Understanding retail quality of sporting goods stores: a text mining approach

Luke Lunhua Mao
University of New Mexico, Albuquerque, New Mexico, USA

Abstract
Purpose – Sporting goods retailing is a significant sector within the sport industry with the total revenue of this sector reaching $52.2 billion in 2018. Beset with formidable competition, sporting goods stores are compelled to augment their merchandise with services and improve retail quality. The purpose of this study is to investigate retail quality of sporting goods stores (RQSGS).

Design/methodology/approach – Based on 27,793 online reviews of 1481 stores in the United States, this study used Leximancer 4.0, a text mining software, to identify critical retail quality dimensions associated with sporting goods stores, and further explored the most salient dimensions among different levels of ratings.

Findings – Customer service and store aspects are the two higher-order dimensions of RQSGS; holistic experience, manager and staff are three themes under customer service, and product, B&M store and online–offline integration are three themes under store aspects. Furthermore, extreme reviews focus more on customer service, whereas lukewarm reviews focus more on store aspects.

Practical implications – Knowledgeable staff, managers and online–offline integration are instrumental in creating superior retail quality. Sporting goods stores should enhance hedonic and social values for consumers in order to ward off online competitions.

Originality/value – This study explored retail quality dimensions that are pertinent to sporting goods retailing utilizing text mining methods. This study to certain extent cross-validated the existing retailing literature that is developed on alternative methods.

Keywords Retail quality, Sporting goods stores, e-Commerce, Leximancer, Text mining

Introduction
Sporting goods stores, coded as 451110 and defined as “business establishments primarily engaging in retailing new sporting goods” by the North American Industry Classification System (NAICS), is an important sector within the sport industry (U.S. Census, 2019; Eschenfelder and Li, 2007). Major players in this sector include Dick’s Sporting Goods Inc., Foot Locker, Academy Sports and Outdoors and Bass Pro Shops Inc. They retail bicycles, camping equipment, exercise equipment, athletic apparel, athletic footwear, sporting guns and the like. There were 40,122 sporting goods stores in the United States and the total revenue of this sector reached $52.2 billion in 2018 (Hyland, 2018). Roughly speaking, this industrial sector is composed of three categories of stores: a) sporting goods superstores, such as Dick’s Sporting Goods and Academy Sports and Outdoors, which typically have a retailing space of over 35,000 square feet and emphasize high volume sales; b) traditional sporting goods stores, such as Big 5 Sporting Goods, which range in size from 5,000 to 20,000 square feet and typically carry a varied assortment of merchandise; and c) specialty stores, which typically carry a wide assortment of one specific product category, such as shoes, golf or outdoor equipment. Examples include Golf Galaxy, Finish Line and lululemon athletica.

Primarily driven by the increased sports participation, health-conscious consumers and a strong rise in demand for athletic footwear from consumers aged 10 to 19, sporting goods retailing was projected to have an annual growth rate of 1.9% for the following five years.
However, sporting goods retailing is highly competitive (Dick’s Sporting Goods, 2018; Gale, 2018). Sporting goods stores compete not only with adjacent stores in their own industry but also with online retailers (e.g., Amazon), mass merchandisers (e.g., Costco, Walmart), department stores (e.g., Kohl’s), direct channels of sports brands (e.g., Callaway, Nike), along with others (Gale, 2018). In fact, Walmart, with $9.8 billion revenue from sporting goods sales, was the biggest retailer in the market in 2015 (Statista, 2018). The most recent US Census Bureau Consumer Spending Report suggested that sporting goods, hobby, musical instrument and book stores is the only category that has not seen a positive month-over-month increase in 2018, and the total sales decreased by 3.4% in the first 10 months of 2019 (U.S. Census, 2018). Based on the business statuses listed on yelp.com, about 15–22% of sporting goods stores might have been closed in the past five years.

Sporting goods retailing is highly competitive and moderately concentrated (Gale, 2018; Hyland, 2018). The IBISWorld report suggests that the four major players of the industry, including Bass Pro, Dick’s Sporting Goods, Academy Sports and Outdoors and Recreational Equipment have a total market share of 48.7%. About 93% of sporting goods stores employ fewer than 20 people and cater to specialty niches and local consumers. As a result, most stores still rely on the revenues from their brick and mortar (B&M) stores even in the era of rapid growth of e-commerce and multichannel retailing (Chiu et al., 2011). For instance, 87.6% of sales in 2017 at Dick’s Sporting Goods and 78.3% of Cabela’s sales in 2016 occurred in their offline channels (Cabela’s, 2016; Dick’s Sporting Goods, 2018). It is reasonable to assume that the proportion of offline sales for independent local stores and smaller chain stores will be greater.

To effectively compete in the sporting goods market, retailers have extensively augmented their products with additional services. In addition to such services as installation, warranty, credit cards, return and exchange, sporting goods retailers may utilize other more idiosyncratic tools (Homburg et al., 2002). For instance, specialty running shoe stores may augment their products with gait analyses; tennis pro-shop with stringing services; bike shops with repair services; yoga equipment with training classes; and general stores with loyalty programs. The demarcation between products and services has been blurred. Service quality matters to sporting goods stores as much as to any pure service firms, because it not only elicits positive behavioral consequences and engenders consumer loyalty (Zeithaml et al., 1996) but also enhances retailing differentiation (Davis-Sramek et al., 2009; Homburg et al., 2002; Jiang and Zhang, 2011).

Retail quality is considered more complex than quality of pure service as retail stores sell products, together with services implicitly or explicitly (Dabholkar et al., 1996). Furthermore, it is generally found that retail quality dimensions are often contextualized: quality dimensions and perceived importance of these dimensions tend to vary across different industries and different types of stores (Mehta et al., 2000). The quality dimensions associated with high fashion retailing (Gagliano and Hathcote, 1994) likely differ from those with grocery retailing (Martinelli and Balboni, 2012); the quality dimensions of an online store (Yang et al., 2003) likely differ from those of a conventional B&M store (Dabholkar et al., 1996); the quality dimensions of service-intensive stores likely differ from those of goods-intensive stores (Mehta et al., 2000). From this reasoning, although we suspect sporting goods stores may have some unique characteristics, we do not have a priori information or belief about how they may differ.

Situating in the wider context of retailing quality (Dabholkar et al., 1996; Simmers and Keith, 2015), the purpose of this study is therefore to empirically examine the quality dimensions of sporting goods retailing. Similar to the conventional content analysis approach (Simmers and Keith, 2015; Yang and Fang, 2004), this study draws insights about retail quality based on consumer self-generated online reviews, specifically a unique real-world dataset released by Yelp on August 1, 2018 for their round 12 dataset challenge competition.
Our specific research questions include: What are the retail quality dimensions as reflected by online reviews? How do reviews differ by ratings?

Yelp is a publicly traded company publishing crowd-sourced reviews for local businesses through yelp.com and their yelp mobile app. By the end of 2017, Yelp has 148 million reviews on their platform. Yelp is one of few online review companies that have built sophisticated machine learning models to filter fake reviews (Kamerer, 2014) and their filtration system has been found reasonably effective by third-party researchers from University of Illinois at Chicago and Google (Mukherjee et al., 2013). Additionally, it is ruled by the California Supreme Court that businesses cannot force Yelp to remove a review (Thanawala, 2018), thus reviews from dissatisfied consumers are kept intact. It can be argued that the reviews have provided a more comprehensive overview of a business. Therefore, this dataset has enough credibility to provide insights about sporting goods business.

This research makes several contributions. First, to our best knowledge, this research is the first attempt to examine retail quality of sporting goods stores. Although service quality is a prominent research area in sport management, the existing literature has primarily focused on spectator sports and fitness clubs. Sporting goods retailing sector, as a major sector of sports industry, warrants an examination. This study provides an early description of retail quality in the context of sporting goods retailing. Second, this study to certain extent cross-validates the existing retailing literature that is developed on alternative methods. Third, this research extends service quality literature by considering the interaction between online–offline channels as the online channels significantly shape consumers’ perceptions of offline service encounters. By doing so, this research updates the theory of retail service quality. Fourth, this research utilizes the text mining methodology and a large data set of self-generated online reviews. Specifically, a probabilistic topic modeling method, as implemented by Leximancer 4.0, is used to uncover the underlying structure of the reviews, which provides a more objective calibration to the existing literature. Finally, our substantive results help owners of sporting goods stores understand the important dimensions within sports retailing.

Relevant literature

Service quality

Service quality is one of the most researched topics and the literature is rich. The focus of this line of research has been primarily on the definition of quality pertaining to services in comparison with quality of physical goods. As service is thought intangible, inseparable, heterogeneous and perishable, it cannot be easily seen, felt or touched; thus the experience of it is elusive and evaluation of it subjective (Kotler and Keller, 2006; Parasuraman et al., 1985). As a result, the prevailing paradigms of service quality acknowledge that service quality is an overall judgment about service superiority. Service quality essentially is perceived quality (Oliver, 1980).

Most leading service quality scholars agree that service quality as a perception stems from some types of comparisons (Grönroos, 1984; Parasuraman et al., 1985; Zeithaml et al., 1990). The prevailing SERVQUAL paradigm, as developed by Parasuraman et al. (1985), follows Oliver (1980) disconfirmation paradigm and contends service quality is a comparison between what consumers feel the service providers should offer (i.e. customer expectations) and their perceptions of how the service deliverer performs (Parasuraman et al., 1985). The initial version of SERVQUAL includes 10 dimensions: reliability, responsiveness, competence, access, courtesy, communication, credibility, security, understanding/knowing the customer and tangible (Parasuraman et al., 1985). In comparison, the revised version of SERVQUAL includes five dimensions of service quality. They are tangibles, reliability,
responsiveness, assurance and empathy (Parasuraman et al., 1988). These dimensions are
generic in nature and are suggested to be applicable to a wide variety of services.

Alternative measures of service quality exist, and they depart from the SERVQUAL either
on the target to be compared or if a comparison is needed. For instance, Teas (1993) considers
service quality as a comparison between perception and ideals; the SERVPERF paradigm
developed by Cronin and Taylor (1992, 1994) contends that service quality is based on
perceptions alone. Furthermore, in sport management literature, Chelladurai and Chang
(2000) proposed a comprehensive evaluation model. They consider multiple standards and
contend that the standard to be used depends on the type of service being evaluated.

Inspired by the SERVQUAL and other service quality paradigms, service quality remains
a prominent research area of sport management since 1990s. In addition to Chelladurai and
Chang (2000) conceptual model, most empirical studies have focused on recreational sports
centers (Kim and Kim, 1995; Ko and Pastore, 2005; Lam et al., 2005), fitness centers (Foroughi
et al., 2019; Papadimitriou and Karteroliotis, 2000), spectator sports (McDonald et al., 1995;
Yoshida, 2017) and sport educational program (Mao and Zhang, 2012). Sancihak and Seuna-
Chana (2013) examined the impact of sponsorship on perceived service quality of sporting
goods stores and they directly applied Dabholkar et al. (1996) scale in measuring sporting
goods retailing quality. Their focus was to examine the impact of sponsorship rather than to
examine the service quality. To our knowledge, no study has focused on sporting goods
stores.

Retail quality

Because retail stores sell products, together with services implicitly or explicitly, the retail
service quality is more complex. Consumers experience, which is the origin of their quality
perception, is determined not only by in-store experiences but also experiences related to the
products. Therefore, as reviewed by Dabholkar et al. (1996) and many others, earlier studies
on retail service quality by directly applying SERVQUAL found that the five dimensions of
SERVQUAL were inadequate to capture the full gamut of quality in retailing. With an
understanding of the gap in the literature, Dabholkar et al. (1996) developed a retail service
quality scale, which would later be cited as RSQS by other researchers.

The RSQS can be considered as a modification and extension to the SERVQUAL.
Dabholkar et al. (1996) did not attempt to conceptually differentiate retail service quality from
the generic service quality. Nor were they bothered by if retailing is a service or more than a
service. Rather, efforts were made to incorporate the fact that retailers sell both products and
services; and identify additional dimensions of service quality that are unique to retailing.

The seminal hierarchical structure of retail quality identified by Dabholkar et al. (1996)
includes five dimensions: physical aspects, reliability, personal interaction, problem solving
and policy. The physical aspects in turn have two subdimensions: appearance and
convenience, encompassing items relating to appearance of the facility, store layouts,
convenience of shopping, etc. The reliability dimension resembles the reliability dimension in
the SERVQUAL; but the RSQS breaks reliability down to “keeping promises” and “doing it
right” and considers merchandise availability as a component of “doing it right”. The
personal interaction dimension also has two subdimensions: inspiring confidence and
courteous/helpful. As explained by Dabholkar et al. (1996), this dimension subsumes the
assurance and responsiveness dimensions of the SERVQUAL. Both problem solving and
policy do not have subdimensions. Problem solving is specifically reserved for handling of
returns, exchanges and complaints, which is common in retailing. Policy refers to the service
quality influenced by store policies. The policy dimension in the seminal article is a curious
one as it does not include return policy, warranty policy or other store policies. Instead, this
dimension includes merchandise quality, parking availability and so on.
In summary, in comparison with the pure SERVQUAL services, the peculiarities associated with retailing as identified by the RSQS include: a) store layouts, convenience of shopping, easiness of finding merchandise (physical aspects); b) merchandise availability (reliability); c) handling of returns, exchanges and complaints (problem solving); and d) merchandise quality, operating hours, parking, credit cards service (policy).

Both the SERVQUAL and the RSQS have inspired much applied research in retailing. Most subsequent investigations of service quality in retailing (Siu and Cheung, 2001), online retailing (Trocchia and Janda, 2003) or a subsector of retailing, such as apparel specialty stores (Gagliano and Hathcote, 1994) and grocery stores (Jain and Aggarwal, 2018; Siu and Chow, 2004) would utilize the SERVQUAL /RSQS or a modified variant of SERVQUAL/RSQS. Mehta et al. (2000) differentiated two types of retailing, service-intensive retailing and goods-intensive retailing, and found that SERVQUAL worked better for the former and RSQS excelled for the later.

With the phenomenal growth of e-commerce and online shopping, many recent studies in the retailing literature have focused on Internet retail (Janda, 2002; Nguyen et al., 2018; Trocchia and Janda, 2003; Yang and Fang, 2004; Yang et al., 2003) or aspects of service quality that are pertinent to online retailing, such as logistics and online order fulfillment (Murfield et al., 2017; Zhang et al., 2019). In the same fashion as RSQS’s extension of SERVQUAL to accommodate characteristics of retailing, these later studies expanded the literature by incorporating additional dimensions associated with information systems, which is a critical component of online retailing.

Another salient development in this literature is to examine the retail quality from the perspective of multi-channel retailing as many retails nowadays have both online and offline channels (Lin, 2012). For instance, Acquila-Natale and Iglesias-Pradas (2020) developed a comprehensive multi-channel retailing quality scale by integrating quality dimensions that are pertinent to both physical stores and online shopping. They identified in-store experience, reliability and fulfillment, service provision policies and customer service are the four retailing quality dimensions.

Content analysis, text mining and leximancer
To uncover retailing quality dimensions, content analysis approach has been utilized by some researchers (Aggarwal et al., 2009; Simmers and Keith, 2015; Yang and Fang, 2004; Yang et al., 2004; Yang et al., 2003). For instance, Yang and his colleagues have conducted a series of content analytical studies on online retailing. Based on a total of 1078 online reviews, with an average length of 65 words, of nine online companies, Yang et al. (2003) explored service quality dimensions associated with online retailing. They classified these reviews into positive and negative reviews and compared quality attributes across these reviews. Yang et al. (2004) subsequently developed an online retailing quality scale based on statements extracted from 848 reviews of online banking services. Of particular relevance to the current study is Simmers and Keith (2015). They utilized consumer comments cards to gain insights of retailing quality. Comments cards are used by many retailing stores to assess service quality in the United States. Based on a content analysis of a sample of sixty comment cards collected from diverse retail stores throughout the United States, Simmers and Keith (2015) compared the dimensions derived from comment cards and the RSQS dimensions. They found that two RSQS subdimensions – convenience (under physical aspects) and promises (under reliability) – were not present in comment cards. Additionally, they identified several new themes that were not included in the RSQS. Their new additions include friendliness and professionalism, check-out, delivery, loading, availability of service, price, selection, value, condition, usability, styling, preference and the location of store facilities. Comparing with the RSQS, Simmers and Keith (2015) had a
broader operationalization of retail quality. More store policies, including refund/exchange, delivery, were also considered.

Complementary to traditional content analysis approach, one salient advancement in marketing research methodology is the use of text mining. Text mining is a process of deriving meaningful information from a large number of unstructured texts using Natural Language Processing (NLP) technology and has been widely used for information extraction and retrieval, clustering and classification and trend detection (Berezina et al., 2016; Yu et al., 2011a). The text mining approach enables researchers to extract structured information from large datasets with objectivity (Duan et al., 2013), and capture semantical nuances that might be difficult to be differentiated by manual coding (Roberts et al., 2014). It has been suggested that text mining is epistemologically compatible with grounded theory and content analysis; it is similar to content analysis in the sense that both aim to extract common themes by counting words (Yu et al., 2011b). There are numerous applications of text mining technology in marketing research. For instance, Aggarwal et al. (2009) suggested that text mining technology can be used by retailers to monitor the positioning of their brands against their major competitors; Netzer et al. (2012) proposed an innovative method to research market structure using text mining; Humphreys and Wang (2018) and Berger et al. (2019) proposed two methodological guidelines of conducting text mining for consumer behavior and marketing scholars; Berezina et al. (2016) used text mining to understand service attributes associated with hotel industry; Chakrabarti et al. (2018) proposed a text mining approach to capture service quality of private banks in India. They found that sentimental scores of consumer reviews significantly correlated with service quality dimensions.

There are various open source programming packages and commercial software available to conduct text mining, such as gensim and Scikit-learn packages in Python, mallet in JAVA, tidytext in R, SPSS Modeler Text Analytics, SAS Text Miner, and Leximancer. Particularly, Leximancer has been extensively used in social science and business research, including sport management research. For instance, Shilbury (2012) and Anagnostopoulos and Bason (2015) used Leximancer to analyze the topics and contents of sport management–related journals; Tseng et al. (2015) used Leximancer to understand destination image based on travel blogs.

Roughly speaking, there are three types of text mining: unsupervised learning, where learning algorithm only draws inferences from input data without labeled responses (analogous to factor analysis which does not have a dependent variable); semi-supervised learning, where learning algorithm draws inferences from a small amount of labeled data with a large amount of unlabeled data; and supervised learning, where learning algorithm draws inferences from data with labeled responses (analogous to regression analysis which has at least one dependent variable) (Ang et al., 2016). Whereas most existing studies have only utilized the unsupervised learning algorithm in Leximancer, Leximancer is able to implement some supervised learning algorithm through “learning from tags”. One advantage of supervised text mining over unsupervised text mining is that it can incorporate information from the labels assigned to the texts. The “learning from tags” function in Leximancer conceptually is similar to Labeled Latent Dirichlet Allocation (Ramage et al., 2009) and supervised topic modeling (Mcauliffe and Blei, 2008). As far as we know, Leximancer is one of very few programs that have implemented this function. Utilizing Leximancer as a text mining tool, the current study attempts to identify key retail quality dimensions of sporting goods stores as reflected by yelp reviews. Furthermore, drawing upon the “learning from tags” function, this study also explores the relationships among these identified quality dimensions with consumers’ summative judgment (i.e. the star ratings).
Methods

Yelp review data
This study utilizes the Yelp reviews to understand the service quality dimensions pertaining to sporting goods retailing under the modern marketing milieu. The Yelp dataset, released by yelp.com in August 2018, includes five json files, 7.3 Gigabyte in total, containing information about businesses, reviews, users, check-ins and tips respectively. Of relevance to this study are the businesses and reviews files. The businesses file is used to filter out the relevant sporting goods stores based on the business category (i.e. primary business activities). Yelp has over 1200 categories of businesses, which were extracted based on consumers search. Whereas Yelp allows a maximum of three categories to be listed for any business, all businesses included in this study contain “sporting goods” in their category list. We identified 1481 sporting goods stores, including both full-line general stores and specialty stores. These stores operate or have operated in 320 zip-codes in 7 states, namely Arizona, Illinois, North Carolina, Nevada, Ohio, Pennsylvania and Wisconsin.

After the identification of the stores, their respective reviews can be identified by linking the businesses file to the reviews file through unique business_id variable in the dataset. Subsequently, 27,793 reviews with a total of 3,080,054 words were identified. The dates of the reviews spanned from April 29, 2005 to July 2, 2018. The length of reviews ranged from 1 to 978 words, with a mean length of 111 words and standard deviation of 100 words. In addition to the text, the star rating of each review is also recorded. Table 1 further presents the word counts of reviews grouped by star ratings.

Yelp did not report the demographic information of the reviewers, but they were a part of the users. Comparing with the general US population, yelp users on average are younger, wealthier and better educated. Plausibly, the demographics of yelp users are what most sporting goods stores target at. In the United States, 33% of yelp users are 18–34 years old; 35% are 35–54 years old, and 32% are 55+ years old. Regarding educational level, 64% of yelp users have some college education, 18% have grad school education and 18% are below college education. Regarding income, 51% of yelp users have an income of greater than 100 thousand USD, 24% users have an income between 60 and 99 thousand USD and 25% have an income below 59 thousand USD (Bennett et al., 2003).

Leximancer analysis
The current study used Leximancer 4.0, a commercial software implementing a patented probabilistic topic modeling algorithm largely based on Naïve Bayes classifier with improvements for seeded classifier approach (Salton, 1989; Smith, 2003). The 27,793 reviews are subsequently analyzed by Leximancer 4.0. Leximancer identifies themes in a text based on the occurrences and co-occurrences of words or text segments. Roughly speaking, an analysis using Leximancer entails the following stages: text preprocessing, automatic concept seed identification, concept seeds editing, thesaurus identification, compound concept editing and concept map and report generation (Nunez-Mir et al., 2016).

<table>
<thead>
<tr>
<th>Star ratings</th>
<th>N</th>
<th>Percentage</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>4,016</td>
<td>14.45</td>
<td>150</td>
<td>137</td>
<td>6</td>
<td>108</td>
<td>978</td>
</tr>
<tr>
<td>Two</td>
<td>1,517</td>
<td>5.46</td>
<td>149</td>
<td>120</td>
<td>9</td>
<td>118</td>
<td>905</td>
</tr>
<tr>
<td>Three</td>
<td>1,712</td>
<td>6.16</td>
<td>131</td>
<td>106</td>
<td>3</td>
<td>103</td>
<td>943</td>
</tr>
<tr>
<td>Four</td>
<td>4,448</td>
<td>16.00</td>
<td>118</td>
<td>94</td>
<td>2</td>
<td>94</td>
<td>957</td>
</tr>
<tr>
<td>Five</td>
<td>16,100</td>
<td>57.93</td>
<td>93</td>
<td>81</td>
<td>1</td>
<td>70</td>
<td>934</td>
</tr>
</tbody>
</table>

Table 1. Descriptive statistics of review length by star ratings
Preprocessing. Preprocessing text file is critical to any Natural Language Processing algorithms. One advantage of Leximancer is that the software can almost automate the critical data preprocessing step with minimal researchers' input. Therefore, the unpreprocessed reviews were supplied to Leximancer.

Concept Editing. In Leximancer, there are three key terms: theme, concept and thesaurus. Themes are clusters of concepts that are selected based on co-occurrences; concepts in turn are defined by a thesaurus of words. Leximancer firstly generates concepts and themes before it identifies the relevant thesaurus of words. After preprocessing, Leximancer can quickly identify the concepts and themes in an exploratory fashion based on frequency and co-occurrences of words. Automatic concept seeds are generated based on a ranked list of these words. If desirable, conceptual seeds can also be manually edited by researchers. For instance, as reviewers often mention about store names in the reviews, we combine all the names into one concept and label it as “storeName”; this also applies to city names (cityName), brand names (brandName), staff names (staffName). The reviewers may also mention about a specific product in the reviews, which may not be of central interest to our research. We may merge all these different product names using one concept “product”. In this study, we used camel case words (e.g. storeName) to indicate user-defined seeds.

Compound Concepts. Although Leximancer can automatically extract some multiple-words proper names and treat them as one concept, the software relies on the capitalization of the words as a hint of proper names (Leximancer, 2011). Some compound concepts, such as “customer service”, were not able to be automatically identified. After the generation of thesaurus of words, researchers can manually add compound concepts. For this study, we added one compound concept “customer service”.

Mapping Concepts. Finally, Leximancer can select concepts and tags to map in the final report generation stage. Tags are special names assigned to a folder, a file, a variable, patterns, non-meaningful texts. By default, words associated with tags will be excluded from the machine learning process, unless the program is instructed to “learn from tags”. In this study, we instructed the program to learn the star rating tags. Furthermore, to be explained in the discussion section, we assigned “gun(s)”, “bike(s)”, and “bicycle(s)” as killed concepts. The final output of Leximancer provides an interactive conceptual map that visually represents the main themes and concepts extracted from the text mining process and contains information about how they are associated.

Results
Validity of extracted information
An analysis of the 27,793 reviews about 1481 stores that sell sporting goods produced a concept map consisting of a maximum of 81 concepts. Before we present the substantive themes and concepts, we attempt to assess the face validity of the extracted information. Text mining is a relatively new method to academic research; as a result, the standards for assessing its validity have not yet fully established. Following the validity types proposed by Krippendorff (2004), Dr. Andrew Smith, the inventor of Leximancer, assessed face validity, stability, reproducibility, correlative validity and functional validity of Leximancer and found the tool reliable and valid (Smith and Humphreys, 2006).

Humphreys and Wang (2018) suggested that three strategies – comparison, correlation and prediction – may be used to assess various types of validity, which would require the document-term matrix or collecting additional external non-textual variables. As a proprietary software, Leximancer does not disclose the exact formula and algorithms used for the analyses; neither does it provide the conventional document-term matrix as far as we know. Therefore, the ability of conducting further analysis is restricted due to the lack of document-term matrix, and we can only conduct some descriptive statistics based on the...
available output. In the spirit of Humphreys and Wang (2018) correlation strategy, we examined the correspondence of the extracted sentimental words and star ratings to assess how well the extracted concepts have preserved information in the raw reviews. It is intuitive to hypothesize that reviews with higher ratings will contain more positive sentimental words than those with lower ratings. The results of our analysis, as shown in Table 2, suggest that higher ratings are indeed accompanied by more positive sentimental words and fewer negative sentimental words. This pattern is largely consistent across all the sentimental words identified Leximancer, which lends credibility to the information extracted by the text mining approach.

Emergent themes and concepts

Figure 1 presents the concept map with theme size at 50% and visible concepts at 100%, which serves as the foundation for further analysis. The Leximancer concept map is interactive in nature: different clusters of concepts (i.e. themes) can be generated by adjusting the “% Theme Size” sliding bar. With theme size at 100%, only one theme, “customer” is produced; with theme size at 0%, each individual concept becomes a theme. Furthermore, although some of the identified concepts were deemed to be generic and less diagnostic (e.g. “feel”, “nice”, “area”), there is nothing to lose by keeping them for potential exploration. Hence, we decided to keep “% Visible Concepts” at 100%.

Five themes, “customer”, “staff”, “prices”, “store” and “return”, with a connectivity rate of 100%, 90%, 71%, 59% and 34% respectively, were identified as dominant by Leximancer. Connectivity, indicating the relative importance of the themes, is calculated based on the co-occurrence counts of the concept with every other concept on the map (Leximancer, 2011).

Customer. The strongest theme, “customer”, is the most important theme. Out of the 12,962 counts of “customer”, 11,021 are paired with “service”, making it a compound theme, “customer service”. This theme, including concepts like “experience”, “care”, “business”, “review”, “purchased”, “owner”, “time”, “local”, “life”, relates to the holistic judgment of service experience and customer care.

Here are some quotations of this theme:

... The customer service is ABSOLUTELY TERRIBLE in all of my experiences. Today i went in to buy some balls and 5 of them were just sitting around not doing anything ...

<table>
<thead>
<tr>
<th>Concept</th>
<th>One-star</th>
<th>Two-star</th>
<th>Three-star</th>
<th>Four-star</th>
<th>Five-star</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAVORABLE</td>
<td>0.65</td>
<td>1.06</td>
<td>1.58</td>
<td>2.03</td>
<td>2.38</td>
</tr>
<tr>
<td>nice</td>
<td>0.08</td>
<td>0.18</td>
<td>0.26</td>
<td>0.27</td>
<td>0.21</td>
</tr>
<tr>
<td>love</td>
<td>0.04</td>
<td>0.09</td>
<td>0.12</td>
<td>0.21</td>
<td>0.23</td>
</tr>
<tr>
<td>amazing</td>
<td>0.02</td>
<td>0.01</td>
<td>0.05</td>
<td>0.05</td>
<td>0.15</td>
</tr>
<tr>
<td>cool</td>
<td>0.03</td>
<td>0.05</td>
<td>0.09</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td>fun</td>
<td>0.02</td>
<td>0.02</td>
<td>0.07</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>UNFAVORABLE</td>
<td>1.07</td>
<td>0.87</td>
<td>0.57</td>
<td>0.35</td>
<td>0.31</td>
</tr>
<tr>
<td>bad</td>
<td>0.14</td>
<td>0.15</td>
<td>0.12</td>
<td>0.09</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Note(s): FAVORABLE, UNFAVORABLE are not words used in the reviews. Instead, they are collections of sentimental words automatically detected by Leximancer. Words in lower cases are the exact words used in the reviews. As reviews of different star ratings on average have different lengths, the frequency was scaled by length of the review using 150, the average length of one-star review, as the baseline. The scaling coefficients for two-star, three-star, four-star and five-star reviews were 150/149, 150/131, 150/118, and 150/93 (see Table 1). The numbers in the table mean that given the level of star rating, how likely these sentimental words would occur in the review text.

Table 2. Occurrence frequencies of sentimental words by star ratings
...he had no intention to service me, he did not care if I left unpleased, there was no sense of urgency in the atmosphere to leave any customer with a lasting impression. A good lasting impression at that. Customer service once meant meeting the needs of the customer by any means, especially when you're more than capable of doing so. . .

**Staff.** The second strongest theme, “staff” (merged with “employee(s)”) has a word count of 14,442, which corresponds to the Sales Staff dimension by Simmers and Keith (2015). The most important concepts under this theme are “helpful”, “friendly” and “knowledgeable”; and the theme connects to “needed” (merged with “need(s)”), “recommend”, and “staffName”. This following quote highlights the quality of sales staff expected by consumers and positive behavioral consequences associated with high quality of sales staff:

...We were at [storeName] the day before and left empty handed because the girl who worked in the snowboarding dept. had no idea what she was talking about. [Another storeName] was the exact opposite. A very helpful and informative guy by the name of [staffName] answered all of our questions. He gave us his personal advice over what size board would be right for me . . . After speaking with him I felt confident in the purchase I was making, and so happy that I ended by a new

---

**Figure 1.**
Leximancer conceptual map with **Visual Concepts** at 100% and **Theme Size** at 50%. The Leximancer output is interactive as the users can decide the number of concepts and the number of themes to be displayed by adjusting the sliding bars (% Visible Concepts and % Theme Size, respectively) below the conceptual map. The themes are displayed in a conventional heat map scheme, with the most important theme displaying in red, followed by orange, green and so on; the size of the circles has no bearing as to its importance in the text (Leximancer, 2011). The words in red starting with “stars_” are the ratings of the businesses and were treated as “tags” by Leximancer. The concepts that cluster near a tag indicate that they are more specific to the tag.

The black line in the middle, which is not a Leximancer output, is manually added by the researcher. Additionally, the researcher also added the names of themes to the original Leximancer output.
pair of pants that I did not really even need. I am just sooooooo impressed by ... the great customer service that was provided.

*Price.* The third strongest theme, “price”, with a word count of 9170, is really one subdimension of “product” which has a word count of 7106. This theme also aligns well with the product dimension identified by Simmers and Keith (2015). Apparently, “price” is the most important concept under this theme. Other highly connected concepts are “selection”, “quality”, “inventory” / “stores” (as a verb, merging with “storing”). Here are a few review excerpts that highlight the product dimension:

...Stopped by today just to check out the store, and I walked out with a new backpacking tent. Prices are very reasonable...Not a huge selection, but they definitely have quality products.

...This place has all the snow and skateboard gear one might need. Great priced helmets for skating and skiing, great penny board selection.” “The [brandName], [brandName], [brandName] and [brandName] shoe selection is insane. Seriously, blows [competitorStore] out of the water. ... kids and adult sizes! And [brandName] are hard to find...my husband loves them ...

*Store.* The fourth theme, “store”, with a word count of 12,851, includes the store facilities dimension by Simmers and Keith (2015) and also overlaps with the service dimension by Simmers and Keith (2015) and policy dimension by Dabholkar et al. (1996). The critical concepts associated with this theme include “sales”, “location”, “items”, “size”, “line” (i.e. checkout line), “order”, “brandName”. More interestingly, this theme is closely related to “online” and “competitorStores”. This theme is also associated with “holidayName” as most stores have sales promotion during holiday seasons. The following are a few reviews that are representative of this dimension:

...The rest of the store I do not mind. It’s well organized and occasionally you can find good sales.

...I really like this [storeId] location ... Great location within the Oakville downtown shops. Only downside is hard to find street parking, and you can only park for 1 hour.

...The only thing this [storeId] store has going for it is the convenience of its location.

*Return.* The last theme, “return”, with a word count of 2931, is closely related to “order”, “trying”, “money”, “manager”, “bad”, “problem”, “perceivedValue”, “online”, “online order”, and time related concepts (i.e. “minutes”, “week”, “days”). Given the fact that this theme is highly associated with one-star reviews, an examination of the reviews indicates that this theme is primarily about problem handling, order returning and re-patronage intention.

Further examination of the concept map also reveals that this theme is split by two relatively distinct clusters of concepts. The upper left sphere of this theme, where adjoins the “customer” theme, clusters concepts of “manager”, “bad”, “problem”; and the lower right sphere of this theme, where adjoins the “store” theme, clusters concepts of “order”, “home”, “online”, “trying”, “online order”. The former, generally relating to the policy and problem-solving dimensions identified by Dabholkar et al. (1996), has highlighted the role of store managers. The reviews reveal that a lack of service orientation from the management level often leads to poor service and negative behavioral consequences. The latter, which has not been documented in the existing literature, reveals how the presence of online channels interacts with the perceived offline retail quality. Given the novelty of this finding, we decide to separate the return theme into “manager” and “online”.

*Online.* In this section we further explore the “online” concept, which has a word count of 1906. Comparing with online stores, better customer service is the reason that consumers might remain loyal to offline channels, which is reflected in many of the reviews. It is also evident that service quality, as consumers’ perception, has been impacted by the online channels in multiple ways. First, many consumers have higher expectation about service
quality in the offline channels, due to the threats of the online channels. The consumers feel they are empowered to switch channels and thus entitled to receive good services at the offline channel. Second, consumers expect that the prices at the online channels to be lower. They constantly compare the prices at the offline stores with online channels. Most sporting goods are shopping goods as they are typically valuable, consumers often have significant shopping involvement, such as quality inspection and price comparisons. Consumers will be pleasantly surprised when they find out that the offline channel has competitive prices vis-à-vis the online channels.

However, with the easiness of price comparison enabled by smartphone and price checker apps, the online channels have pushed the offline channels to bundle their products with creative value-adding services. Branded variants (e.g. exclusive products; Bergen et al., 1996; Shugan, 1983), in-house brands (i.e. private labels) and service bundles are effective tools to avoid fierce price competition. However, smaller store may not have the bargaining power with the manufacturers to create their own product variants or private products, only the service bundle may be feasible. This again highlights the importance of customer service.

Third, although most transactions of sporting goods occur in the offline channels, the e-commerce is on the rise. Leading sporting goods chains are prioritizing their investment on online channels (Dick’s Sporting Goods, 2018). Many sporting goods stores have their own online stores; they may have “buy online pick up in store” service; they may have “in store preorder” service if they do not have items in stock; they may have an “online stock check” service. To compete with dominant online channels, going online is another strategic move for most sporting goods stores. However, the integration of the online channels and offline channel becomes a new challenge in running a sporting goods store, as evident in the following reviews.

Returning an online order to an offline channel:

...Ordered a pair of shoes online and they did not fit brought them to a [brandName] corporate store which they are listed on [brandName] website as and the brilliant manager proceeds to tell me he has never encounter a return like this and do not know how to do it or even go about finding a resolution to my problem...

Online inventory check:

...I called store based on online inventory indicating that the store had 1 pair of the gloves that I was looking for... After driving the 30 miles to this location they did not have anything on hold for me...

In store ordering:

...So because their inventory was so messed up, we proceeded to check out... we were buying 2 pairs of shoes from the store, then 3 pairs from online since sizes were not available in the store.... This was a 30 minute + process and was A TOTAL WASTE OF TIME!

Through this analysis, it becomes apparent that online channels have exerted significant influences on consumers’ evaluation of retail quality. More importantly, an effective integration of online–offline channels has become a new dimension of retail quality (Chiu et al., 2011; Ofek et al., 2011).

Retail quality of sporting goods stores
As indicated in Figure 1, some identified circles overlap with each other, whereas others do not. The overlapped areas suggest that the themes are semantically related to each other; the nonoverlapping on the hand suggests distinctiveness of the themes. There are some
interesting observations by examining the overlapped themes and concepts: “love” is at the intersection of “staff” and “product”; “online” and “perceivedValue” are at the intersection of “store” and “return”; “problem” and “review” are at the intersection of “customer” and “return”; “product”, “inventory” and “deal” are at the intersection of “prices” and “store”. Furthermore, we can hypothetically draw a line to separate the five themes into two (Figure 1): the upper left being customer service and the lower right being store aspects, which more or less is consistent with the existing conceptualization of retailing quality (Westbrook, 1981). The current analysis, however, provides not only empirical evidence for the relative separation of these two themes but also insights about where these two themes interact. It seems that there are two main interaction areas (i.e. overlapping areas): the first being “staff” and “product”, and the second “manager” and “online”. Additionally, this separation can be confirmed by adjusting the theme size of the mind map to 97%, when only two themes would emerge.

Although the themes and concepts identified by Leximancer have provided valuable insights about quality of sporting goods retailing, they are indicative in nature. Based on the above analysis, we re-name the themes to better reflect the meanings of the concepts and re-cluster them in a hierarchical structure. Table 3 presents the final conceptual model of Retail Quality of Sporting Goods Stores (RQSGS). On the very top level, customer service and store aspects serve as the two primary dimensions of RQSGS. Under customer service, three main themes, including holistic experience, manager and staff have emerged. Holistic experience focuses mainly on customer care and shopping experience. Manager highlights the role of managers in elevating retail quality; problem handling, terms of sales, and a prompt refunding policy have been particularly pronounced in this dimension. Staff are the frontline service individuals who typically provide direct customer service to consumers. Helpfulness, knowledge and friendliness are the three major positive qualities that are valued by customers. Under store aspects, three main themes, including product, B&M store and online–offline integration, have emerged. Regarding product, consumers care about price, selection, quality and brand. Regarding B&M store, consumers care about inventory,
location, sales promotion and perceived value. Finally, online–offline integration (Nguyen et al., 2018) becomes a salient dimension in RQSGS.

RQSGS vis-à-vis star ratings

Based on the relative positions of the star rating tags in Figure 1 and the numeric representation of conditional probability of word occurrence in Table 4, it is evident that different star ratings have focused on different aspects of retail quality. A general conclusion is that one-star and five-star reviews focus more on customer service, whereas two- to four-star reviews focus more on store aspects. Figure 2 presents a heat map of the occurrences of concepts by normalizing the data points in Table 4. Green color on the figure indicates more likely to occur and red color least likely to occur. By grouping concepts into their respective dimensions, Figure 3 presents the probability of dimensions by star ratings. It is evident that from one-star to five-star ratings, some dimensions are monotonically increasing (i.e. Staff), some monotonically decreasing (i.e. Manager), some peaks at the middle (i.e. Product and B&M Store) and others peak at the extremes (i.e. Holistic Experience).

Customer Service. Holistic experience, as indicated by “care” and “experience”, is mentioned more often in one-star and two-star reviews, followed by five-star reviews, three-star and four-star reviews do not discuss this theme as often. Regarding positive characteristics of staff, as indicated by “helpful”, “knowledgeable” and “friendly”, we can roughly observe a positive correlation between star ratings and frequency of word occurrences.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Counts</th>
<th>One-star</th>
<th>Two-star</th>
<th>Three-star</th>
<th>Four-star</th>
<th>Five-star</th>
</tr>
</thead>
<tbody>
<tr>
<td>Care</td>
<td>1970</td>
<td>10</td>
<td>9</td>
<td>6</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>experience</td>
<td>4856</td>
<td>21</td>
<td>19</td>
<td>15</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>helpful</td>
<td>5514</td>
<td>6</td>
<td>11</td>
<td>17</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>knowledgeable</td>
<td>4674</td>
<td>4</td>
<td>9</td>
<td>9</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>friendly</td>
<td>7157</td>
<td>13</td>
<td>19</td>
<td>21</td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>manager</td>
<td>1805</td>
<td>20</td>
<td>12</td>
<td>7</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>owner</td>
<td>2300</td>
<td>14</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>problem</td>
<td>1834</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>money</td>
<td>2227</td>
<td>20</td>
<td>14</td>
<td>9</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>return</td>
<td>2937</td>
<td>28</td>
<td>20</td>
<td>9</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>week</td>
<td>2464</td>
<td>17</td>
<td>15</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>days</td>
<td>1791</td>
<td>13</td>
<td>10</td>
<td>7</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>minutes</td>
<td>2724</td>
<td>20</td>
<td>18</td>
<td>10</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>prices</td>
<td>9170</td>
<td>29</td>
<td>41</td>
<td>49</td>
<td>42</td>
<td>29</td>
</tr>
<tr>
<td>selection</td>
<td>4936</td>
<td>6</td>
<td>21</td>
<td>35</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>quality</td>
<td>2092</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Brand</td>
<td>1842</td>
<td>6</td>
<td>9</td>
<td>11</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>brandName</td>
<td>936</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Inventory</td>
<td>684</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Location</td>
<td>3293</td>
<td>13</td>
<td>17</td>
<td>21</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>holidayName</td>
<td>432</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Sales</td>
<td>4116</td>
<td>22</td>
<td>24</td>
<td>25</td>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td>competitorStores</td>
<td>479</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>perceivedValue</td>
<td>436</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Online</td>
<td>1906</td>
<td>11</td>
<td>12</td>
<td>10</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>

Note(s): The numbers on the columns 3–7 mean the frequencies of word occurrences. For instance, in each 100 one-star reviews, the concept “care” would occur 10 times; whereas in each 100 four-star reviews, the same concept would only occur four times. Whereas some concepts are distributed relatively evenly among all groups, other concepts are more or less prominent with certain group. For instance, “knowledgeable” and “sales” with five-star reviews; “manager” with one-star reviews.

Table 4. Frequencies of concept occurrences in every 100 reviews
occurrences under this theme. Five-star and four-star reviews are strongly related to the positive characteristics of staff. Whereas four-star reviews are more likely to use “helpful” and “friendly” than five-star reviews; five-star reviews are more likely to use “knowledgeable”. “Knowledgeable” is one the two most diagnostic terms (the other being “sales”) that separates five-star reviews from the rest. With regard to the manager dimension, we observe a roughly monotonic decreasing occurrence of such words as “manager”, “owner”, “problem”, “money”, “return”, “week”, “days”, and “minutes”, from one-star to five-star reviews. One-star and two-star reviews tend to focus much more on this dimension than the rest.

Store Aspects. Regarding price, selection, quality and brand of the product dimension, we observe that three-star and four-star reviews tend to discuss most often, followed by two-star reviews; one-star and five-star reviews tend to discuss least. The same phenomenon holds for the location, promotion aspects of the B&M store dimension. Furthermore, five-star reviews tend to discuss “sales” the least. Three-star and four-star reviews are strongly related to the B&M store dimension. Particularly, these reviews are closely related to “competitorStore” and “online” concepts. In those reviews, a comparison with competing stores, such as Amazon and Walmart, or other online channels is often mentioned, indicating an assessment of quality is anchored with competing stores and influenced by online channels. Two-star and one-star reviews are strongly related the manager dimension. Essentially, the lack of service from the management level leads to order returning behavior and negative repatronage intentions. Value-laden words, such as “worth” and “value”, are often mentioned in these reviews. “Online” again is a closely connected concept, indicating impacts of online channels.
Sporting goods retailing is a significant supporting sector of the sports industry (Eschenfelder and Li, 2007). The main objective of this research is to investigate the retail quality of sporting goods stores using a relatively novel approach – text mining. Drawing upon themes and concepts extracted by Leximancer on voluminous consumer self-generated online reviews, this study identifies key quality dimensions within sporting goods retailing. We subsequently re-theme and re-cluster the Leximancer map and propose a model of Retail Quality of Sporting Goods Stores (RQSGS). Consistent with the existing retail service quality literature (Dabholkar et al., 1996; Simmers and Keith, 2015), this model takes a broad perspective of retail quality and identifies customer service and store aspects as two higher-order dimensions within RQSGS. Customer service includes holistic experience, manager and staff; store aspects include product, B&M store and online–offline integration. These two dimensions are relatively distinct, but they also interact through store, manager, staff and product. Furthermore, this study explores the relative importance of each concept to different levels of ratings. To our knowledge, this is also the first study examining service quality in the context of sporting goods retailing.

**Theoretical and practical implications**

The current study to certain extent cross-validated the existing retailing literature that is developed on alternative methods (Acquila-Natale and Iglesias-Pradas, 2020; Dabholkar et al., 1996; Simmers and Keith, 2015): most themes identified in RQSGS are consistent with the existing literature. Whereas the emerging multi-channel retailing quality literature has focused on the integration of consumer experiences in both physical stores and online stores.
(Acquila-Natale and Iglesias-Pradas, 2020; Murfield et al., 2017; Zhang et al., 2019), it does not fit the sporting goods retailing completely as sporting goods stores rely on revenues from physical stores. The current study in the context of sporting goods retailing revealed several points that had not been discussed in the existing literature. We focus our discussion on these findings.

First, knowledgeable staff is ultimate important to create satisfied consumers in sport retailing. Staff has been the focus of service quality literature: four out of the five SERVQUAL dimensions, namely reliability, responsiveness, assurance and empathy, are staff related. In retailing literature, the importance of sales staff has also been highlighted (Simmers and Keith, 2015). However, the existing literature often considers only the sales staff. In the current context, we do not limit staff to sales staff. Instead, staff are the frontline employees who help with the sales and/or provide other value-added services (e.g. a biomechanics specialist in a specialty running shoe store, a club repairer in a golf pro shop). Whereas three positive characteristics of the staff – knowledge, friendliness, and helpfulness – were identified important in sporting goods retailing business, “knowledgeable”, however, is one the most diagnostic concepts that separate five-star reviews from four-star reviews.

Second, managers play a critical role in elevating consumers’ satisfaction. Manager, loosely referring to both owners and professional managers, has three subdimensions: problem handling, terms of sales and promptness. Whereas the role of frontline staff has been acknowledged in the literature, the role of manager is less clear. Our results suggest that consumers tend to blame managers/owners for poor service and experience. It might be due to the fact that many sporting goods stores are small local businesses. Traditionally, sporting goods retailers are merchandise movers and are typically low in service orientation (Garg and Chan, 1997). Faced with severe competition and easiness of price comparisons due to national branding, sporting goods stores are urged to differentiate themselves from their local and online competitors. Among exclusive distribution, branded variants, private labels and service bundling, augmenting merchandises with value-adding services become one of very few viable approaches sporting goods stores might take (Homburg et al., 2002). Therefore, the demarcation of merchandise and service has become increasingly blurred. Sporting goods stores are no longer in the merchandise mover business; rather, they are shifting towards service business. As a result, managers and owners’ service orientation must be emphasized in the current marketing landscape (Garg and Chan, 1997; Homburg et al., 2002). Store policies, such as return/exchange, warranty, credit cards, installation, are the materialization of manager’s service orientation philosophy.

Third, online–offline integration is a new dimension of retail quality. Consistent with the existing literature, the online channels have considerable impacts on consumers’ shopping behavior. Consumers rely on the online channels to find out product information, stocking information, store information, and pricing information; or to place orders online and then pick them up in store. As a result, traditional B&M stores have started to expand their online presence. However, the lack of effective integration between the online and offline channels has created much consumers’ dismay. An effective integration of the channels is likely an important area of investment for sporting goods stores.

Finally, hedonic and social values are important to create superior satisfaction. Ostensibly, from our data, extreme reviews (i.e. one-star and five-star) tend to focus on customer service and mediocre reviews (i.e. two-star, three-star and four-star reviews) tend to focus on store aspects. Fundamentally, we argue that different value dimensions have played a role in consumers’ satisfaction. Rintamäki (2006) has decomposed values of shopping into utilitarian, hedonic and social values. Utilitarian values focus on monetary savings, convenience of shopping; hedonic values focus on emotional, entertainment aspects of shopping; and social values focus on the symbolism and relationships. It is evident from the review data that utilitarian value is the bedrock of shopping experience, but is in itself often
unable to produce ultimate satisfaction. Verhoef et al. (2009) suggest social environment, service interface, retail atmosphere, assortment, price, customer experiences in alternative channels and retail brand are the major drivers of customer experiences. In the context of sport retailing, there are still rooms to grow consumers’ experience by utilizing the drivers mentioned in Verhoef et al. (2009).

It should also be noted that sporting goods retailing only has subtle differences with general retailing. These differences are not only nuanced but also may be shared by other specific retailing contexts. For instance, sporting goods are mostly shopping goods therefore free riding is more likely; grocery retailing typically does not share this characteristic but electronics retailing might; in the grocery retailing case, friendly staff might be more important than knowledgeable staff but in the latter case the opposite might be true. Piano is often considered as shopping goods, but piano retailing might have much less free riding due to limited distribution, therefore online–offline integration might be less an issue. Many sporting goods stores are small in scale and cater to local markets, therefore managers and owners may play a more significant role in delivering services. Sporting goods stores are often augmenting their products with ancillary services such as gait analysis, repair services, renting services or training classes, therefore the definition of staff may be much broader than other retailing contexts. Sporting goods are mostly national/international brands that are distributed across many channels therefore price competition is severe in comparison with other retailing such as high fashion. Nonetheless, it is almost always possible to find another context that will share the same characteristics with sports retailing.

Limitations and future research

One challenge and limitation of this research is the identification of sporting goods stores. Although the NAICS has provided a straightforward definition of sporting goods stores, from market definition point of view, limiting stores to NAICS 451110 may misidentify true players in this industry, thus inappropriate (Shugan, 2011). For instance, both Walmart and Amazon do not belong to NAICS 451110. As a compromise, this study relies on stores’ self-identification. We include all the stores who claim they sell sporting goods. On the Yelp platform, many gun shops and shooting ranges are listed as sporting goods stores; granted, shops sell sporting guns are NAICS 451110 per the NAICS. Also, in the dataset, we have a large proportion of bike shops which both sell and repair bikes. We even included a small number of vision centers that may sell eyeglasses for sports use. As Leximancer extract concepts based on frequency, the large amount of gun shops and bike shops we have in the dataset could have biased the findings. Hence, we excluded some bike and gun shops by assigning “gun(s)” and “bike(s)” “bicycle(s)” as killed concepts. Essentially, we underweight concepts associated with bike and gun shops. Due to the lack of a precise definition of sporting goods stores, the applicability of RQSGS to certain sporting goods stores warrants further research.

A second challenge of this research is the representativeness of Yelp reviewers. We do not believe Yelp reviewers are a representative sample of all consumers. However, as the unit of analysis is individual review instead of individual reviewer, at the issue really is the representativeness of the reviews; and the exact nature of this concern cannot be empirically assessed. We do have confidence in our findings: as in many qualitative studies, a saturation of themes can often be achieved with a small number of participants. Given the large amount of reviews we have, the topics that are covered by the current study are likely more extensive than the traditional methods. The chance of omission will be smaller. This approach essentially has the same spirit as the “maximum variance sampling” (aka heterogeneous sampling) technique: the goal is not to be representative of the average views on an issue but to look at it from all angles. The key empirical task is to determine the weights of these
concepts. Although a closer examination of rare concepts will be of interest in our future research, it is beyond the scope of the current project.

The generalizability of RQSGS beyond the sporting goods retailing industry is a secondary consideration of this project. We do not intend to claim the framework works for other retailing businesses, as different sectors likely have their own salient characteristics. However, we do believe that the framework provides a supplementary foundation for other retailing contexts; and the RQSGS may be generalizable to other retailing businesses that share similar characteristics with sporting goods retailing (such as music instruments stores).

This research was exploratory with the goal of producing a better understanding of service quality in sporting goods retailing. We neither examine service quality relating to specific types of stores nor quantify the relative importance of each dimension. Lastly, validating the findings from text mining remains a promising task. Cross-validating findings based on Humphreys and Wang (2018) strategy can be fruitful. Future studies may address these topics.

References


Eschenfelder, M.J. and Li, M. (2007), *Economics of Sport*, Morgantown, Fitness Information Technology, WV.


Grönroos, C.J.E.J.o. m. (1984), “A service quality model and its marketing implications”, Vol. 18 No. 4, pp. 36-44.


**Corresponding author**

Luke Lunhua Mao can be contacted at: lmao@unm.edu