through the lenses defined by the researcher, and to rate these brands on a pre-defined scale, creating researcher/manager bias, while maps produced from aggregated online reviews incur another bias by ignoring reviewer differences in writing style and/or perceptual acuity. On the other hand, survey-based product positioning maps are more useful than those based on product reviews when introducing new brands or brand extensions because, at that point, one cannot find online reviews posted by consumers.

Even though the starting point in our proposed framework is text selected from the Internet in the form of product reviews, the end result is a perceptual/preference map similar to those commonly used by marketing researchers and brand managers. We see this as a benefit, as marketing practitioners are already familiar with how to interpret these positioning maps and make decisions based on the maps (Lilien & Rangaswamy, 2006). Besides, the parsimony with our space reduction approach will be more dramatic as the numbers of reviewers, brands under review, and topics extracted from the reviews increase.

Marketing researchers have used several psychometric mapping techniques to turn product reviews into product maps, such as correspondence analysis (Lee & Bradlow, 2011) and multidimensional scaling (Netzer et al., 2012; Tirunillai & Tellis, 2014). However, these studies did not consider reviewers' idiosyncratic writing styles and, accordingly, did not place reviewers on the maps they created. By contrast, our reviewer–product dual space-reduction model generates a comprehensive map that simultaneously locates brands, reviewers (or meta-reviewers), and product features. This allows us to obtain insights into how these three dimensional objects can interact, as shown in Figs. 2, 4, and 5. Moreover, we translate product reviews into a product positioning map; at the same time, we adjust these reviews for perceptual and stylistic differences among the reviewers so that the resulting product õpositioning map reflects perceptions and preferences rather than linguistic differences.

Our research shares noteworthy similarities and dissimilarities with a recent study by Tirunillai and Tellis (2014) (refer to Table 1). Both studies generated a product map using product attribute terms directly elicited from consumers' product reviews. However, they differ in two ways. First, in eliciting product attribute terms, our research used ontology learning, whereas TT used latent Dirichlet allocation. Second, in generating a product map, our research used a dual space-reduction mapping technique to generate both reviewer and product locations in the same space, whereas TT used multidimensional scaling to generate only brand locations.

From a managerial perspective, our framework for mapping individual or meta-reviewer preferences and perceptions on the same latent space can produce useful insights that would not be otherwise obtained, as discussed in our empirical application for hotel reviews. When we identify proper meta-reviewers, each meta-reviewer group can be seen as a legitimate consumer segment deserving a separate marketing strategy. In our hotel application, we used a combination of two criteria (traveler type and review source) to identify 18 meta-reviewer groups. The bottom panel of Fig. 4 suggests that there are notable differences in preferences across these groups. A hotel manager, given his or her hotel position in the same space, can develop an informed market strategy targeting proper consumer segments (e.g., business travelers vs. family travelers, Hotels.com customers vs. Priceline customers). For example, family travelers seek fun activities, whereas business travelers focus on getting their job done. Furthermore, managers can potentially identify weak competition areas, where consumer demand exceeds hotel supply on the positioning map. Because the map allows managers to identify what types of consumer segments have a strong demand in the weak competition areas, it may be used to develop a niche marketing strategy targeting those consumer segments as primary targets.

Based on their learning from our approach grounded in consumer reviews revealing their nuanced experiences, online hotel websites (as the review source in our application) can possibly implement their own business policies (including their reservation and price policies), differentiated from generally similar competitors (e.g., deep discounts for promoted hotels, easy cancelation for loyal members, highlighting free parking vs. valet parking, rich travel information emphasizing nearby attractions, website designs). Besides, some websites allow anyone to post their reviews, whereas others only allow their true clients to post reviews after a confirmed hotel stay. These subtle but important differences among hotel websites attract different types of customers, creating different positions for such meta-reviewers. Our approach allows hotel managers to look into these customer differentiation issues and develop more informed brand positioning strategies. Thus, our use of the dual-space reduction method to capture both brands and reviewers on the same map, while correcting for linguistic differences can be more effective in matching existing brands to their proper target consumers based on a vast volume of product reviews. On the other hand, those product reviews can contain "fake" and "extreme" reviews that can distort observed perceptions and preferences. These distortions can be minimized by analyzing a large volume of online reviews and by adjusting for the idiosyncrasies of the websites and reviewer types, as we propose.

6.2. Limitations and future research avenues

Our framework requires the creation of a domain-specific taxonomy based on a large amount of consumer or expert reviews. The analyst needs to have or develop strong substantive knowledge about the domain to create a domain taxonomy, which can initially take a good amount of time. However, once a domain taxonomy is created, the taxonomy can be repeatedly used for additional reviews. Similarly, it takes some time to collect a large amount of reviews by automatic web scrawling because review websites may possess some unique structures that cannot be perfectly captured by a general web crawling script. Although our framework can initially process a huge amount of information efficiently, it requires a significant amount of time in data collection and taxonomy building. Besides, our dependence on consumers' self-reported post-consumption experiences precludes the application of our proposed framework for new product development, unless the goal of the new product is to satisfy unmet expectations identified in the product reviews. For such a situation, when the manager needs to develop product positioning maps based on pre-consumption perceptions and preferences (i.e., new product development), marketing researchers can take advantage of various

MDS techniques developed for the spatial analysis of multi-way dominance data involving consumers' stated preferences, choices, consideration sets, and purchase intentions (DeSarbo, Park, & Rao, 2011).

Our empirical analysis involving two product categories, wines and hotels, did not consider dynamic or geographical extensions of the product positioning maps. Future extensions of our proposed framework include a dynamic comparison of product positioning maps across multiple time periods for the same product category to help managers understand consumers' shifts in perceptions and preference changes (Lee & Bradlow, 2011; Netzer et al., 2012; Tirunillai & Tellis, 2014). Extending our framework to longitudinal comparisons would be straightforward. In addition to a temporal extension of our research, we can conceive of geographical extensions. However, unlike the temporal extensions, these geographical extensions would involve multiple languages and cultures. Because translation between languages does not convey word meanings perfectly, and consumer reactions vary based on their cultural background, various issues arising from translations and cultural interactions would complicate the potential extensions of our framework. We leave these extensions for future research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijresmar.2016.05.007.

References

Aggarwal, P., Vaidyanathan, R., & Venkatesh, A. (2009). Using lexical semantic analysis to derive online brand positions: An application to retail marketing research. *Journal of Retailing*, *85*(2), 145–158.

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research*, *13*(4), 411–454.

Anderson, C. K. (2012). The impact of social media on lodging performance. *Cornell Hospitality Report*, *12*(15), 6–11.

Archak, N., Ghose, A., & Ipeirotis, P. G. (2011). Deriving the pricing power of product features by mining consumer reviews. *Management Science*, *57*(8), 1485–1509.

Buitelaar, P., Cimiano, P., & Magnini, B. (2005). *Ontology learning from text: Methods, evaluation and applications.* IOS Press.

Camacho, N., De Jong, M., & Stremersch, S. (2014). The effect of customer empowerment on adherence to expert advice. *International Journal of Research in Marketing*, *31*(3), 293–308.

Carroll, J. D., & Green, P. E. (1997). Psychometric methods in marketing research: Part II. Multidimensional scaling. *Journal of Marketing Research*, *34*(May), 193–204.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(August), 345–354.

Cimiano, P. (2006). *Ontology learning and population from text: Algorithms, valuation and applications.* Germany: Springer.

Decker, R., & Trusov, M. (2010). Estimating aggregate consumer preferences from online product reviews. *International Journal of Research in Marketing*, *27*(4), 293–307.

DeSarbo, W. S., Park, J., & Rao, V. R. (2011). Deriving joint space positioning maps from consumer preference ratings. *Marketing Letters*, *22*(1), 1–14.

Faure, C., & Natter, M. (2010). New metrics for evaluating preference maps. *International Journal of Research in Marketing*, *27*(3), 261–270.

Feldman, R., & Sanger, J. (2007). *The text mining handbook: Advanced approaches in analyzing unstructured data.* Cambridge University Press.

Ghose, A., Ipeirotis, P. G., & Li, B. (2012). Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. *Marketing Science*, *31*(3), 493–520.

Holbrook, M. B., Lacher, K. T., & LaTour, M. S. (2006). Audience judgments as the potential missing link between expert judgments and audience appeal: An illustration based on musical recordings of "My Funny Valentine,". *Journal of the Academy of Marketing Science*, *34*(1), 8–18.

Kamakura, W. A., & Wedel, M. (2000). Factor analysis and missing data. *Journal of Marketing Research*, *37*(4), 490–498.

Kamakura, W. A., Basuroy, S., & Boatwright, P. (2006). Is silence golden? An inquiry into the meaning of silence in professional product evaluations. *Quantitative Marketing and Economics*, *4*(2), 119–141.

Kostyra, D. S., Reiner, J., Natter, M., & Klapper, D. (2016). Decomposing the effects of online customer reviews on brand, price, and product attributes. *International Journal of Research in Marketing*, *33*(1), 11–26.

Lau, K. -N., Lee, K. -H., Lam, P. -Y., & Ho, Y. (2001). Web-site marketing: For the travel-and-tourism industry. *The Cornell Hotel and Restaurant Administration Quarterly*, *42*(6), 55–62.

Lawrence, R. D., Almasi, G. S., & Rushmeier, H. E. (1999). A scalable parallel algorithm for self-organizing maps with applications to sparse data mining problems. *Data Mining and Knowledge Discovery*, *3*(2), 171–195.

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

Liebermann, Y., & Flint-Goor, A. (1996). Message strategy by product-class type: A matching model. *International Journal of Research in Marketing*, *13*(3), 237–249.

Lilien, G., & Rangaswamy, A. (2006). *Marketing engineering* (Revised 2nd ed.). Bloomington, IN: Trafford Publishing.

Liu, Y. (2006). Word-of-mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing*, *70*(3), 74–89.

Maedche, A. (2002). Ontology learning for the semantic web. *The Kluwer international series in engineering and computer science*. New York: Springer Science + Business Media.

Moon, S., Park, Y., & Kim, Y. S. (2014). The impact of text product reviews on sales. *European Journal of Marketing*, *48*(11/12), 2176–2197.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine your own business: Market structure surveillance through text mining. *Marketing Science*, *31*(3), 521–543.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1–2), 1–135.

Pazienza, M., Pennacchiotti, M., & Zanzotto, F. M. (2005). Terminology extraction: An analysis of linguistic and statistical approaches. *Knolwedge mining. Of the series studies in fuzziness and soft computing, Vol. 185.* (pp. 255–279).

Peck, J., & Childers, T. L. (2003). Individual differences in haptic information processing: The "need for touch" scale. *Journal of Consumer Research*, *30*(3), 430–442.

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.

Staab, S., & Studer, R. (2004). Handbook on ontologies. *International handbooks on information systems*. Springer.

Stone, H., & Sidel, J. L. (2004). *Sensory evaluation practices* (3rd ed.). *Food science and technology, international series*. California: Elsevier Academic Press.

Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent Dirichlet allocation. *Journal of Marketing Research*, *51*(August), 463–479.

Wedel, M., & Kamakura, W. A. (2001). Factor analysis with mixed observed and latent variables in the exponential family. *Psychometrika*, *66*(4), 515–530.