

Valeria Franceschi

China as an International Travel Destination

A Corpus-based Analysis of Online Travel Blogs

Abstract

User-generated content in tourism has increasingly become an important source of data for multiple aspects of tourism research, including destination image, electronic Word-of-Mouth (eWoM), and destination marketing. Such online material as travel blogs, guides, video blogs and reviews may influence readers and prospective visitors, shaping their gaze on the destination and informing their travel choices. User-generated travel blogs and guides indeed provide a combination of knowledge and information as well as emotions, and are often deemed more reliable than professionally-produced tourist materials. This paper aims at investigating the tourist gaze on China and its population as emerging from travel blogs and travel guides produced by non-professional internet users. Specifically, it wants to determine how popular Chinese destinations and local people are perceived through the eyes of international tourists. To this purpose, a corpus of blogs on travel and life in China was compiled using the text mining software BootCat and then analyzed with the Sketch Engine corpus analysis tool. A quantitative and qualitative mixed method approach was adopted for this study, analyzing keywords and collocational profiles in addition to carrying out content analysis on concordance samples to shed light on how visitors gaze upon major Chinese destinations. Results indicate that visitors have an overall positive perception of China and its people but describe the country in contrasting terms, as seen in previous studies on this destination. Tourists appear to look for elements signifying authenticity, such as historical heritage sites and locals engaging in stereotypical activities.

Keywords: *tourism discourse, collocation analysis, content analysis, tourist gaze, tourism in China*

In the past decade China has established itself as a major player in the tourism industry “both as one of the most important generators of outbound tourism and as a significant global tourism destination” (Tseng et al. 2015, 347). In the final year before the Sars-Cov-2 pandemic that caused international tourism to grind to a halt in 2020, the Chinese National Bureau of Statistics reported 145.31 million inbound tourists from outside mainland China (including

Hong Kong, Macao and Taiwan), of which 31.88 million are foreign tourists.¹ China's global tourism competitiveness similarly increased, with its ranking rising from 71st in 2007 to 13th in 2019 (Zhao and Liu 2020, 337).

China's rising importance on the global tourist scene makes the perception of international travelers worthy of investigation, especially in the current era of user-generated content (UGC): tourist perspectives can influence destination image and, in turn, affect prospective traveler behavior in terms of destination and travel planning through electronic Word-of-Mouth (eWoM) (Ran et al. 2021; Chang and Wang 2019; Xiang et al. 2015).

This paper aims at contributing to existing research on China as an international travel destination by investigating the way visitors perceive and describe China and its people. This has been done by carrying out a linguistic analysis on a corpus of English-language travel blog entries, specifically travel accounts or travel guides to Chinese destinations. A mixed method approach combining corpus linguistics tools (keyword, collocation and concordance analysis) and content analysis has been applied in order to identify the recurring themes in the dataset and investigate how tourists perceive and reproduce their gaze (Urry 1990). This study sets itself apart from previous studies on travel blogs about China by looking at a bigger dataset, with a different approach, and a wider—and more recent—timeframe. While the notion of gaze is very closely intertwined with the visual mode and often includes photographs and images, due to the process of data retrieval, which does not include visual elements, it was not possible to carry out a multimodal analysis.

1. Literature review

1.1 *China as an international destination*

There has been no lack of academic interest in China as a tourist destination, especially in the years immediately following the Beijing Olympics. These studies focus preeminently on US citizens (e.g. Pan and Li 2011; Li et al. 2009) and investigate destination image in general, exploring how subjects who both have and have not traveled to China perceive the country as a whole or individual locations and/or attractions. A range of different approaches and data sources are employed, i.e. surveys (Becken et al. 2017; Chen, Hua and Wang 2013), TripAdvisor reviews (Wu, Wall and Pearce 2014), newspaper discourse (Xiao and Mair 2006), and blogs, as with the current research (Tseng et al. 2015; Li and Wang 2011).

¹ <https://news.cgtn.com/news/2020-02-29/Chinese-tourists-made-169-million-outbound-trips-in-2019-report-OtIYWzZmOQ/index.html>. Last visited 01/06/2021.

We will briefly focus on Li and Wang's analysis as it informs part of this study. Content analysis was carried out with Atlas.ti 6.0 on a corpus of 89 travel blog entries written by Western bloggers between February 2008 and 2009. A frequency list of nouns and verbs was classified in ten categories representing the major themes discussed in the blogs (2011, 697):

- Cities, Attractions, & Other Places
- People & Life
- Travel & Transportation
- Accommodation, Restaurants, & Food
- Entertainment & Activities
- Price & Quality
- Infrastructure
- History & Culture
- Environment & Safety
- Policies & Regulations

A separate analysis of frequent adjectives and verbs showed that “bloggers employed more positive descriptors than negative ones, implying that their perceptions of China as a tourism destination tend to be positive” (2011, 699).

Content analysis was then carried out on the entirety of the corpus through coding and categorization of the content according to image attributes and the positive/negative aspect of the attributes (2011, 700).

Their work, alongside previous research focusing on people who had already visited the country, seems to highlight the construction of a mixed image of China. They suggested a positive perception of China as a tourism destination but more negative as a country overall (2011, 705). Other studies suggested that tourist perceptions were informed by stereotypes and influenced by a Western, ‘Orientalist’ perspective: Xiao and Mair’s analysis of newspaper travel writing identified a “sustaining image of paradox” (2006, 10) in the representation of China, most notably in the “perceptions of the changing versus the unchanged” (2006, 8), or the modern versus the primitive, with contrasting images of China and its people being highlighted in the dataset. Tseng et al.’s study of travel blogs similarly underlined contradictions between old and new in China, with bloggers struck by “a fascinating mixture of time periods that can be uniquely experienced in China” (2015, 355). A common trend across research seems to underline cultural heritage and historical sites as the prominent positive aspects in perceptions of China.

For example, Li and Wang state that “[t]he most popular attractions were related to the cultural and historical aspects of China’s image (e.g. Great Wall, Tiananmen Square, Terra Cotta), and the natural features of China (e.g. river, mountain)” (2011, 705). Other recurring positive aspects include friendliness, accommodation, and restaurants (Tseng et al. 2015; Chen, Hua and Wang 2013; Xiao and Mair 2006).

Similarities across studies were also identified in the negative aspects emerging from the data, with political stability, environment and cleanliness, accessibility and transportation often mentioned as drawbacks during trips, or as factors that may steer prospective travelers away from selecting China as a destination. For instance, Becken et al. (2017) focused specifically on the impact of poor air quality on destination image and subsequent behavior in travel choices: survey results showed that air quality would indeed influence respondents’ decision to visit China (2017, 142).

1.2 The tourist gaze

As detailed above, previous research on tourism in China has been carried out prevalently on destination image or country image. Both concepts are extremely relevant to the tourist industry, as the former especially can inform people’s destination choices and travel planning. However, we should consider that destination image is informed, among other elements, by what Urry (1990) defines as the ‘tourist gaze,’ that is, the way that people frame and understand their tourist experience. Gazing is not a passive experience, but rather, it is a “performance that orders, shapes and classifies, rather than reflects the world” (Urry and Larsen 2011, 2). In D’Egidio’s words, the tourist gaze is “one way of understanding the experiential elements of tourism and [it] also helps explain why people visit certain environments and attach meanings to tourist settings” (D’Egidio 2014b, 59). The gaze is neither neutral, universal, nor static. Instead, there are many ‘gazes,’ multi-faceted and wrought through “a particular filter of ideas, skills, desires and expectations, framed by social class, gender, nationality, age and education” (Urry and Larsen 2011, 2). Gazes are shaped by both professionally and institutionally produced promotional materials, as well as other types of discourses and media. Indeed, the authors underline how, in the post-modern world, the tourist gaze is increasingly “media-mediated,” with people creating an image of the destination in their own minds through exposure to a wide variety of different materials:

[t]hrough representational performances, most tourist places have over time been inscribed with specific ‘imaginative geographies’ materialised and mobilised in and through books, brochures, postcards and photo albums. (Urry and Larsen 2011, 116)

To these, we can add other media such as video games, music videos, movies, and TV series, which all contribute to the shaping and negotiating of people's expectations and of the meaning attached to tourist destinations. These imaginary travels in turn mediate how people gaze upon places when they physically experience them as tourists, and influence how such locations are perceived and represented in travel narratives. Such narratives, often in multimodal form, are another type of text that has become increasingly popular in the past few years, since the advent of the Web 2.0 and the proliferation of UGC. Indeed, travelers nowadays share their experiences online in a multitude of ways: on social media platforms, dedicated forums, review platforms, blogs and micro-blogs, and media-sharing platforms such as YouTube, where trips are documented through video blog entries. Such content, produced by travelers themselves rather than industry professionals, both express the travelers' own gazes and potentially contribute to shaping the gaze of prospective visitors.

For these reasons, UGC in general and travel blogs in particular have become a prominent source of data for scholars and professionals in the tourist industry, as their popularity can influence destination image and prospective travelers' destination choices and plans (Alarcón-Urbistondo, Rojas-de-Gracia and Casado-Molina 2021; Lam, Ismail and Lee 2020; Tomaz and Wattanacharoensil 2020; Bosangit, Dulnuan and Mena 2012).

While destination image studies show a more direct link with travel intentions, the tourist gaze is also worth investigating. In point of fact, narratives constructed by travelers may be influenced by the degree of satisfaction they get from gazing at destinations and attractions through a pre-existing filter. Indeed, as Urry and Larsen (2011, 51) state that

satisfaction stems from anticipation, from imaginative pleasure-seeking. People's basic motivation for consumption is not simply materialistic. Rather, people seek to experience 'in reality' the pleasurable dramas they have already experienced in their imagination.

Therefore, evaluations of attractions and destinations may be influenced by their adherence to a pre-existing image built in the tourist's mind over time. Indeed, a gap is often identified between the projected destination image (marketer-based) and the tourist destination image (TDI) (e.g. Song and Kim 2016). Investigating the tourist gaze may also contribute to shedding light on what tourists expect from destinations, hence informing processes of destination branding.

1.3 Blogs as drivers of eWoM in tourism

As seen in the previous section, non-professionally produced texts may have entered the world

of marketing due to their potential in influencing customer decisions. Indeed, UGC has expanded the traditional phenomenon of Word-of-Mouth to the digital environment, in what is known today as electronic Word-of-Mouth (eWoM) or Word-of-Mouse. Word-of-Mouth consists in interpersonal communication “whom the receiver perceives as non-commercial, regarding a brand, a product or a service” (Arndt 1967, 3) or, as in our case, a destination. The worldwide extent of this phenomenon has come to the attention of scholars in multiple fields, trying to determine the role of blogs in influencing consumer attitudes and decisions. EWoM has two main functions that attract and influence readers (Ran et al. 2021, 5): a utilitarian function, which “provides useful elements such as price and variety,” and a credibility function. Clearly, the more credible a text, the more likely it is to influence its readers.

Word-of-Mouth plays a similarly significant role in tourism, and blogs seem to be a widespread and popular tool to share content with vast international audiences (Wenger 2008). The information shared in such texts appears to be perceived as more reliable than marketer-produced knowledge: Akehurst (2009) states that UGC and blogs, in particular, are more credible than traditional marketing material. This observation is corroborated by D’Egidio: “freedom and spontaneity in the act of writing may make travel blogs and trip reports a more reliable source of information on real travel experiences” (2014a, 147; see also Carson and Schmallegger 2007, 100). Mack, Blose and Pan (2008, 141) similarly point out that within blog types—Business to Customer, Government to Customer, and Customer to Customer—the latter was considered more credible, even though not as much as traditional Word-of-Mouth. As a consequence, it would be advisable that the role of eWoM in the evolution of destination image be taken into consideration by destination managers (Jalilvand and Heidari 2017, 728).

2. Data and methodology

2.1 Corpus design and compilation

The corpus for this study was compiled using the text mining software BootCat (Baroni and Bernardini 2004) for the purpose of linguistic investigation through a mixed-method approach; the analysis is, however, prevalently quantitative, integrating qualitative analysis of concordance lines to clarify quantitative results.

BootCat, written in Perl programming language, allows “automated extraction of specialized corpora and technical terms by web-mining” (Baroni and Bernardini 2004, 1313). The process works by inputting a number of keywords, or ‘seeds,’ which are then combined into sequences—‘tuples’—of different length. The resulting tuples, which may be manually selected, are in turn

used to run online queries, retrieving a number of URLs that may be potentially relevant to the creation of the corpus.

For the purpose of this study, the seeds used include major Chinese cities and attractions (Beijing, Shanghai, Hong Kong, Xi'an, the Great Wall of China) alongside other terms to orient the search towards travel accounts or guides. Five different searches were carried out using different sets of seeds, as reported in Table 1 below:

Set	Seeds
1.	travel blog, tour, China, Beijing, Great Wall, Xian, Hong Kong
2.	Beijing, China, my trip, travel blog, traveling
3.	Shanghai, city, itinerary, travel guide, travel tips, traveling to Shanghai, visit
4.	China, city, travel guide, travel tips, traveling to China, visit
5.	travel blog, Shanghai, my trip

Tab. 1: Seeds employed in BootCat

The choice of using different sets of seeds to create four smaller corpora to merge into a bigger one was made with the purpose of retrieving a more diversified range of texts that would include entries both in the style of travel guides and travel accounts. As the American spelling of 'traveling' was used as a search term, the corpus is expected to include mainly texts written by either American bloggers or international bloggers favoring American spelling conventions, which may be seen as a limitation in this study. The overall number of URLs retrieved was 1152: as this study focuses on texts written by travelers rather than institutional or professional sources, the output was manually sifted to exclude texts irrelevant to the scope of this study. Such URLs include social media platforms and media-sharing websites (e.g. Facebook, YouTube, Pinterest, TripAdvisor), tour operators, travel agencies, individual travel agents, and DMOs (Destination Marketing Organization). Some bloggers also offer customized travel planning services, but in those cases, the text was retained. In the few cases where the retrieved URLs were linked to the homepage of a viable data source, URLs linking to relevant entries within the websites were added manually to the final list. The final number of URLs used to create the corpus was 401. When the corpus was compiled on Sketch Engine, the option to remove duplicated documents was flagged to avoid skewing results with repeated material. For this reason, the final corpus includes 333 documents, totaling 472,445 words.

It should be noted that, while web-mining software allows users to compile relatively large specialized corpora in a time-saving way, it also has multiple limitations, such as little control and lack of metadata for the texts included in the sample. Another limitation of web-mined

corpora is the lack of ethnographic information about the bloggers. Possessing such information would have been useful in order to explore whether and how different backgrounds may influence travel narratives, especially in cases of travelers of Chinese descent who may be traveling to visit family or see the country of their ancestors.

2.2. Analytical approach

The analysis of the compiled corpus was carried out in three stages, employing the Keywords, Word Sketch, and Concordance tools in Sketch Engine to explore how and what type of gaze emerges from travel narratives on China.

In a first step, a list of keywords was retrieved to identify recurring themes in the dataset, using the EnTenTen18 corpus. The ‘Simple Maths’ (Kilgarriff 2012) parameter was set to 100. This “parameter give[s] prominence to different frequency ranges” (Kilgarriff 2012, 6), allowing users to produce keyword lists including more or less common words. A preference for rarer keywords in this study would have skewed results towards Chinese toponyms, so a setting of 100 was selected to provide a more balanced output.

In a second step, a collocation analysis was carried out on the selected destinations by using the destinations as node words in Sketch Engine’s Word Sketch tool, which calculates collocations using the logDice statistic measure and categorizes them according to grammatical relations to the node. A subsequent qualitative content analysis of a concordance sample allowed the identification of recurring positive and negative aspects of the individual places under examination and of China as an overall destination. Analyzing the context of each occurrence of all 5 destinations would have been too time-consuming due to the size of the corpus and the number of occurrences. For this reason, a concordance sample was created by limiting output to those instances where the node word occurs within three items (left or right) of an adjective, in order to focus on those passages that are more likely to contain an evaluation.

The third stage of this research looked at the collocational profile and concordance output of three words: *people*, *local**, *Chinese*, with the aim of clarifying how visitors gaze upon the local population.

3. Results

3.1 Keywords

To calculate keywords, Sketch Engine’s EnTenTen18 corpus of the English web was chosen as a reference corpus. Keyness score cutoff was set at 3,000, which resulted in 116 keywords² that

² The words *hour*, *day*, *night*, *spend*, as time references, were expunged from the list.

were classified according to a slightly modified version of Li and Wang's (2011) categories. A choice was made here to include all parts of speech (POS), as may be seen in Table 2 below:

Cities, Attractions, & Other Places	beijing, hong kong, china, shanghai, wall, temple, city, tower, xian, mutianyu, forbidden, street, attraction, terracotta, kowloon, chengdu, buddha, jinshanling, peak, badaling, mountain, museum, palace, square, great, park, pagoda, entrance ³ , hutong, warrior, bund, macau, island, lijiang, section, nanjing, gate, garden, guilin, destination, tiananmen, yangshuo, road, panda, yunnan, village, skyline, victoria, town, lantau, province
People & Life	chinese, tourist, crowd, blog, traveler, you
Travel & Transportation	bus, trip, travel, itinerary, tour, train, walk, subway, taxi, metro, ferry, ride, tram, cable, around, arrive, flight, locate, journey
Accommodation, Restaurants, & Food	hotel, hostel, restaurant, stay, food, dumpling, eat
Entertainment & Activities	visit, hike, explore, guide, shopping, shop, climb
Price & Quality	ticket, rmb, yuan, cheap, usd, hkd, cny
History & Culture	dynasty, tea, ancient, Ming, emperor, dragon
Infrastructure	station, airport, mtr
Policies & Regulations	visa
Evaluative Terms	famous, recommend, tip, amazing, beautiful, check, definitely, popular

Tab. 2: Keywords

While the category of *Evaluative Terms* was added to include those terms that did not fit in any of the others, it should also be noted that no keywords were identified that fit Li and Wang's category of *Environment & Safety*. From a look at the other categories, it is immediately noticeable that some of the keywords pertain to the blog genre, such as *blog* (*People & Life*) itself. As part of the corpus includes blog entries in the style of travel guides, the presence of the second personal pronoun *you* (*People & Life*), the noun *tip* (*Evaluative Terms*), and the verbs

³ It should be noted that around a third of the instances of the word *entrance* referred to admission fees (or lack thereof).

recommend and *check* (*Evaluative Terms*)—the latter often found in the phrasal verb form ‘check out’—reflects this type of texts. Indeed, the pronoun *you* is often used to address the blogger’s imagined audience (Scheidt, 2006, 116) directly, providing advice on how to plan their trip to China and what to visit (e.g. “While in Hong Kong, *you* should not miss;” “we *recommend* the tea houses;” “*check* out the Hutong area”).

In the first category—*Cities, Attractions, & Other Places*—we predictably find the destinations included as BootCat seeds, with the capital of China, Beijing, showing the highest keyness score. The category also includes proper names of locations and attractions within cities (e.g. *hutong, Kowloon, the bund, Victoria*), general names of places of interest (e.g. *temple, museum, palace, pagoda*), other cities (e.g. *Chengdu, Nanjing*) and of natural attractions (e.g. *mountain, peak*). The presence of *village* and *town* also suggests that travelers visited other types of destinations in addition to the bigger cities in China. *People & Life* includes references to both tourists (*tourist, traveler*) and local people (*Chinese*) as well as to the readership (*you*). The keyword *crowd* refers to the large population of big Chinese destinations as well as to the number of tourists at popular destinations. As will be seen, it is one of the recurring negative perceptions about China. *Travel & Transportation* as well as *Infrastructure* show that tourists move around China using a variety of means of transportation, and the words *itinerary* and *tour*, in addition to *road* and *journey*, seem to indicate that travelers visit multiple destinations during their time in China.

Accommodation, Restaurants, & Food includes keywords pertaining to accommodation (*hotel, hostel*) and eating (*restaurant*) options, including a type of food, *dumpling*, which exists in many varieties and that appears to be popular with tourists.

Keywords in the *Entertainment & Activities* category include activities that can be carried out in big cities (*shopping, hiring a guide*), at the Great Wall or in the wilderness (*climb, hike*), and in both (*visit, explore*).

Price & Quality consists of keywords denoting currencies (*hkd, usd, yuan, cny, rmb*), access to attractions and transportation (*ticket*), and an evaluation of such prices (*cheap*), as bloggers describe their own experience or give prospective tourists practical information.

The only keyword in the *Policies & Regulations* is *visa*: instructions on how to obtain one, or situations in which one is not needed, are often discussed in the corpus.

Keywords pertaining to the *History & Culture* category seem to be in line with previous studies that suggest an interest in ancient China and its history (*dynasty, Ming, ancient, emperor*), and culture (*tea, dragon*). Tseng et al.’s study indicated that the theme of place, as emerging from their Leximancer analysis of travel blog posts, was related to *old* and *ancient*, which links the

theme of place to “somewhere ancient that is able to represent Chinese culture and history” (2015, 351).

The category of *Evaluative Terms* tells us that the locations and attractions visited are defined as *famous* and *popular*, which is to be expected as the corpus was built to include some of China’s most renowned tourist destinations. The locations are also evaluated positively, as may be inferred from the presence of *amazing* and *beautiful*.

3.2 Destinations

3.2.1 Collocational profiles

To take a more detailed look at the main destinations mentioned in the blog entries, their collocational profile was considered using Sketch Engine’s WordSketch tool, with 5 as the minimum frequency for collocates.

The selected destinations are the following:

- Beijing (1,998 occurrences)
- Shanghai (1,093 occurrences)
- Hong Kong (1,278 occurrences)
- Xi’an (618 occurrences)
- Great Wall (1,055 occurrences)

While WordSketch classifies collocations according to their grammatical relation to the node, for reasons of space and scope of this paper, collocations⁴ are summarized according to destination in Table 3.

Destination	Collocations
Beijing	Tour, Hotel, central, layover, City, old, Beijing, China, Shanghai, itinerary, Airport, zoo, subway, Station, Line, Day, attraction, Park, guide, visit, leave, feel, get, Xi'an, Kong, easy, home, stay, arrive, hour, see, market, Palace, live, spend, week, time, weekend, people, restaurant, visit, close, trip, travel, flight, return, head, go, train, bus, accessible, Wall, further, distance, area, part, Xian, Badaling, kid

⁴ Collocates that appeared in multiple grammatical relations in the WordSketch and those that appeared in both lowercase and with an uppercase initial were reported only once. Auxiliary verbs (*be*, *have*, *do*) were expunged from the list, as well as words not pertaining to the content of the texts (e.g. *courtesy*, when used in picture attribution). The ‘prepositional phrases’ element of the WordSketch was also not included.

Shanghai	Garden, Old, Beijing, YuYuan, bund, Shanghai, Tower, Disneyland, Museum, Line, Center, skyline, Shanghai, City, itinerary, Pudong, China, Airport, Street, attraction, visit, explore, Garden, bund, Guangzhou, Xi'an, Pudong, Kong, city, day, arrive, try, street, time, see, hotel, stay, travel, trip, go, kid
Hong Kong	Peninsula, Tsui, Central, Shenzhen, Kowloon, Macau, visit, enter, know, leave, make, see, Shanghai, Beijing, China, easy, destination, city, day, Stay, hotel, attraction, visit, live, place, building, transport, area, restaurant, arrive, travel, visitor, flight, trip, go, get, view, part, budget
Xi'an	Open, Shaanxi, attraction, wall, House, Wall, China, Station, Tower, guide, station, city, locate, visit, Beijing, Shanghai, famous, warrior, day, stay, come, trip
Great Wall	Mutianyu, Badaling, Jinshanling, China, Jiankou, Juyongguan, Simatai, visit, hike, climb, walk, see, explore, reach, City, section, part, portion, view, trip, tour, Beijing, get, journey, go, camp, camping

Tab. 3: Collocations for the major Chinese destinations in the corpus

As may be seen from the table, the majority of collocates identified for the five destinations pertain to different aspects of traveling, including specific locations and attractions within and near the destination (e.g. *market*, *YuYuan*, *Simatai*), nouns pertaining to *Travel & Transportation*, *Accommodation*, *Restaurants*, & *Food*, *Entertainment & Activities*. The Great Wall, as an individual attraction rather than a city, diverges from this especially in terms describing activities, with *hike*, *climb*, and *camp* and *camping* appearing only in this destination. The four cities have more similar profiles, differing mainly in the proper names of locations and attractions. It should also be noted that the main nodes tend to collocate with *China* and other destinations, which may be part of the bloggers' travel itineraries or simply mentioned, for example in relation to transportation (e.g. "fly into *Hong Kong*, *Beijing*, *Shanghai*, or *Guangzhou*"), or in a list of recommended destinations (e.g. "*Beijing*, *Shanghai* and *Xi'an* are must-see cities"). If we look at the adjectives in the collocations, we find *old* (Beijing and Shanghai), *easy* (Beijing and Hong Kong), *famous* (Xi'an). The latter refers to attractions in or nearby Xi'an (the terracotta warriors) that are renowned. *Easy* pertains mostly to how simple it is to reach and navigate the cities. This seems to be an improvement on previous studies pointing out transportation as a problematic issue for travelers in China (Li and Wang 2011, 702-703). *Old* instead highlights the interest in the historical sights of the cities, that, as will be seen, are associated with a desirable gaze on the more traditional side of China.

In order to shed light on how the tourist gaze is manifested with regard to these destinations, a choice was made to complement the collocational analysis with content analysis.

3.2.2 Content analysis

A sample of statements was identified by running a concordance search for adjectives within three tokens—left and right—of each destination. The output was first filtered to select instances where the adjectives identified, referred to and evaluated the locations under examination. Those instances were then analyzed individually to determine whether the stance expressed was positive, neutral or negative, and, where relevant, which aspects of the destination were discussed. The recurring aspects in the output, which were used to categorize the statements, were: *modernity, culture/history, expenses, food, ease of navigation, locals, pollution, safety, crowdedness, language barrier, museums*, for a total of 308 statements. Where the immediate context was not sufficient to determine the blogger’s stance, the extended context was also taken into consideration. A summary of the overall stance on the destinations under investigation can be seen in Fig. 1.

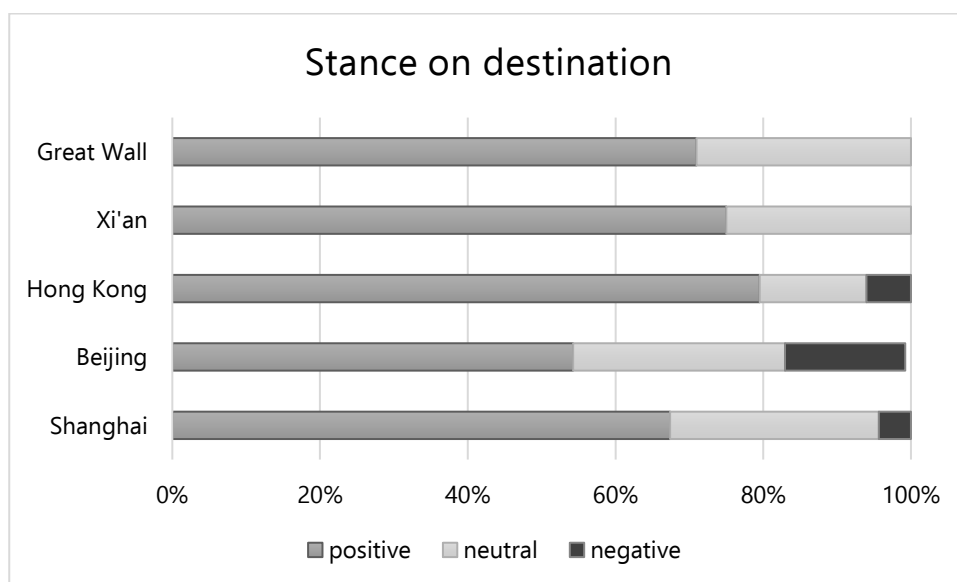


Fig. 1: Content analysis on sample concordances

As may be seen, the overall evaluation is visibly positive for the concordances analyzed, with only the three bigger cities including instances of negative evaluation. If we look more closely at which aspects of the city were evaluated negatively, we find that in Shanghai, the 2 negative evaluations related to *crowdedness* (“yet another overly populated city, Shanghai”) or a general evaluation of the city (“I’m not a huge fan of Shanghai”).

The 5 negative statements about Hong Kong are related to pollution (1) and expenses (4). Indeed, Hong Kong is compared to other Chinese cities in terms of prices, and it is mentioned

that it may not be the right destination for budget travelers and backpackers. It should also be noted, however, that prices in Hong Kong are evaluated positively in 7 other instances.

The city that received the most negative evaluations is Beijing, with 21 overall negative statements (16% of all items analyzed) pertaining to *pollution* (“Pollution is a real problem in Beijing”), *crowdedness* (“Chaotic, crowded Beijing”), *language barrier* (“Language can be a pretty big barrier in Beijing”), and other aspects such as *size*, *locals*, and *ease of navigation* (e.g. wheelchair accessibility).

No instances were marked as negative for Xi’an and the Great Wall. However, in the latter, some recommendations were made as to which sections of the wall to visit in order to avoid crowds and have a more ‘unique’ experience.

It should be said that while the elements marked as negative do highlight certain downsides to these destinations, as may be seen from the examples above the criticism is very often mitigated, with a few exceptions (“obscene smog;” “nasty smog”). As some of the blog articles in the dataset act as travel guides, such elements should not be necessarily seen as complaints but rather as warnings for prospective travelers, to whom bloggers also provide potential solutions, as in example (1) below:

- (1) Language can be a pretty big barrier in Beijing [...] so make sure have Google Translate with the offline translating

As the concordances analyzed here are a sample, it cannot be said with certainty that the ratio of negative assessments would be the same for the entire corpus. However, on the basis of the data examined, it may be observed that travel bloggers in China express an overall positive opinion on the destinations visited, accompanied by factual information (neutral statements). Even when underlining setbacks or downsides to their trips, visitors do not do so simply to complain, but mainly with the purpose of providing useful information to future travelers, in accordance with the utilitarian function of eWoM.

3.2.2.1 Culture and history vs. modernity

Concordance analysis showed that both the more traditional aspects of the destinations (*culture/history*, 40 occurrences, of which 29 positive, 2 negative and 9 neutral) and the more modern (*modernity*, 49 occurrences, of which 39 positive, 1 negative and 9 neutral) were evaluated positively, with some instances highlighting the contrast between the traditional parts of the city with the modern, more cosmopolitan areas. The coexistence of tradition and modernity in contemporary China is often mentioned by bloggers: in 17 instances in the

analyzed concordance sample, these two contrasting aspects of Chinese destinations are highlighted, described as “an interesting mix of the old and new,” with historical heritage sites “bursting with culture and history” and traditional activities (calligraphy, tea drinking) being juxtaposed to “great transport infrastructure,” “huge shopping streets and malls,” and “glittering” skylines. This appears to be in line with previous studies on perceptions of China, with travelers pointing out the contrast between the two aspects of China, the historical where traditional culture is still discernible, and the contemporary, feeding into the narrative of paradox that Xiao and Mair had put forward in 2006.

In Beijing and Shanghai, where both these aspects are clearly manifested, some statements appear to have an underlying nostalgic stance, pointing out that traditional, ancient China is increasingly less visible in these big cities, swallowed by skyscrapers, technological advancement and cosmopolitanism. More in detail, Old Beijing, defined as “authentic,” is also described as “slowly disappearing” and “harder and harder to find.” Prospective travelers are then encouraged to visit the Hutongs, narrow streets or alleyways with traditional courtyard residences where they can still get “a glimpse of ancient China.” Some bloggers therefore appear to underline an incongruity between the pre-constructed, exoticized image of China in their minds and the reality of the destination. Tourists “seeking authenticity in other ‘times’ and other ‘places’ away from [their] life” (Urry and Larsen 2011, 10; see also MacCannell 1999) therefore struggle to reconcile their mental image of the destination with the reality and point prospective travelers’ gaze to where they can satisfy their thirst for their idea of authenticity.

3.3 Perception of local people

This study also aims to investigate how tourists perceive and gaze at the local population of China. To this purpose, concordance searches were run of the lemma *local* and the words *people* and *Chinese*. Contrary to the analysis illustrated in the previous section, in this case the search was intentionally not restricted to any specific context, as these items would be likely to be accompanied not only by adjectives, but also verbs. This choice was made because people would not only be described through adjectives but also and especially through their actions (or as recipients of actions).

The search yielded 576 instances for *local*, 902 for *people*, and 1174 for *Chinese*, which were then manually scanned in order to identify occurrences where the local population was described and/or evaluated. These occurrences were classified according to a number of recurring topics, which are detailed in Fig. 2 below alongside their frequencies.

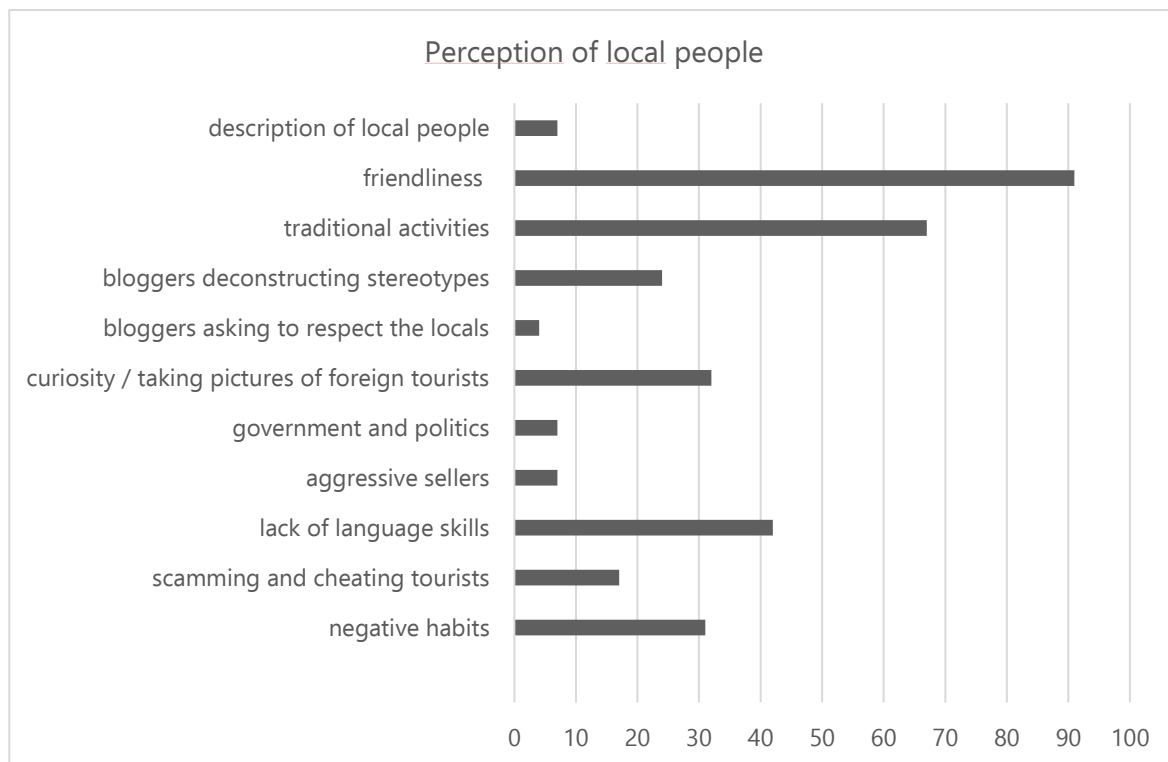


Fig. 2: Perceptions of local people

The overall outcome here is that the local population is described in very positive terms, as may be seen from the number of instances referring to local people's *friendliness*. They are described as "very friendly" and "very approachable and willing to help." The categories *description of local people* and *traditional activities* also include positive statements, as will be seen in greater detail. Certain bloggers take an active role in breaking down stereotypes, reminding their readers to always respect the locals ("don't treat the *locals* there as a zoo exhibit") or by stating that common stereotypes are either misconceptions or should be understood within the context of a different cultural frame ("Don't be surprised if a *local* tries to snap a sneaky photo of you"). The aspects that emerge from this analysis in one case overlap with the downsides already discussed in the destination analysis above, such as *lack of language skills* (e.g. "Not all *locals* in Beijing speak English," "most *people* in China don't speak English"). Recurring criticism in this dataset reflects habits and customs that have emerged as negative aspects by travelers to China in previous studies (Tseng et al. 2015, 355; Li and Wang 2011, 703-704): these include spitting (e.g. "*people* do spit in the streets"), line-cutting (e.g. "long lines disrespected"), jostling people around ("Don't be surprised if you're jostled or shoved") littering (e.g. "I see *people* throwing stuff on the ground"). Behavior that attracted tourist attention is the curiosity that the locals seem to have for non-Asian tourists and especially children, staring at them and sometimes asking for pictures. This, however, seemed to be a source of amusement for most

bloggers, who recount the experience as a specific aspect of their trip rather than complaining about the unwanted attention: bloggers claimed they felt “like a rockstar,” or that they had become their “new white star.”

The category of *Government and politics* was rarely mentioned in the sample analyzed, less so than in other studies where this aspect was more prominent (Tseng et al. 2015; Li and Wang 2011; Xiao and Mair 2006). One or two bloggers recommend refraining from talking about politics to the locals to avoid making them uncomfortable, and as both locals and foreigners can “be punished for inappropriate remarks.” However, no general discussion of policies and politics is made unless essential for traveling: *visa*, for example, is a corpus keyword, and requirements to travel to and through China are frequently addressed in the dataset. *Scamming and cheating tourists* (e.g. “local drivers will often hike prices for tourists”), as well as *aggressive sellers* (e.g. “people will strongly encourage you to pay a ticket”) are common complaints in many tourist destinations (Pearce 2011, 147), including Europe,⁵ so while they can impact overall destination image, these issues are not peculiar of China. As with the destinations, it should be noted that criticism is usually expressed in very mild terms, and, in blogs written in guise of travel guides, scams and frauds are conveyed as possible experiences prospective travelers might have and are often complemented with suggestions on how to avoid or solve them.

The most significant data pertaining to the tourist gaze on Chinese people comes from the category of *Traditional Activities*. Travelers appear to find pleasure in watching locals perform activities they recognize as being part of traditional Chinese culture, such as exercising or practicing tai-chi or kung-fu in parks, flying kites, playing a game of mah jong. In the data sample, such sights were more likely to be experienced in the Hutongs of Beijing, which are “a great way to see how the *locals* really live,” in city parks, where you can “get to know the *locals*,” and in temple grounds. Chinese people are also described as having “an amazing sense of community,” and they are indeed often described as taking part in group activities:

- (2) It’s a beautiful park where you can see the *locals* doing Tai Chi and other forms of exercise.
- (3) I saw large groups of *Chinese* [...] *people* doing Tai Chi.

⁵ A survey carried out by money.co.uk on 2,000 UK travelers to European destinations showed that more than one in five was victim of a scam. According to *The Telegraph*, which reported about the results, among the scams included in the survey were “included pickpocketing, paying a hidden tourist tax on arrival at a hotel, hire car excess costs, pushy street vendors, overcharging taxi drivers and street crime” (Morris 2015).

Other types of activities are mentioned in the dataset, such as hanging laundry out to dry, going shopping or to markets, all activities that catch the bloggers' attention and meet their need to experience what they perceive as traditional activities. These portraits of groups of people performing activities that are considered as quintessentially Chinese—practicing tai chi and kung-fu, playing mah jong, cycling—satisfy the tourist's need for that disappearing old China and, in a way, objectify the locals, whose function in the landscape is to “signify authenticity, induce romanticism and bring life to the scene” (Urry and Larsen 2011, 175). Tseng et al. similarly point out that “gazing upon groups of local people while going about their usual lives is sometimes the difference that many tourists seek to experience” (2015, 354), providing the sense of authenticity that they set out to fulfill. The locals are also described as friendly and extremely helpful, going out of their way to help tourists: this plays into a dynamic of reciprocity that links visitors and hosts, and that Maoz (2006) terms the ‘mutual’ gaze. Here, the locals are depicted as ‘collaborative,’ contributing to the tourists’ enjoyment and meeting their needs. Guests are also gazed upon, as could be seen from the data, with locals staring at western tourists and asking for pictures. The reactions to this kind of attention are of surprise and amusement, but not negative, which suggests that being the object of the locals’ gaze may highlight the sense of unfamiliarity and ultimately enhance the travel experience.

4. Concluding remarks

This study aimed to contribute to research on China as a destination, attempting to shed light on how international tourists perceive China and its people and what type of gaze they express and perpetuate through the texts they publish online. The image of China emerging from the collocational profiles and concordance analysis of some of the main destinations and attractions in the country appears to be consistent with existing research on visitor perceptions and destination image, with recurring positive and negative aspects matching results of previous studies. An exception may be identified in the category of *ease of navigation*, which emerged as a positive aspect of the bloggers’ trips but had been identified in the past (Li and Wang 2011) as a problematic aspect for visitors.

Recurring negative aspects of the trip that had already been identified in other studies were *pollution*, *crowdedness*, and *language barrier*, as well as certain local habits such as spitting and line cutting. However, it should be said that the language used does not characterize such statements as complaints but rather as warnings to prospective travelers, who often also find solutions to these problems within the same blog. Nevertheless, contrasting representations, as identified by Xiao and Mair (2006) and Becken et al. (2017) among others, seem to characterize

perceptions of China, as could also be seen in the data for this study. While recognizing and appreciating the metropolitan, technologically-advanced character of major Chinese destinations like Beijing, Shanghai, and Hong Kong, cultural and historical sites appear to draw the bloggers' attention and fascination: the juxtaposition of 'old' and 'new' China is recurrent in the data, alongside a certain nostalgic feeling for a more traditional China. This suggests that tourists in this dataset gaze at China looking for signs of what they perceive to be authentically and quintessentially Chinese; in Culler's words, they are "in search of [...] exemplary Oriental scenes" (1981, 127). This is more clearly observable in the perceptions of the local population, where visitors take pleasure in observing 'stills' of Chinese life that signify authenticity and constitute something "visually extraordinary" (Urry and Larsen 2011, 15), "something which for them is unusual, authentic in its otherness, a sign of an alien culture" (Culler 1981, 5) that, as Culler remarks, tourists emphasize when recounting their travels (1981, 5), as shown in these data. These timeless portraits of people hanging their laundry outside in the Hutongs or practicing martial arts in city parks contribute to the construction of Chinese people as an exoticized Other; this view is reinforced by the frequent comments about friendliness and helpfulness of the 'collaborative' local (Maoz 2006). The mismatch between the images signifying an idealized authenticity, and those describing current China, both in positive (cosmopolitan character of the cities, efficient public transportation) and in negative terms (poor air quality, lack of English speakers, local habits), confirms an idea of China as a 'polarized' destination (Becken et al. 2017, 134), with travelers appreciating the modern sides of China but longing to see a representation of the preestablished images of traditional China in their minds.

Acknowledgements: The Author wishes to acknowledge the funding that the former Italian Ministry of Education, University and Research granted to the Department of Foreign Languages and Literature at the University of Verona under the 'Departments of Excellence' plan for the project 'Digital Humanities Applied to Foreign Languages and Literature' (2018–2022).

Valeria Franceschi *is a temporary assistant professor in English Language and Translation at the Department of Foreign Languages and Literatures of the University of Verona, where she teaches at undergraduate and graduate level. Her research focuses on English as a Lingua Franca, especially in relation to plurilingual practices and on the business context (BELF). Her research areas of interest also include corpus linguistics and, more recently, tourism discourse.*

Works cited

- Akehurst, Gary. "User Generated Content: The Use of Blogs for Tourism Organisations and Tourism Consumers." *Service Business* 3 (2009): 51-61.
- Alarcón-Urbistondo, Pilar, María-Mercedes Rojas-de-Gracia and Ana Casado-Molina. "Proposal for Employing User-Generated Content as a Data Source for Measuring Tourism Destination Image." *Journal of Hospitality & Tourism Research* (2021): 1-22.
- Arndt, Johan. *Word of Mouth Advertising: A Review of The Literature*. New York: Advertising Research Foundation Inc., 1967.
- Baroni, Marco and Silvia Bernardini. "BootCaT: Bootstrapping Corpora and Terms from the Web." *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*. Edited by Maria Teresa Lino et al. Lisbon: European Language Resources Association (ELRA), 2004. 1313-1316.
- Becken, Susanne, et al. "Urban Air Pollution in China: Destination Image and Risk Perceptions." *Journal of Sustainable Tourism* 25.1 (2017): 130-147.
- Bosangit, Carmela, Juline Dulnuan and Miguela Mena. "Using Travel Blogs to Examine the Postconsumption Behavior of Tourists." *Journal of Vacation Marketing* 18.3 (2012): 207-219.
- Carson, Dean and Doris Schmallegger. "Blogs in Tourism: Changing Approaches to Information Exchange." *Journal of Vacation Marketing* 14.2 (2007): 99-110.
- Chang, Jung-Hua and Shuan-Huei Wang. "Different Levels of Destination Expectation: The Effects of Online Advertising and Electronic Word-of-Mouth." *Telematics and Informatics* 36 (2019): 27-38.
- Chen Po-Ju, Nan Hua and Youcheng Wang. "Mediating Perceived Travel Constraints: The Role of Destination Image." *Journal of Travel & Tourism Marketing* 30.3 (2013): 201-221.
- Culler, Jonathan. "Semiotics of Tourism." *American Journal of Semiotics* 1 (1981): 127-40.
- D'Egidio, Angela. "The Language of Tourists in English and Italian Travel Blogs and Trip Reports: A Corpus-Based Analysis." *Lingue Culture Mediazioni/Languages Cultures Mediation* 1.1-2 (2014a): 145-161.
- . "The Tourist Gaze in English, Italian and German Travel Articles about Puglia: A Corpus-based Study." *ICAME Journal* 38 (2014b): 57-72.
- Kilgarriff, Adam. "Getting to Know your Corpus." *Text, Speech and Dialogue. TSD 2012. Lecture Notes in Computer Science*. Edited by Petr Sojka et al. Berlin: Springer, 2012. 3-15.

- Jalilvand, Mohamed Reza and Ali Heidari. "Comparing Face-to-Face and Electronic Word-of-Mouth in Destination Image Formation: The Case of Iran." *Information Technology & People* 30.4 (2017): 710-735.
- Lam, Jason M. S., Hishamuddin Ismail and Sean Lee. "From Desktop to Destination: User-Generated Content Platforms, Co-Created Online Experiences, Destination Image and Satisfaction." *Journal of Destination Marketing & Management* 18 (2020): 100490.
- Li, Xiang R., et al. "The Effect of Online Information Search on Image Development: Insights from a Mixed-Methods Study." *Journal of Travel Research* 48.1 (2009): 45-57.
- Li, Xu and Youcheng Wang. "China in the Eyes of Western Travelers as Represented in Travel Blogs." *Journal of Travel & Tourism Marketing* 28.7 (2011): 689-719.
- MacCannell, Dean. *The Tourist*. New York: Schocken, 1999.
- Mack, Rhonda W., Julia E. Blose and Bing Pan. "Believe it or Not: Credibility of Blogs in Tourism." *Journal of Vacation Marketing* 14.2 (2008): 133-144.
- Maoz, Darya. "The Mutual Gaze." *Annals of Tourism Research* 33 (2006): 221-239.
- Morris, Hugh. "Europe's Worst Countries for Tourist Scams." *The Telegraph* 21 July 2015. <https://www.telegraph.co.uk/travel/news/Europes-worst-countries-for-tourist-scams/>. Last visited 17/08/2021.
- Pan, Bing and Xiang R. Li. "The Long Tail of Destination Image and Online Marketing." *Annals of Tourism Research* 38.1 (2011): 132-152.
- Pearce, Philip. "Tourist Scams: Exploring the Dimensions of an International Tourism Phenomenon." *European Journal of Tourism Research* 4.2 (2011): 147-156.
- Ran, Li, et al. "Marketing China to U.S. Travelers through Electronic Word-of-Mouth and Destination Image: Taking Beijing as an example." *Journal of Vacation Marketing* (2021): 1-20.
- Scheidt Lois A. "Adolescent Diary Weblogs and the Unseen Audience." *Digital Generations: Children, Young People and New Media*. Edited by David Buckingham and Rebekah Willett. London: Lawrence Erlbaum Associates, 2006. 193-210.
- Sketch Engine. <https://www.sketchengine.eu/>.
- Song, Seob-Gyu and Dae-Young Kim. "A Pictorial Analysis of Destination Images on Pinterest: The Case of Tokyo, Kyoto, and Osaka, Japan." *Journal of Travel and Tourism Marketing* 33.5 (2016): 687-701.
- Tomaz, Kolar and Walanchalee Wattanacharoensil. "One Does not Simply... Project a Destination Image Within a Participatory Culture." *Journal of Destination Marketing & Management* 18 (2020): 100494.

- Tseng, Chi, et al. "Travel Blogs on China as a Destination Image Formation Agent: A Qualitative Analysis Using Leximancer." *Tourism Management* 46 (2015): 347-358.
- Urry John. *The Tourist Gaze: Leisure and Travel in Contemporary Societies*. London: Sage, 1990.
- Urry, John and Jonas Larsen. *The Tourist Gaze 3.0*. London: Sage, 2011.
- Wenger, Anita. "Analysis of Travel Bloggers' Characteristics and their Communication about Austria as a Tourism Destination." *Journal of Vacation Marketing* 14.2 (2008):169-176.
- Wu, Mao-Ying, Geoffrey Wall and Philip L. Pearce. "Shopping Experiences: International Tourists in Beijing's Silk Market." *Tourism Management* 41 (2014): 96-106.
- Xiao, Honggen and Heather L. Mair. "A Paradox of Images' Representation of China as a Tourist Destination." *Journal of Travel & Tourism Marketing* 20.2 (2006): 1-14.
- Xiang Zheng, et al. "What Can Big Data and Text Analytics Tell us about Hotel Guests Experience and Satisfaction?" *International Journal of Hospitality Management* 44 (2015): 120-130.
- Zhao, Yanyun and Binjie Liu. "The Evolution and New Trends of China's Tourism Industry." *National Accounting Review* 2.4 (2020): 337-353.