Mining Brand Perceptions from Twitter Social Networks

Aron Culotta

Department of Computer Science, Illinois Institute of Technology, Chicago, IL, 60616 aculotta@iit.edu

Jennifer Cutler

Kellogg School of Management, Northwestern University, Evanston, IL, 60208 jennifer.cutler@kellogg.northwestern.edu

Consumer perceptions are important components of brand equity and therefore marketing strategy. Segmenting these perceptions into attributes such as eco-friendliness, nutrition, and luxury enable a fine-grained understanding of the brand's strengths and weakness. Traditional approaches towards monitoring such perceptions (e.g., surveys) are costly and time-consuming, and their results may quickly become outdated. Extant data mining methods are not suitable for this goal, and generally require extensive hand-annotated data or context customization, which leads to many of the same limitations as direct elicitation. Here, we investigate a novel, general, and fully automated method for inferring attribute-specific brand perception ratings by mining the brand's social connections on Twitter. Using a set of over 200 brands and three perceptual attributes, we compare the method's automatic ratings estimates with directly-elicited survey data, finding a consistently strong correlation. The approach provides a reliable, flexible, and scalable method for monitoring brand perceptions, and offers a foundation for future advances in understanding brand-consumer social media relationships.

Key words: social media; brand image; market structure; Twitter; attribute ratings; perceptual maps; big data; data mining; social networks

1. Introduction

Understanding how consumers perceive brands is fundamental to much of marketing strategy. A central analytical tool used to do so is *perceptual mapping*, which organizes brands according to how consumers rate them with respect to attributes such as eco-friendliness or luxury (Green et al. 1989, Shocker and Srinivasan 1979, Steenkamp et al. 1994). Consumer ratings are typically collected through surveys or other elicitation means (Aaker 1996, Hauser and Koppelman 1979, Lehmann et al. 2008, Steenkamp and Van Trijp 1997); however, these data are costly and time-consuming to collect and may quickly become outdated.

2

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

The recent proliferation of social media use by both marketers and consumers offers a promising data source to understand consumer perceptions; yet the noise, volume, and ambiguity of such data pose substantial challenges to algorithmic solutions. In this paper, we introduce and validate a fully-automated and highly generalizable method for estimating brand perceptions along a perceptual attribute[1] of choice from publicly available secondary social media data, specifically Twitter. To the best of our knowledge, there are no extant data mining approaches developed for this task.

At a high level, our algorithm takes as input a brand name and a query specifying the attribute of interest (e.g., "eco-friendliness"). It then returns a real value indicating the strength of association between the brand and the attribute. The main source of evidence used by our approach is the similarity between a brand's Twitter account and a set of *exemplar* accounts representing a perceptual attribute — e.g., the similarity between Smart Automobile's account and those of the EPA and GreenPeace may signal its perceived eco-friendliness. The method we develop is innovative in several ways. First, while most extant methods analyze user-generated text, we instead rely only on the structure of the brand's social network, which offers advantages in simplicity and scale. Second, we focus our analysis on the platform Twitter, which has received limited attention in the marketing literature, but offers advantages in data relevance and accessibility. Third, we introduce a fully-automated and highly generalizable process that requires only a keyword as input to generate near real-time estimates of brand ratings for an attribute mapping to that keyword. By leveraging the crowd-organization of social media, we circumvent the often extensive manual tuning and customization requirements of extant data mining approaches, thus providing a versatile and scalable method that can be applied to a range of marketing inquiries.

To validate the effectiveness of the method, we use it to estimate perceptual ratings along three example attributes (eco-friendliness, luxury, and nutrition) for over two hundred brands across four sectors[2] (Apparel, Cars, Food & Beverage, and Personal Care), and collect directly elicited survey ratings for the same set of brands and attributes. We find

---

[1] For consistency, we use the word *attribute* throughout the paper, though we mean it to include any specific aspect of brand identity that can be identified through a keyword and rated along a continuum. These might also map to perceptual *dimensions* or *associations*, as they are referred to in other areas of the literature.

[2] As will be explained in §4, attributes were tested only for sectors that made sense–e.g., nutrition perceptions were not estimated for car brands.

an average correlation over all sector-attribute combinations of 0.72, indicating that this fully-automated approach provides a reliable signal of current brand perceptions. This correlation meets or exceeds standards set in prior literature for related (though distinct) tasks, despite the manual customization required to implement these extant methods.

Our core contribution, then, is a new methodological tool to quantify consumer perceptions of brands with respect to a specified attribute. Our approach is a real-time, low-cost alternative to extant methods that firms and researchers can use for a number of common marketing tasks, such as generating perceptual maps, monitoring market structures, and informing research models (Green 1975, Green et al. 1989, Hauser and Simmie 1981, Schmalensee and Thisse 1988). In addition to the algorithm itself, our scientific contribution consists of a multi-faceted empirical validation against primary survey data, including an exploration of how a number of algorithmic variants affect accuracy.

In the next section, we discuss relevant work from the marketing literature, and describe how our contributions add to this work. In §3, we discuss the theoretical foundations motivating the approach, and in §4 we describe the methodology in detail. We describe our validation methodology in §5 and the main results in §6. §7 provides a series of sensitivity analyses. Finally in §8, we summarize the implications of this work, note its limitations, and provide recommendations for future research.

2. Background and Related Work
2.1. Brand Attribute Ratings

Marketing managers have long relied on estimates of consumer perceptions of brands along attributes of interest to inform marketing strategy (John et al. 2006, Lancaster 1971, Lehmann et al. 2008). Perhaps most notably, such estimates are used as the primary input for generating perceptual maps, which have been used by managers since at least the 1970s to understand the relative positioning of competitive brands (Hauser and Koppelman 1979, Johnson and Hudson 1996), and are widely considered a foundational analytical tool in marketing research (Green et al. 1989, Shocker and Srinivasan 1979, Steenkamp et al. 1994). Developing improvements to perceptual mapping techniques consistently remains a priority for marketing researchers (Bijmolt and van de Velden 2012, Day et al. 1979, Dillon et al. 1985, Kaul and Rao 1995).

Researchers have proposed both compositional and decompositional techniques for eliciting brand ratings from consumers. Both require recruitment and interaction with a large

and diverse set of participants; in the former, users are asked to directly rate brands on a numeric scale according to their strength in a given attribute, while in the latter, users are asked to perform sorting tasks on brands, from which attribute ratings are inferred (Huber and Holbrook 1979). Some research suggests that compositional methods (i.e., rating brands via surveys) can provide greater validity (Bottomley et al. 2000, Hauser and Koppelman 1979), and these are widely used in marketing practice (Steenkamp et al. 1994).

Yet, compared to the wealth of research focused on the advancement of techniques for making inferences from such brand attribute ratings, surprisingly little research has focused on advancing methods for obtaining the ratings themselves (Steenkamp and Van Trijp 1997). At the same time, many researchers have called out substantial limitations resulting from the requirement of collecting primary data to inform these analyses, including difficulty and expense in recruiting sufficient participants, and in maintaining participants attention and cooperation during tasks (Day 1975, McDaniel et al. 1985, Steenkamp and Van Trijp 1997). Ultimately, current methods of eliciting attribute ratings from consumers require substantial "trade-offs between completeness, cost, and feasibility" (Aaker 1996).

It is these difficulties that motivate the research goal of this paper— to develop a flexible and automated means of estimating brand attribute ratings from publicly available secondary social media data. In the following sections, we describe extant data mining approaches that have emerged in recent years in the marketing literature, discuss obstacles to applying them to the current research goal, and motivate the new approaches we introduce.

2.2. Text Mining

Text analysis of user-generated content (UGC) is a frequently used approach in the marketing literature for mining consumer perceptions from social media data (Fader and Winer 2012). One technique receiving notable attention is associative analysis. Here, researchers have employed clustering and semantic network techniques on UGC to discover how product features or brands are perceptually clustered by consumers (e.g., Archak et al. 2011, Lee and Bradlow 2011, Netzer et al. 2012). This approach is not suitable for our research goal, however, as we seek to estimate the strength of perceived brand ratings along predetermined attributes of interest.

Another popular technique is sentiment analysis—quantifying the overall positive and negative sentiments expressed online about a brand (e.g., Sonnier et al. 2011, Tirunillai

and Tellis 2012, Ludwig et al. 2013). The amount of manual input needed to tune these analyses is often substantial, and the accuracy and generalizability of the models across platforms and contexts is debated. For example, Das and Chen (2007) compared five sentiment classification algorithms and reported accuracy rates ranging from 25–40% (up to 67% when ambiguous messages are pre-filtered) for out-of-sample validation.

Unlike classifying sentiment, which is inherently context-neutral, classifying attribute-relevance through text requires that users author content about the brand that is relevant to the attribute of interest. This substantially limits (and likely biases) the data available, and potentially excludes many brand-attribute combinations from effective analysis. Furthermore, the problem of how to classify UGC by relevance to a perceptual attribute remains. The most common approach in the literature to automated text classification by topic involves matching the text against a pre-defined keyword list (Tang and Guo 2013). There are many limitations of such approaches. First, relevant keyword lists require substantial time and effort to curate for each topic. Second, they are static and may not adequately reflect the often rapidly evolving linguistic idiosyncrasies of "netspeak" inherent to many social media sites (Crystal 2001). Third, the accuracy of such models is limited and variable. The Linguistic Inquiry and Word Count (LIWC) is perhaps the most popular tool used in social science for keyword-based topic classification, containing keyword lists for almost forty topics. However, external validation tests are only reported for about a third of its categories, and where they are reported, the correlation coefficients between the tool's output and human judges' ratings range from .07 (sadness) to .87 (family) with the average of .45 (Pennebaker et al. 2007).

Thus, given the limitations of using UGC to infer brand attribute perceptions, we choose to explore an alternate source of information—the social connections of a brand's supporters. Using social structures offers distinct advantages over content analysis for inferring user perceptions, as we outline below.

2.3. Social Network Mining

When inferring consumer perceptions from UGC, one is limited, by definition, to only incorporating information provided by active, content-producing consumers. However, research has shown that fewer than 50% of Twitter users actively post content (Toubia and Stephen 2013), and the vast majority of posts come from a small minority of elite users (Wu et al. 2011). Yet, the silent majority can have a substantial impact on brand image through their

"mere virtual presence" (Naylor et al. 2012) in the brand's network, as the composition of a brand's online follower base has been shown to both reflect and influence brand image (Kuksov et al. 2013). Thus, by looking to the social structure of a brand's follower base rather than the text of UGC, we can capture potentially useful information from every single fan, regardless of whether they create or consume content. This is particularly useful for estimating perceptual attributes that consumers may be less likely to directly mention in brand conversations than core product features. Although some marketing researchers have begun to use online social network data for such purposes as predicting consumer behavior (Goel and Goldstein 2013) and understanding information diffusion (Goel et al. 2012), we believe we are the first to use a brand's social connections as a measure of brand perceptions.

2.4. Twitter

Although analyses of Twitter data are relatively rare in the marketing literature (with notable exceptions including Toubia and Stephen (2013) and Stephen et al. (2010)), we find that Twitter is an ideal platform for our analysis for four reasons. First, it is popular. Approximately 20% of U.S. adults were active on Twitter in 2014, and that percentage is growing steadily (Duggan et al. 2015). As of mid-2013, 77% of Fortune 500 companies maintained active Twitter accounts, compared to 70% that had Facebook pages (Barnes et al. 2013). Second, it is relevant. Twitter is used extensively for brand image and personality development, as frequent and conversation-like messages can be delivered at low cost to a large brand community (Etter and Plotkowiak 2011, Kim and Ko 2012, Kwon and Sung 2011). Accounts can be maintained at the firm level or the brand level, allowing communities to develop at scale appropriate to a firms' brand strategy—an important distinction when studying brand image perceptions, as brands can be dominated to varying degrees by their parent corporate brands (Berens et al. 2005). Third, its social connections are public. Except for a small minority of protected accounts—estimated at 8% (Cha et al. 2010)—followers of Twitter accounts are publicly visible and can be programmatically accessed through Twitter's API. This strengthens the relationship between social network and brand image, as the social signal of "who follows a brand" can be a strong component of brand image (Naylor et al. 2012), and also allows social network information to be easily extracted by marketers interested in implementing the method for novel research. Finally, it is organized. Because Twitter accounts are commonly organized by users into

topic-based Lists, accounts users deem relevant to a perceptual attribute can be identified programmatically, eliminating the need for manual curation. While we focus on Twitter for these reasons, the core idea of using follower connections to infer brand perceptions could be extended to other platforms, and we encourage future research to explore this further.

A question commonly faced by researchers interested in mining consumer insights from any online source is: to what extent can the perceptions or behaviors inferred extend to more general populations? Initial inquiries to this question have been encouraging for generalizability, showing a positive relationship between online and offline consumer loyalty (Danaher et al. 2003) and brand image (El Gazzar and Mourad 2012). Furthermore, consumers are increasingly looking to a brand's social media presence to form judgments about the brand (Baird and Parasnis 2011, Naylor et al. 2012). In our validation section, we investigate this issue further with our own data by validating our Twitter-based estimates against survey results obtained through a separate population, and examining similarities in brand attribute ratings across numerous demographic categories.

3. Theoretical Foundations

Our proposed approach is motivated by a wide-range of research in the social and computational sciences suggesting that proximity in a social network can be indicative of similarity.

A wealth of research shows a tendency for people to express affinity towards those whom they perceive to be similar (Lydon et al. 1988, Morry 2007, Naylor et al. 2012). This property of value homophily has been observed widely in sociology, social network analysis, and computational science (McPherson et al. 2001). When a user follows an organization or brand on Twitter, it provides explicit evidence of a user's voluntary public association with that entity. This is generally interpreted as an expression of affinity (Naylor et al. 2012, Kuksov et al. 2013), and survey research (conducted on the related social networking platform, Facebook) supports that the primary reason users connect to a brand is that they like its products, and that most fans are customers (see Pereira et al. 2014). While this is not always the case (e.g., an environmentally-conscious user might deliberately choose to follow a brand because it is environmentally unfriendly to track its claims), we expect this behavior is likely too uncommon to affect overall trends.

Thus, through the lens of value homophily, we expect that the followers of Twitter accounts that are widely acknowledged as exemplifying a particular attribute (e.g.,

8

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

environmentally-focused non-profit accounts that exemplify the perceptual attribute of eco-friendliness) are, in aggregate, likely to particularly value that attribute.[3] Similarly, we expect that a brand that has an unusually large following of users who value a particular attribute is likely to be perceived as strong in that attribute. This line of reasoning is further supported by a wealth of research that shows a strong relationship between brand image and the characteristics and identities of the brand's supporters and followers (Bearden and Etzel 1982, Berger and Heath 2007, Childers and Rao 1992, Escalas and Bettman 2003, Kuksov et al. 2013, Naylor et al. 2012). Taking these principles together, we attempt to infer the perceived strength of a brand for a given attribute based on the extent to which its follower set overlaps with that of a large set of accounts that exemplify the attribute.

We note that on the surface this is a simple approach, as motivations for following brands and exemplar accounts can be varied and complex—many factors are likely at play, beyond shared value for a perceptual attribute. However, the strength of our approach is its use of "big data"; by examining millions of social links we are able to overcome the noise introduced by infrequent spurious follower connections. Thus, by aggregating over many links, we hypothesize that we can generate meaningful estimates of brand perceptions at a scale and frequency not possible using extant means. In the next section, we explain the implementation of the proposed method in detail.

4. Social Network Mining Methodology

Given a brand (e.g., Smart Automobile) and a perceptual attribute (e.g., eco-friendliness), our goal is to develop an automated method to assign a score to the brand, where a high score indicates a strong perceived relationship between the brand and the attribute.

Our proposed approach is based on the notion of an attribute *exemplar*. An exemplar is an individual or organization that is known to be strongly affiliated with an attribute. For example, the Environmental Protection Agency and the Sierra Club may be said to exemplify the eco-friendliness attribute. As identifying appropriate exemplars may not always be practical *a priori*, our approach allows users to instead specify the attribute using a search query (e.g., "environment" for eco-friendliness). Exemplars are then found based on this keyword, as described below in §4.1.

---

[3] Although this may not be the case for any specific user following a particular exemplar account, we look to an aggregate increase in attribute valuation for the collection of all followers over a wide set of exemplars.
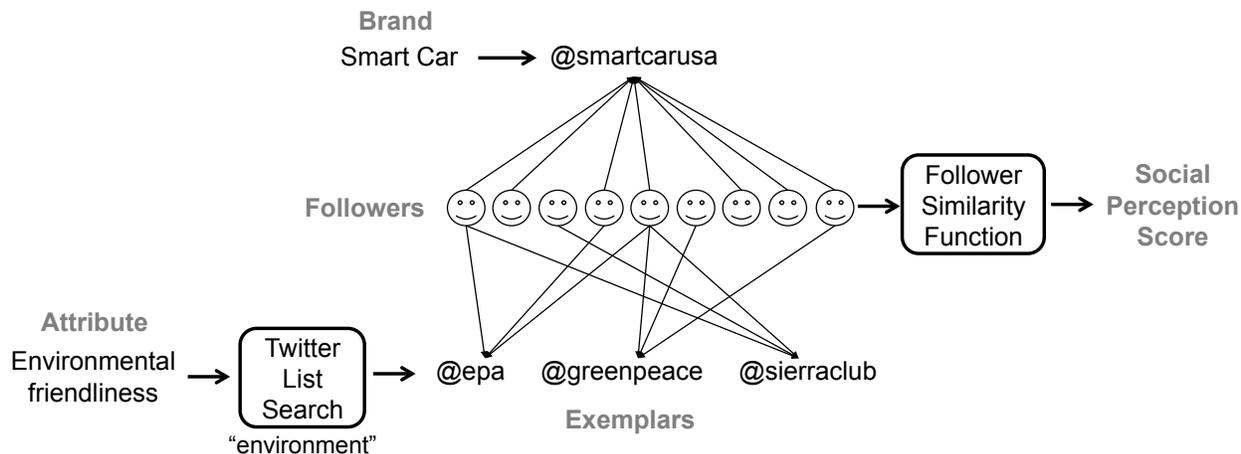
Figure 1: An overview of our approach. Given a brand's Twitter handle and a search query representing a perceptual attribute, our algorithm first collects exemplar accounts representing the attribute, then computes a similarity function between the followers of the exemplars and those of the brand.

To assign a score to a brand, our method first identifies Twitter accounts for the brand and a set of attribute exemplars. It next collects social network information for each account; specifically, it collects the followers of the brand and the followers of each exemplar. Finally, a node affinity score is computed between the brand's account and the exemplar accounts, using standard graph-theoretic measures from the social network analysis literature. We denote this final affinity score the *Social Perception Score* (SPS). Our central hypothesis is that the higher a brand's SPS is for a perceptual attribute, the more strongly consumers associate the brand with that attribute.

In its most generic form, our approach requires two inputs from the user: (1) the Twitter handle for a brand of interest,[4] (2) a search query representing a perceptual attribute. With these, relevant Twitter data are collected and analyzed to produce the SPS. Figure 1 depicts an overview of our approach. We enumerate the four steps below:

1. **Input**
   - B: the Twitter handle for a brand (e.g., *@SmartCarUSA*)
   - Q: a search query representing a perception attribute (e.g., "environment")

2. **Collect Exemplars**

---

[4] Alternatively, a brand name may be provided; we offer a process in Appendix A to automatically identify the corresponding Twitter account.

10

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

- Use Q to retrieve from Twitter a list of exemplar accounts $\mathbf{E} = \{E_1 \ldots E_k\}$ that reflect the specified perception attribute, e.g., {*@epa, @greenpeace, @sierraclub*}.

3. **Collect Followers**

- Collect F_B, the set of Twitter accounts that follow brand B.
- Collect $F_\mathbf{E} = \{F_{E_1} \ldots F_{E_k}\}$, the Twitter accounts that follow each exemplar in $\mathbf{E}$.

4. **Compute Follower Similarity**

- Compute the similarity between brand followers F_B and exemplar followers $F_\mathbf{E}$.
- Return the resulting *Social Perception Score*, $SPS(B, \mathbf{E})$.

In the following subsections, we expand on these steps in more detail.

4.1. Selecting Exemplar Twitter Accounts

The foundation of our approach requires a set of Twitter accounts that exemplify the attribute to be rated. In some use cases, this may be manually provided. For example, the attribute of eco-friendliness may reasonably be exemplified by selecting the Twitter accounts of known environmental non-profits. There are at least three reasons we may want to automate this step: (1) for some attributes, it may be difficult to identify exemplar accounts; (2) automation allows us to scale the approach to produce social perception scores for many attributes; (3) less well-known accounts may often be more valuable in computing SPS (as results in §7 suggest).

Our approach requires as input a query term or phrase representing the attribute. For example, in our validation below, we use the term "environment" as a search term representing the eco-friendliness attribute. With this term, our program automatically queries Twitter to identify accounts representative of the query. To do so, we rely on Twitter Lists.

A Twitter List is a manually-curated collection of Twitter accounts. Any user can create their own List or subscribe to others' Lists. Thus, Lists are used to follow the posts of a related set users. Lists can be understood as a crowd-sourced method to categorize accounts, i.e, a "folksonomy" (Peters 2009).[5]

We use Twitter Lists to programmatically collect exemplar accounts for each attribute as follows: Given a query reflecting an attribute (e.g., "environment"), we submit the query to Twitter's search engine, which returns Lists as well as tweets. We iterate through the first 50 List result and retain accounts which appear on at least two different Lists (to

---

[5] While we were unable to find a publicly reported count of the total number of Twitter Lists that have been created, a Google search suggests there at least 5 million (using the query `site:twitter.com/*/lists/`).

reduce the number of false matches). The resulting accounts become exemplars for the subsequent stage of analysis.

4.2. Computing follower similarity

For each brand account B and exemplar account E_i, we use the Twitter API to programmatically download the list of users that follow each account. The final step, then, is to compute a score indicating the similarity between the followers of a brand and the followers of an exemplar set. Viewed abstractly, the graph-theoretic problem is to determine the similarity between two nodes based on their neighboring nodes. This problem is central to a number of social network analysis problems, including community detection, link prediction, and recommendation engines (Pan et al. 2010, Grabowicz et al. 2012).

Based on this literature, we select a common and empirically successful similarity function, the Jaccard index. The Jaccard index defines the similarity of two sets as the cardinality of their intersection divided by the cardinality of their union. For two sets X and Y, the value is:

$$J(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}$$

Thus, letting F_B be the set of followers of brand B, and F_{E_i} be the followers of an exemplar account $E_i \in \mathbf{E}$, we compute the Jaccard index $J(F_B, F_{E_i})$. In addition to its wide use in link prediction and community detection (Pan et al. 2010, Grabowicz et al. 2012), this metric has the additional advantages of being scalable and transparent. Further, Jaccard scores are normalized appropriately so that we can compare brands with different numbers of followers.

Once we have computed $J(F_B, F_{E_i})$ for each exemplar $E_i \in \mathbf{E}$, we next must determine how to combine the Jaccard values for each exemplar of an attribute. While simply taking the average score seems natural, we introduce a modification to encode our intuition that being similar to niche exemplars is more important than being similar to popular exemplars. For example, one exemplar account for eco-friendliness is @DarrenGoode, an environmental reporter for the website Politico. This account has under 8K followers. Contrast this with another exemplar, @AlGore, which has nearly 3M followers. A user who follows both @DarrenGoode and brand B provides a stronger source of evidence of the environmental affinity of B than a user who follows both @AlGore and brand B.

A simple way to incorporate this intuition is to weight each exemplar inversely proportional to its number of followers. This is analogous to the "inverse document frequency"

12

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

adjustment used in information retrieval to encourage documents containing rare query terms to be ranked higher than documents containing common query terms (Manning et al. 2008). The resulting social perception score for a brand B and exemplar set $\mathbf{E}$ then becomes the weighted average:

$$SPS(B,\mathbf{E}) = \frac{\sum_{E_i \in \mathbf{E}} \frac{1}{|F_{E_i}|} J(F_B, F_{E_i})}{\sum_{E_i \in \mathbf{E}} \frac{1}{|F_{E_i}|}} \tag{1}$$

As compared to survey results, we observed that SPS tends to have more positive skew and greater coefficient of variation (the ratio of standard deviation to mean). To mitigate this, our final score takes the square root of the quantity computed in Equation 1 above.

While we have motivated this final SPS computation based on the literature and our intuition, there are a number of different choices one could make that would result in a somewhat different function (e.g., choice in similarity metric, averaging, and square root transformation). In §7.4, we empirically compare our proposed measure with a number of competing alternatives to investigate how robust the results are to these choices.

5. Validation Methodology

The next step in our process is to validate the extent to which the SPS values match actual perceptions. To do this, we compare SPS values with directly elicited survey ratings for each brand and attribute in our test set, as described below.

5.1. Attributes and Exemplars

To test the generalizability of our approach across different perceptual attributes, we considered three attributes: eco-friendliness, luxury, and nutrition.

Using the Twitter List search methodology described above, we used the queries "environment," "luxury," and "nutrition" to collect exemplars for each of the three perceptual attributes. For each of the exemplar accounts, we collect the IDs of up to 50,000 of their Twitter followers. In total we have 74 eco-friendly exemplars (2.0M followers, 1.0M unique), 110 luxury exemplars (4.4M followers, 2.3M unique), and 405 nutrition exemplars (4.7M followers, 2.7M unique). Figure 2 shows the distribution of the number of followers for these accounts.
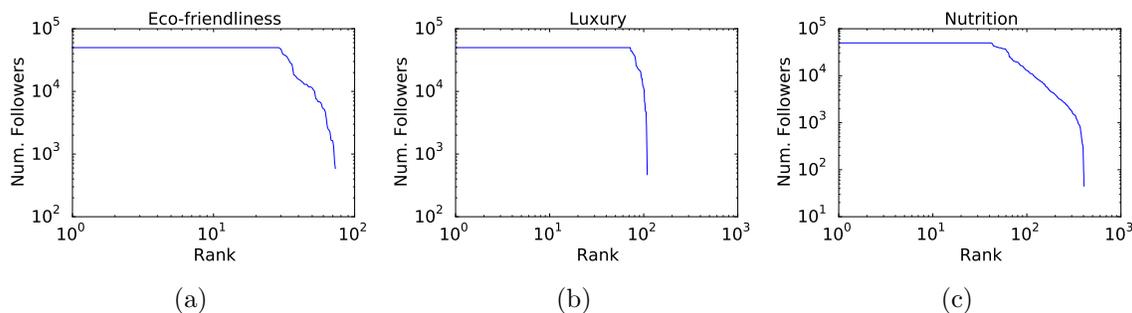
Figure 2: Distributions over the number of followers collected for exemplars of each attribute. We limit our analysis to at most 50,000 followers per account.

5.2. Brand Selection

To test the generalizability of our approach across brands, we desire a wide range of brands from a variety of sectors. To collect brands, we used the website GoodGuide.com, which maintains a large selection of brands categorized by sector. We first downloaded all brands listed under the four largest sectors: Car, Apparel, Food & Beverage, Personal Care. Because Personal Care contained many brands primarily known for products in different sectors, we eliminated Personal Care brands that were not primarily known for hair or skin care products. Next, we used a semi-automated script to match brand names with their corresponding Twitter accounts—this process is described in Appendix A. We manually validated all matches, then discarded any brands for which we could not find an active, English-language account (where active is defined as having at least 1,000 followers and 100 tweets). If a brand had accounts for multiple locations including the U.S., the U.S. version was used. Next, we eliminated sub-brands that matched only to their parent brand's Twitter account. Finally, if more than 70 brands remained in a sector, we randomly selected 70 to keep for our analysis. These eliminations resulted in a test set of 239 brands. Table 1 lists the number of brands per sector, the perceptual attributes the sector was tested for[6], and examples of brands included.

We used Twitter's API to collect up to 500,000 followers for each brand in our test set. In total, we collect Twitter user IDs for 30.6M brand followers (14.6M of which are unique). Figure 3 shows the distribution of follower counts for the brands.

---

[6] Note that certain attributes are only relevant to certain sectors (e.g., while we can consider the eco-friendliness of cars, clothes, and food, it does not make sense to consider the nutrition of cars).

14

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

| Sector	N	Attr.	Example Brands
Apparel	70	Eco, Lux	Ann Taylor, Calvin Klein, Champion, Chanel, GAP, Hanes, J.Crew, Levi's, Nike, North Face, Ralph Lauren
Cars	37	Eco, Lux	Audi, Bentley, BMW, Buick, Cadillac, Honda, Jeep, Kia, Lexus, Mini, Porsche, Rolls Royce, Subaru, Tesla, Volvo
Food & Bev.	70	Eco, Nut	Cheerios, Dannon, Doritos, Godiva, Motts, Oscar Mayer, Snapple, Sunchips, Triscuit, Red Bull, Stouffers
Pers. Care	62	Eco	Aveda, AXE, Burt's Bees, Clearasil, Clinique, Dove, Herbal Essences, L'Oreal, Old Spice, Pantene, Suave

Table 1: Brand examples by sector
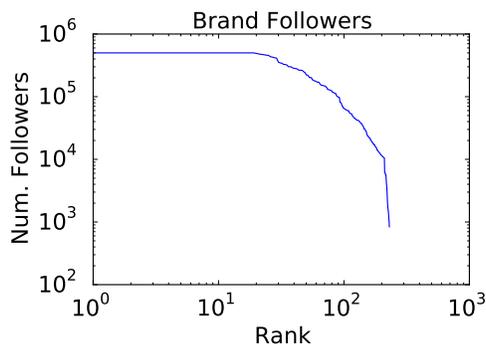
Figure 3: Distribution over the number of followers collected for the 239 brands in our test set. We limit our analysis to at most 500,000 followers per account.

5.3. Survey Design

Given our list of 239 brands and three attributes, we next directly elicited survey ratings to determine how strongly consumers associate each brand with each attribute. We administered the surveys through Amazon Mechanical Turk (AMT), which has been shown to be a reliable source for social science data collection (e.g., see Buhrmester et al. 2011). Brands were grouped into sets by sector and attribute, and five hundred AMT participants were recruited to rate each set. Participants were required to be located in the United States and to have a successful track record on AMT (they must have completed at least one hundred prior assignments with an acceptance rate of at least ninety-five percent). Participants were asked to rate each brand in the set on a scale of one to five according to how much they believed the brand aligned with the perceptual attribute at hand, and were provided a
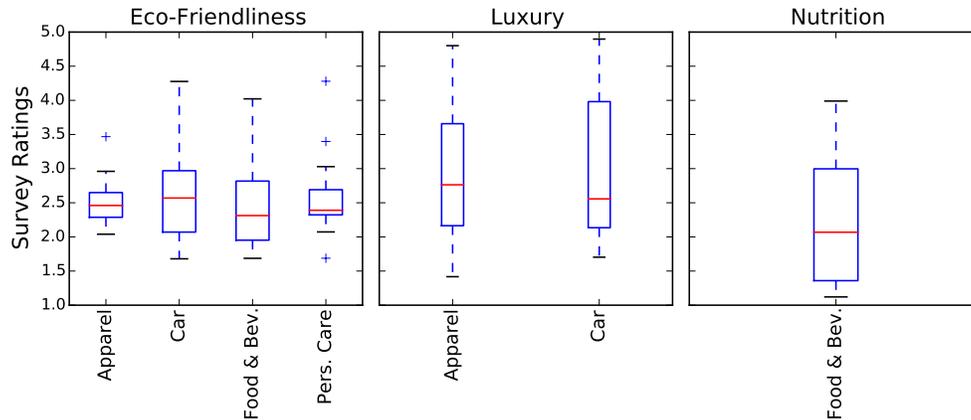
Figure 4: Boxplots for the survey results.

separate column to select if they did not recognize a brand. Each participant rated between 39 and 70 brands, all within the same sector and for the same attribute. Brand order was randomized for each participant. After rating all the brands in the set, participants were asked to indicate categories for their age, gender, education, and household income. We note that we did not ask participants whether they followed brands on Twitter; we expect that the percentage of positive responses for any given brand in this context would be too small for meaningful analysis. However, we encourage future researchers to explore more direct connections between survey respondents and Twitter activity.

Numerous attention filters were included to ensure valid responses[7], and the responses of any participants who did not pass these checks were discarded. On average, 340 participants per set passed the attention filters (68%).

Brand recognition rates varied from 100% for brands such as BMW, Honda, and Pepsi to less than 10% for brands such as Rodial, Elemis, and Bumble Bar, with an overall average recognition rate of 75%. As some of the brands in the list were somewhat obscure, and we wanted to ensure that ratings were generated by a large sample of users, we further filtered out brands that were not recognized by at least two hundred participants. The final brand counts are included in Table 2.

Finally, we computed the average rating for each brand for each perceptual attribute. Figure 4 summarizes the survey responses by sector and attribute. These plots suggest that

---

[7] The attention filters used included asking participants to select a particular response for a given line, to identify a brand they rated one turn prior, and to appropriately indicate that they didn't recognize a nonsense word inserted in place of a real brand.

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

16

Attribute	Sector	r	N
Eco	Apparel	0.62	39
	Car	0.75	37
	Food & Beverage	0.73	62
	Personal Care	0.82	20
Luxury	Apparel	0.62	47
	Car	0.68	37
Nutrition	Food & Beverage	0.80	55
	Average	0.72	

Table 2: Pearson correlation coefficients (r) for SPS and survey scores by perceptual attribute and sector. N is the number of brands remaining in the set after unfamiliar brands were filtered out.

a number of attribute-sector combinations exhibit positive skew; e.g., there are a small number of brands with very high eco-friendliness ratings. In Appendix C, we perform an additional demographics analysis of the survey responses to validate the representativeness of these data.

6. Validation Results

To evaluate the overall accuracy of SPS, we look at each sector-attribute combination individually and compute the Pearson correlation coefficient between the average survey ratings and the SPS estimates for each brand. Table 2 lists these results.

Across all sectors and attributes, the average correlation coefficient between SPS and the survey averages is 0.72, with the strongest correlation for the eco-friendliness of personal care brands (0.82), and the weakest for the eco-friendliness and luxury of apparel brands (0.62). We find these consistently high correlations to be encouraging for the use of this automated methodology in marketing practice. As automated attribute-specific brand perception estimation is a novel contribution to the marketing literature, there is not a clear external benchmark to directly compare our performance against. For an indirect comparison, we look to Netzer et al. (2012), who recently presented a methodology for a related goal of estimating car brand co-association sets (i.e., car brands that consumers cluster together in purchase consideration sets) using text analysis of online user forum
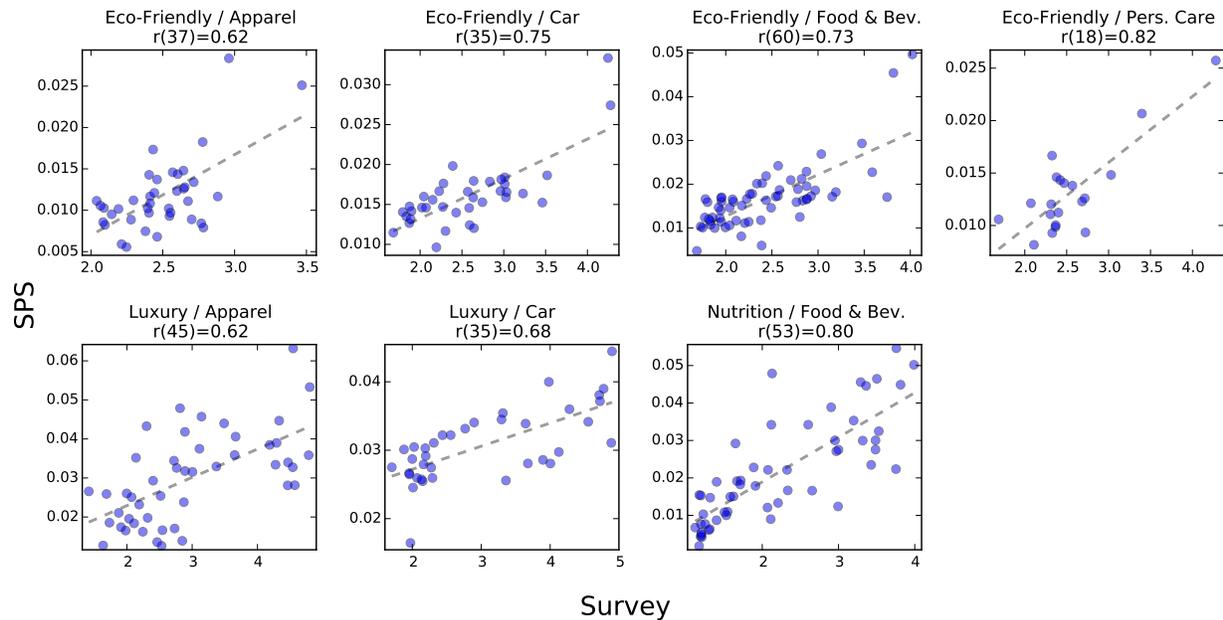
Figure 5: Twitter-based SPS estimates of perception against survey ratings of perceptions for each attribute and sector combination. Each data point represents a brand in that sector. A dashed regression line is included for reference.

posts. In their validation, they obtained correlations between their estimates and survey results that ranged from 0.43 to 0.55. The nature of their goal allowed them to further validate some of their results against brand-switching data. Here, their correlation was 0.75 (slightly higher than our 0.72). However, we note that the comparison is indirect—their goal was different than ours (brand co-associations vs. perceptual attribute ratings).

To verify that the correlations obtained are indicative of attribute-specific perceptual information captured by SPS, we performed three ancillary tests. First, to explore the possibility that the results might be due to a halo effect around brand popularity–i.e., that popular brands are perceived as stronger in all attributes–we computed the Pearson correlation between a brand's number of Twitter followers (as a measure of popularity) and its average survey rating for each attribute. We found a reassuringly small correlation coefficient of 0.107 (p = 0.06), indicating that attribute ratings are not reflections of popularity. Second, to explore whether SPS might be capturing a more general dimension of brand perceptions, we tested whether the SPS estimates for one attribute substantially predicted the survey ratings for a different attribute. For each sector in which multiple attributes were rated, we computed the Pearson correlations between each cross-attribute

combination of survey and SPS scores (e.g., luxury survey ratings and eco-friendly SPS estimates). The average correlation coefficient for such pairs was 0.059, indicating that the SPS estimates provide information specific to the designated attribute. Finally, to explore whether the results were being driven primarily by demographic differences (e.g., if SPS might be reflecting that eco-friendly exemplars and eco-friendly brands both appeal to younger users, independent from environmental values), we ran a series of regression analyses, including age, gender, and income profiles for a subset of brands for which we could obtain this information. We found that for each attribute, the coefficient for SPS as a predictor of average survey scores is positive and highly significant while controlling for demographics. Details of this analysis are provided in Appendix B.

To explore our results in more detail, Figure 5 shows scatterplots of SPS versus survey rating for each sector-attribute combination. The first row of plots indicates that SPS for eco-friendliness also follows the positive skew as observed in the survey results. That is, while most brands have low-to-moderate eco-friendliness, there are a handful of brands that have very high eco-friendliness ratings. These include The North Face and Timberland (Apparel), Tesla and Smart (Car), Organic Valley and Nature's Path (Food & Beverage), and Burt's Bees and Aveda (Personal Care). To investigate the possibility that these highly rated brands are driving the observed correlations, we also compute the Spearman rank coefficient, which is less sensitive to outliers. The average Spearman coefficient across all attribute-sector pairs is 0.64, indicating that a strong relationship still remains. We note, however, that these brands are not "outliers" in the sense of poor quality data; rather, these reflect the fact that only a small number of brands have cultivated a very strong perception of eco-friendliness. Considering this fact, and that the magnitude of such perceptual differences—not just the rankings—are important for informing marketing strategy, we continue our analysis using Pearson correlation.

Finally, we look in more detail at the scatter plots to better understand when the method aligns with survey results and when it does not. Figure 6 shows a close-up of the plots for car brands for attributes eco-friendliness and luxury. From these figures, we can identify some brands for which the SPS values need improvement. For example, in the bottom plot of Figure 6, it appears that Lamborghini has a lower SPS value than we might expect, given that it is well-known for its luxury sports cars and is highly rated by survey respondents. Examining the Twitter presence of Lamborghini, we observe that
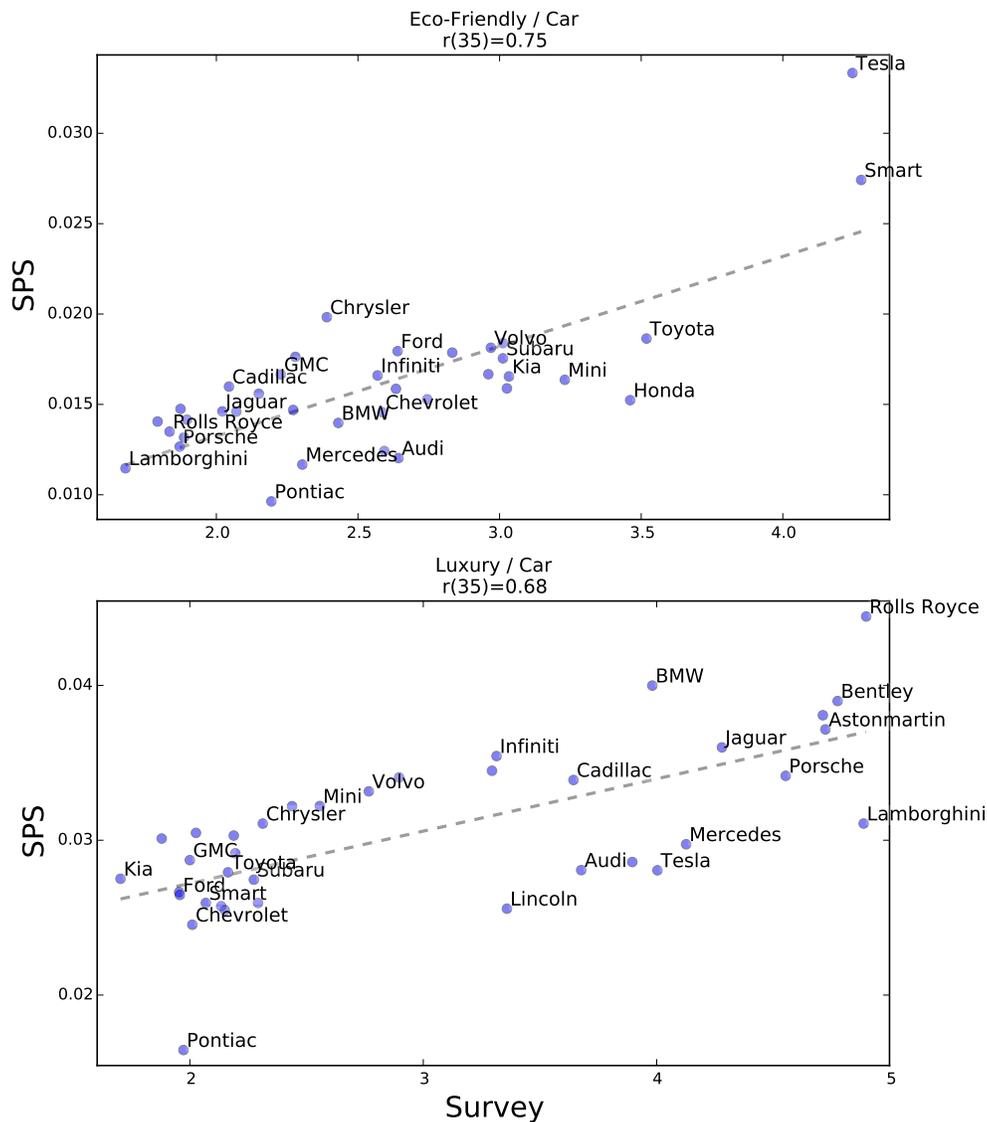
Figure 6: SPS vs. average survey ratings for eco-friendliness and luxury perceptions of car brands. For readability, only some brand name labels are displayed.

many of the communications involve pictures and discussions of auto shows, which focus on cutting-edge technology, rather than cars that one would actually purchase today. Thus, we conjecture that the followers of Lamborghini may comprise more sports car technology enthusiasts, rather than people interested in purchasing a Lamborghini. This is a potential limitation of the simplicity of the approach — if people follow one brand for systematically different reasons than they follow other brands, it may make it difficult to compare SPS values. While this appears to occur rarely in our data, we encourage future work in this area to identify and adjust for such instances.
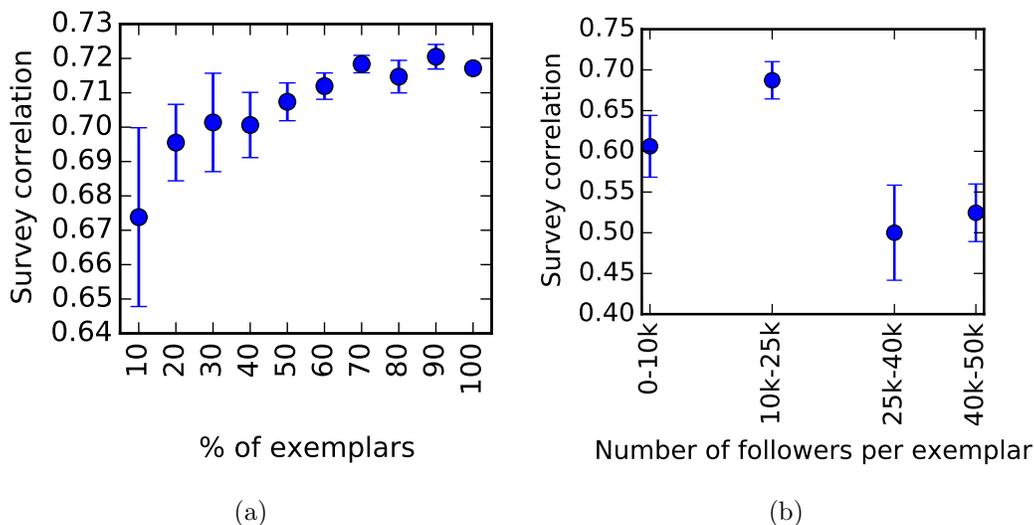
Figure 7: (a) Survey correlation (with standard errors) by percentage of exemplars sampled, averaged across all attributes and sectors. (b) Survey correlation (with standard errors) by number of followers per exemplar; we sample five exemplars for each bin (10K, 25K, 40K, 50K) and plot the average correlation across all attributes and sectors, averaged over four random trials.

Overall, our validation results suggest that the automated method we propose can deliver a viable and industry-acceptable signal of attribute-specific brand perceptions, providing marketing managers interested in monitoring specific attributes of brand image with a fast, flexible, and low-cost alternative to survey administration.

7. Robustness Checks

In this section, we perform additional analyses to understand how the different factors that makeup SPS affect results. We consider several aspects of the exemplar set (including quantity, quality, and number of followers), as well as variants of the SPS function (using different similarity metrics, averaging, and transformations), and report how these aspects affect the quality of the resulting estimates of brand image perception.

7.1. Sensitivity to number of exemplars

In the first analysis, we consider how the quantity of exemplars influences accuracy. One would expect that more exemplars provide a more representative set of accounts, making SPS less sensitive to poorly-chosen exemplars. To investigate this, we select random subsets of exemplars to consider when computing SPS values, then compare how the resulting

scores correlate with survey results. In Figure 7a, we plot the correlation averaged over all sector-attribute combinations as the percentage of exemplars used increases, averaged over four trials. That is, to generate the first value, we select a random 10% of exemplars for each attribute when computing SPS values; standard errors are computed from these four trials.

We can make several observations from this figure. First, as expected, the quality of the SPS values tends to increase as the number of exemplars used increases. However, the quality appears to plateau around 70%, suggesting that at a certain point additional exemplars become redundant. Second, the standard error decreases as the number of exemplars increases, which is expected given the greater sample size. Finally, and perhaps more interestingly, we observe that correlation can be very high using only 10% of the exemplars. Of course, there is large variance, which indicates that *which* 10% we choose matters. We consider this further in the next section.

7.2. Sensitivity to choice of exemplars

In this section we examine more closely how SPS quality varies by choice of exemplars. First, we consider how quality varies by the number of followers an exemplar has. Recall that our proposed SPS score computes the weighted average over exemplar similarity, where the weight is the inverse of the number of accounts that follow the exemplar. Our intuition was that following "niche" exemplars was a stronger indicator than following more popular exemplars. To investigate this intuition, we consider filtering exemplars by their number of followers. Specifically, we partition exemplars into bins of sizes {0-10k, 10k-25k, 25k-40k, 40k-50k}, based on the number of followers. For each attribute, we sample five exemplars from each bin and then compute the SPS values for each relevant sector and the correlation with the survey values.[8] We repeat this four times and report the average correlation per bin, again averaged over each sector-attribute pair. Figure 7b plots these averages with standard error bars.

The overall correlations are lower because we select only five exemplars per bin (consistent with the conclusions of Figure 7a). In addition, exemplars with the most followers (>40k) tend to be the least useful for estimating perception. This relationship appears to be non-linear, with the 10k-25k bin resulting in the highest correlation. We speculate

---

[8] We sample only 5 exemplars per bin to control for the number of exemplars; for example, only 8 luxury exemplars have fewer than 10K followers, compared with 286 nutrition exemplars.
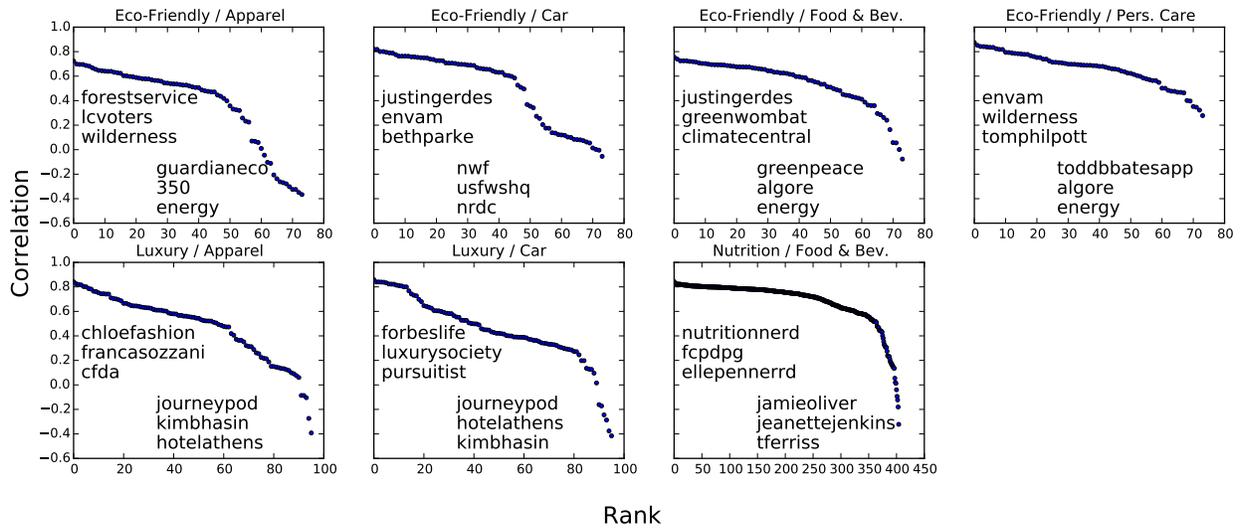
Figure 8: Correlation between survey responses and SPS values, using a single exemplar at a time, sorted by rank. The top three and bottom three exemplar accounts are displayed in each plot.

that there is a "sweet spot" in which an exemplar has enough followers to calculate reliable statistics, but not too many as to dilute the cohesion of its followers. Thus, these "niche" exemplars do indeed appear to be instrumental to quality SPS values.

Figure 7 also provides guidance on some of the practical considerations made in §4. Due to Twitter rate limits[9], we restricted exemplars to those appearing in the first 50 results of the Twitter List search; furthermore, we limited our collection to at most 50k followers per exemplar. The plateauing correlation in Figure 7a suggests that collecting additional exemplars will have limited value. Furthermore, Figure 7b suggests that very popular exemplars are the least valuable, so collecting more than 50k followers per exemplar is unlikely to improve accuracy. Thus, while we originally chose these cutoffs for convenience, these results suggest that more data are not likely to significantly increase the quality of our estimates.

Next, to better understand the variation of SPS quality by exemplar, we consider the correlations obtained using each exemplar in isolation. For example, we consider the correlation between eco-friendliness surveys and the SPS values generated using only a single

---

[9] For example, the Twitter API allows us to collect about 300k follower IDs per hour. See `https://dev.twitter.com/rest/reference/get/followers/ids`.
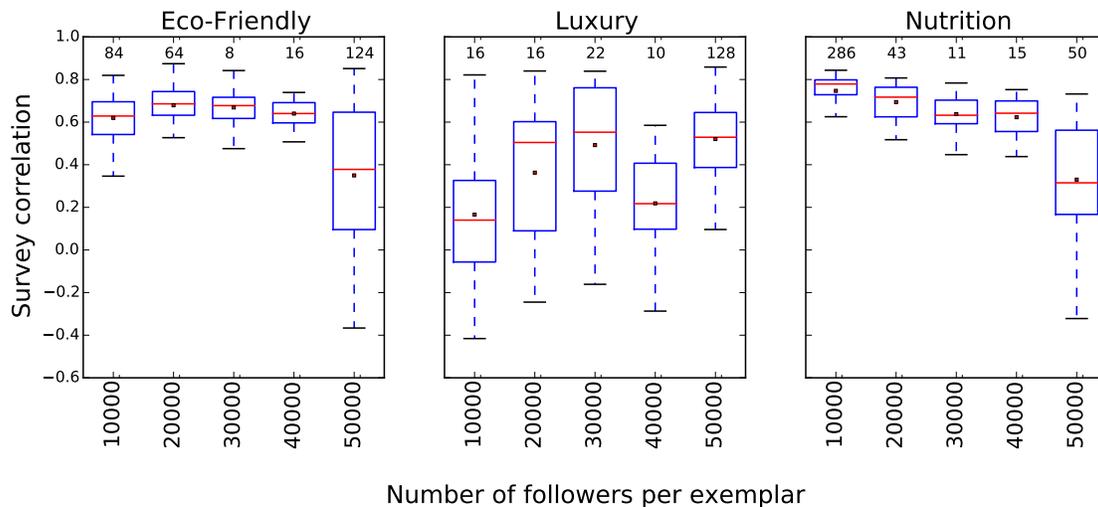
Figure 9: Boxplots of the survey correlation obtained using a single exemplar at a time, averaged across sectors. Exemplars are binned by number of followers. For example, the first boxplot in the first panel shows the distribution over survey correlations using eco-friendly exemplars having 0-10,000 followers. There appears to be a non-linear relationship between number of followers and correlation that varies by attribute; exemplars with between 10,000-20,000 followers appear to perform well across all attributes.

exemplar (such as @greenpeace). In Figure 8, we plot these results for each of the seven sector-attribute pairs; each figure plots the exemplars in descending order of the resulting correlation with survey responses. Additionally, we have indicated the names of the three highest and three lowest ranked exemplars. This figure allows us to visualize the variation in SPS quality by choice of individual exemplar. "Elbows" in these curves tend to occur after half of the exemplars are considered, indicating that most of the exemplars found by our proposed search method can lead to high quality SPS values. It is also clear that some exemplars have been poorly selected — averaging over exemplars therefore appears critical to maintaining useful estimates in the presence of this noise.

Looking at the top-ranked exemplars, we see that there are some exemplars that are highly ranked across sectors — e.g., Justin Gerdes, an environmental journalist, is the top exemplar for eco-friendly Cars and Food & Beverages. However, in general different sectors are best reflected by different exemplars. For example, in eco-friendliness, the top exemplars for apparel have an outdoors focus (Forest Service, The League of Conservation Voters, the Wilderness Society), while for cars, the top exemplars tend to have more to

do with global warming and energy (Beth Parke is also an environmental journalists). Similarly, for luxury, the top exemplars for Apparel typically pertain to fashion (CFDA is a fashion trade association, Franca Sozzani is the editor of the fashion magazine Vogue); the top exemplars for Cars focus more on high-end technology products. These findings suggest that practitioners may be able to apply domain knowledge to refine the exemplar query to tailor it to a sector of interest. Although our presented approach used only one keyword per attribute, combinations of terms such as "environment, conservation" or "environment, clean energy" could be input to generate for more specific and sector-relevant exemplar lists. Domain experts interested in such customization could refine their queries through manual examination of returned exemplars, and/or by validating results against a smaller set of brands for which perceptual ratings may already be known.

Finally, we revisit the relationship between the number of followers an exemplar has and the quality of the resulting SPS values. Figure 9 again considers the survey correlation for each exemplar in isolation, but here we group exemplars by the number of followers (in bins of size 10K). For example, the first boxplot in Figure 9 shows the distribution over correlations obtained by considering individual eco-friendly exemplars with between 0 and 10K followers. These plots suggest a difference between luxury and the other two attributes. While eco-friendliness and nutrition display a mild negative correlation between the number of exemplars and survey correlation, luxury displays a mild positive correlation. Thus, popular exemplars appear to serve as high quality exemplars for the luxury attribute, but less popular exemplars appear to be beneficial for eco-friendliness and nutrition.

7.3.   Manual versus automatic exemplar selection

We also compare our automatic exemplar selection algorithm against a manually collected set. In cases where it is practical to manually identify such a set, we expect that the resulting exemplars may serve as better representations of an attribute than exemplars collected automatically.

For eco-friendliness, a natural starting point is to identify non-profit organizations that support environmental causes. To obtain a list of environmental nonprofits, we use the CharityNavigator[10] API to collect the names of all national and international nonprofits assigned to the Environmental Protection and Conservation sector. We then manually

---

[10] http://charitynavigator.org

Attribute	Sector	r	N
Eco	Apparel	0.74	39
	Car	0.80	37
	Food & Beverage	0.76	62
	Personal Care	0.80	20
Average		0.78	

Table 3: Pearson correlation coefficients (r) for SPS and survey scores using as exemplars the list of environmental non-profits from CharityNavigator.com. N is the number of brands remaining in the set after unfamiliar brands were filtered out.

identify the Twitter accounts for each, where possible, resulting in a total of 79 exemplars (comparable to the 73 accounts identified using the automated method). We compute SPS scores as before using the followers of these exemplars. Table 3 shows the correlations with survey results for the eco-friendliness attribute.

We can see that hand-selecting exemplars can result in more accurate SPS values. Compared with the results using auto-generated exemplars in Table 2, the CharityNavigator exemplars result in an average correlation that is .05 higher (.73 versus .78), averaged across the four sectors for the eco-friendliness attribute. These results should give the practitioner some guidance in balancing the cost-benefit tradeoff of computing such estimates. If it is not too burdensome to identify a large sample of exemplars, then this may improve the accuracy of the perception estimates; however, if it is difficult to identify exemplars, the automated approach produces competitive results.

7.4. Sensitivity to similarity metric

In §4.2 we proposed an SPS value that used Jaccard similarity, averaging over exemplars weighted by the number of followers. In this section, we revisit some of these algorithmic choices to determine how they affect survey correlation.

We consider two alternative similarity metrics commonly used in social network analysis:

- **Cosine similarity (cosine):** If we consider each list of followers as a binary vector, then the cosine similarity between a brand's followers F_B and and an exemplar's

followers F_{E_i} is the cosine of the angle between the two vectors, which can be written as:

$$C(F_B, F_{E_i}) = \frac{|F_B \cap F_{E_i}|}{\sqrt{|F_B|}\sqrt{|F_{E_i}|}}$$

This metric is often use as a measure of similarity in information retrieval and clustering problems (Manning et al. 2008).

- **Conditional probability (cnd-prob):** Another intuitive measure is the empirical conditional probability[11] that the follower of a brand B also follows exemplar E_i:

$$P(x \in F_{E_i} | x \in F_B) = \frac{|F_B \cap F_{E_i}|}{|F_B|}$$

We can see that these two alternative metrics differ from Jaccard only in the denominator. (Recall that the denominator for Jaccard is $|F_B \cup F_{E_i}|$.) Thus, these different normalizations will affect the interplay between the number of followers of each exemplar and the number of followers of a brand.

In addition, we compare both weighted average (as proposed in §4.2) with a simple average of exemplar similarities. Finally, we also consider variants that optionally transform SPS values with square root versus without.

Table 4 displays survey correlations for the twelve systems resulting from all combinations of similarity metric, averaging strategy, and transformation. Averaged across all attribute-sector pairs, the correlations range from .63-.72, with our proposed system (jaccard/wt-avg/sqrt) tied for the highest with the another configuration (cosine/wt-avg/sqrt). In aggregate, the survey correlations appear robust to these algorithmic decisions, suggesting that the value of this approach is not limited to one particular implementation.

Examining individual columns reveals some qualitative differences between attributes and sectors. For example, while **cnd-prob** is competitive with the other metrics for eco-friendliness and nutrition, it performs substantially worse for luxury. We suspect this is due to the fact that, by only normalizing by the number of followers of a brand, **cnd-prob** imposes too large of a penalty on popular brands. That is, **cnd-prob** is biased to give lower scores for popular brands. Indeed, upon further analysis, we find a mild negative correlation

---

[11] We thank the anonymous reviewer who recommended this metric.

| Method | Variant | Eco-Friendliness | | | | Luxury | | Nutr. | |
		Appar.	Car	Food	PC	Appar.	Car	Food	Avg.
jaccard	simp-avg	0.53	0.54	0.62	<u>0.87</u>	0.67	0.63	0.70	0.65
	wt-avg	0.62	0.73	0.67	0.86	0.59	0.70	0.73	0.70
	simp-avg, sqrt	0.44	0.53	0.64	0.82	0.69	0.62	0.77	0.64
	wt-avg, sqrt	**0.62**	**<u>0.75</u>**	**<u>0.73</u>**	**0.82**	**0.62**	**0.68**	**<u>0.80</u>**	**<u>0.72</u>**
cosine	simp-avg	0.40	0.49	0.59	0.85	0.70	0.75	0.69	0.64
	wt-avg	<u>0.64</u>	0.63	0.65	0.86	0.61	<u>0.87</u>	0.73	0.71
	simp-avg, sqrt	0.31	0.47	0.60	0.79	<u>0.73</u>	0.72	0.75	0.63
	wt-avg, sqrt	0.61	0.68	0.70	0.81	0.65	0.84	0.78	<u>0.72</u>
cnd-prob	simp-avg	0.62	0.68	0.67	0.83	0.54	0.44	0.71	0.64
	wt-avg	0.61	0.73	0.67	0.80	0.42	0.47	0.73	0.63
	simp-avg, sqrt	0.58	0.67	0.72	0.79	0.55	0.44	0.78	0.65
	wt-avg, sqrt	0.61	0.74	0.72	0.75	0.44	0.47	0.79	0.65

Table 4: Pearson correlation coefficients (r) for SPS and survey scores across a range of similarity functions. Our proposed method is highlighted in bold. The largest values in each column are underlined.

between the number of followers of a brand and its survey rating for eco-friendliness and nutrition; however, for luxury, there is a mild positive correlation.

This may also explain the difference between the top two configurations, (jaccard/wt-avg/sqrt) and (cosine/wt-avg/sqrt). While the average correlation is the same, we can see that jaccard outperforms cosine for eco-friendliness and nutrition attributes, while cosine outperforms jaccard for luxury. The denominator for jaccard is linear in the number of followers of a brand, while for cosine it is sublinear (square root). Thus, brands with more followers are penalized less by cosine than jaccard, partly explaining these differences.

In summary, these robustness checks suggest that our main conclusions hold across a wide range of similarity functions, as well as across a number of alternative methods of collecting attribute exemplars.

28

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

8. Discussion

In this paper, we investigated a novel approach to estimating attribute-specific brand perceptions from social media to provide a low-cost, real-time alternative to traditional elicitation methods. We validated our estimates against survey data for three attributes and over two hundred brands. With an average correlation coefficient of .72, the results indicate that the method provides a reliable means for automatically estimating attribute-specific brand ratings. These results appear robust to a number of alternative choices of exemplar selection and affinity metric.

In addition to being, to the best of our knowledge, the first data mining attempt in the literature that addresses this important goal, the approach is innovative in several ways. First, we use social connections to infer brand image. While most extant data mining methods developed for other brand perception goals focus on analyzing the text of user-generated content about the brand, we instead consider an alternate source of information—the social connections of a brand's supporters. Prior research has shown that a brand's image is deeply connected to its social media network, and by analyzing these network connections, we can exploit the social network positions of millions of consumers, the majority of whom do not actively author content, to inform brand image insights. This allows insights to be gained for topics that consumers do not write about concurrently with brand mentions, reduces bias in the data, and allows for more efficient analysis.

Second, we focus our analysis on the social media platform Twitter. Though few studies in the marketing literature have focused on Twitter, the platform is highly used by marketers for brand image marketing, and thus is well-suited for mining consumer perceptions. Additionally, the open API maintained by Twitter ensures that the data required by this method can be easily accessed by researchers and practitioners desiring to implement it.

Third, we provide a fully automated and highly generalizable method. Extant data mining approaches in the marketing literature require context-specific manual tuning and/or data-annotating to implement, which can be as or more costly and time consuming as the manual direct-elicitation methods they aim to replace. By leveraging the user-generated account organization of Twitter, the method we present can automatically identify exemplar accounts for an attribute of interest to researchers based on a single keyword input. This automation allows researchers and practitioners to generate estimates for a range of

attributes and brands of interest frequently and easily, allowing for rich marketing insights that can be kept up-to-date over time.

The use-cases for the developed method are many. Most directly, marketing practitioners can use the method to automatically generate and update perceptual maps for large sets of brands for numerous attributes of interest. Such perceptual maps have long been a "major analytical tool in marketing research" (Steenkamp et al. 1994), and, by reducing the need to directly elicit perceptions from consumers, the proposed method can enable richer and more varied market studies that incorporate data from a much larger set of users, and can be continuously kept up to date to monitor evolving perceptions. Marketing researchers can also use the method to easily extract perception data to be used to inform more substantiative marketing research analyses and models (for example, to study how brand perceptions change in response to marketing initiatives).

We also speculate that the approach can be adopted to other common marketing tasks beyond those explored here. For example, instead of computing brand-exemplar affinities, one could compute brand-brand affinities using the same data. Clustering the resulting weighted graph may be used to generate competitive market structures and brand associative networks (Henderson et al. 1998, Netzer et al. 2012, Urban et al. 1984). Similarly, clustering consumers instead of brands may aid in marketing segmentation and personalization.

We note several limitations with this work. First, as presented, it is inherently limited to analyzing brands that maintain a Twitter presence. While we have found this to be the majority of brands, there are undoubtedly some that cannot be analyzed this way. We note that the general approach of using a brand's social connections to infer brand perceptions is versatile across platforms and social connection types. Though we use Twitter "follow" relationships, Facebook "fan" relationships could also be used to the extent that they can be publicly mined. Extending further conceptually, networks based on authorship of product reviews on Amazon or Yelp could be used; product-level perceptual attribute ratings could potentially be inferred through the degree to which a product's reviewers also review products that are known exemplars of those attributes. Analyzing comment authorship on blog or news platforms could be used in a similar manner. We encourage future researchers to pursue these directions and hope that the general method we introduce provides a foundation for enabling richer consumer insights across many domains.

Another limitation is that attributes may vary in the reliability of the estimates they provide via this method. Indeed, some may not be amenable to the automated version of the method at all, if an appropriate keyword cannot be identified or if the "Lists" returned for the keyword are not of high quality. We found the automated method to work well for the three attributes tested, and also discuss how an exemplar set can be constructed manually to increase accuracy, noting that the manual curation of a list of accounts is likely to be substantially less work compared to the manual curation tasks typically required in data mining applications. We encourage future research to explore methodological advances that can further increase generalizability.

Finally, while the simplicity of our approach has benefits in terms of transparency and implementability by practitioners, we encourage future researchers to develop this work further and investigate more nuanced approaches. For example, second degree connections can be explored, or the social network analyses can be supplemented with text analyses to improve accuracy and versatility, and to gain deeper theoretical understanding into the nature of consumer-brand social network relationships. Additionally, given our analysis on the varying quality of exemplars, future work may consider hybrid approaches that use a small amount of survey data to guide selection or weighting of exemplars.

Overall, it is our hope that the methods introduced in this paper provide a useful tool for marketing researchers and practitioners interested in automatically monitoring brand image perceptions, and also provide a foundation for future research advances in understanding the nature of brand-follow relationships, exploiting social media structure to more fully automate data mining algorithms, and using social network data to gain insights about consumers and brands.

Acknowledgments

Appendix A: Collecting Twitter Accounts of a Brand

Our analysis requires a Twitter user name for each brand under consideration, which is straightforward to obtain by manually searching Twitter.com. However, to aid scalability, we describe the semi-automated procedure we use here. We write a script that searches Yahoo.com using as a query the brand name and the word Twitter. The script then checks if the first result is from Twitter.com. If so, it returns the suffix of the URL as the user name (e.g., `http://twitter.com/patagonia`). To validate the retrieved account, we write a second script to find the homepage of the brand. Heuristic brand name combinations are used to find active candidate websites, and the title tag of the page is checked for relevance to the sector. We also search for links to their Twitter accounts from their homepage, and compare with the user names determined through the Yahoo! search. For sites that do not link to a Twitter account, we search the Twitter profile for links to the home page. This process produced valid Twitter accounts for approximately 80% of accounts we considered. For the purposes of this study, we additionally perform a manual validation; however, we offer the approach above to enable larger studies.

Appendix B: Regression with SPS and Brand Demographics

One interpretation of the high correlations between SPS and average survey scores could be that they are capturing demographic differences in followership, rather than attribute-specific perceptions. To investigate this, we collected gender, age, and income profiles for a subset of brands in our test set that had website user demographic information publicly available from Compete.com (N = 101). For each attribute, we performed an OLS regression according to the following equation:

$$Survey = \beta_0 + \beta_1 SPS + \beta_2 Gen + \beta_3 Age_1 + \beta_4 Age_2 + \beta_5 Age_3 + \beta_6 Age_4 + \beta_7 Inc_1 + \beta_8 Inc_2 + \beta_9 Inc_3 + \epsilon$$

Survey is the brand's average survey rating for the specified attribute. The demographic predictors indicate the percentage of a brand's website traffic that comes from distinct gender, age, and income brackets, as reported by Compete.com. Specifically, Gen is the percentage of the brand's website users that are male; Age_{1-4} are the percentages that are 18-24, 25-34, 35-44, and 55-64 years old; and Inc_{1-3} are the percentages with annual household incomes of \$0-\$29,999, \$30,000-\$59,000, and \$60,000-\$99,999. All variables were standardized ($mean = 0$, $SD = 1$), and the model was fit separately for each attribute.

The resulting estimates for β_1 are positive and highly significant for all attributes (for Eco, Nut, and Lux, respectively, $\beta_1 = \{.81, .62, .77\}$, $SE = \{.12, .16, .15\}$, $\{p < .0001, p < .01, p < .0001\}$). Demographic predictors are only significant at the .05 level in two cases: a more male audience positively predicts luxury ratings ($\beta_2 = .52$, $SE = .14$, $p < .001$), and a more youthful audience positively predicts eco-friendliness ratings ($\beta_3 = .20$, $SE = .08$, $p < .05$). The R^2 measures for the models ($\{.44, .72, .54\}$ for \{Eco, Nut, Lux\}), are larger by an average of .29 compared to those for models run with just the demographic predictors. These results support that SPS is capturing attribute perceptions beyond demographic similarities.

Appendix C: Twitter and Survey Demographics

According to a recent report by Pew Research, approximately twenty percent of U.S. adults used Twitter in 2014, and use is growing year over year. It is particularly popular with individuals who are college-educated,

32

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

under fifty, and earn more than $50,000 a year, though popularity among other groups, such at those aged 65+, is steadily increasing (Duggan et al. 2015). Beyond its general popularity, Twitter is an ideal platform for our goal of developing a methodology for mining brand image perceptions as, among social media platforms, it is used heavily for brand marketing; its follow connections are publicly accessible; and accounts are user-organized through the Lists features into topic-relevant sets. While many brand managers might consider monitoring the perceptions of Twitter brand communities (many of which have millions of members) to be a worthwhile goal in itself, a natural question that arises is whether brand perceptions estimated in this way can generalize beyond Twitter users. We investigate this here.

We first note that the current "gold standard" for measuring brand perceptions is asking consumers through surveys, and this, itself, is subject to selection bias, as only the opinions of individuals who agree to complete the survey can be counted. We also note that the "general" population to generalize to is somewhat ill-defined—we are not interested in customers only, per se, but the aggregate perceptions of the greater community for which customers might be drawn. With these limitations in mind, we explore the extent to which bias in our presented methodology might be cause for concern.

Our first step is in validating the Twitter-based perception estimates with surveys administered to a different population—specifically, participants recruited through Amazon Mechanical Turk (AMT). AMT has been shown to be a reliable source for collecting social science research data, comparable to traditional laboratory methods (e.g., see Buhrmester et al. (2011), Mason and Suri (2012), Sprouse (2011)). As shown in our results section, we find encouragingly high correlations between the Twitter-based SPS and average AMT ratings.

Although AMT is generally considered a reliable source for social science data, we asked our survey participants to identify categories for their gender, age, education, and household income so that we could examine demographic bias in our sample. Figure 10 shows the distributions of the survey respondents along these variables.

The gender distribution of the survey sample is fairly balanced (with slightly more males) and the median income category is $30,000-$59,999, which maps to the U.S. median income of $51,915[12]. Similar to the Twitter population, the sample is notably young (with large representation in the 25-34 range, and few seniors) and educated (more than half have a college degree).

We next investigate whether differences in such demographic categories affect the measure of interest—i.e., ratings of brand perceptual attributes. To explore this, we computed the mean rating of each brand (for each perceptual attribute) by each demographic category, and examine the correlations between the average ratings of the different demographic groups (e.g., we computed the correlation coefficient between the average luxury rating for each brand by participants aged 18-24, with the average luxury ratings by participants aged 65+, and so on). The Pearson correlation matrices showing the coefficients for all ninety-six pairs (32 demographic category pairs times 3 attributes) are provided in Table 5.[13]

---

[12] http://www.census.gov/content/dam/Census/library/publications/2014/acs/acsbr13-02.pdf

[13] Due to the small sample size for Education=Less than High School Degree, we combined that category with High School Degree of Equivalent to simplify to a single No College category.
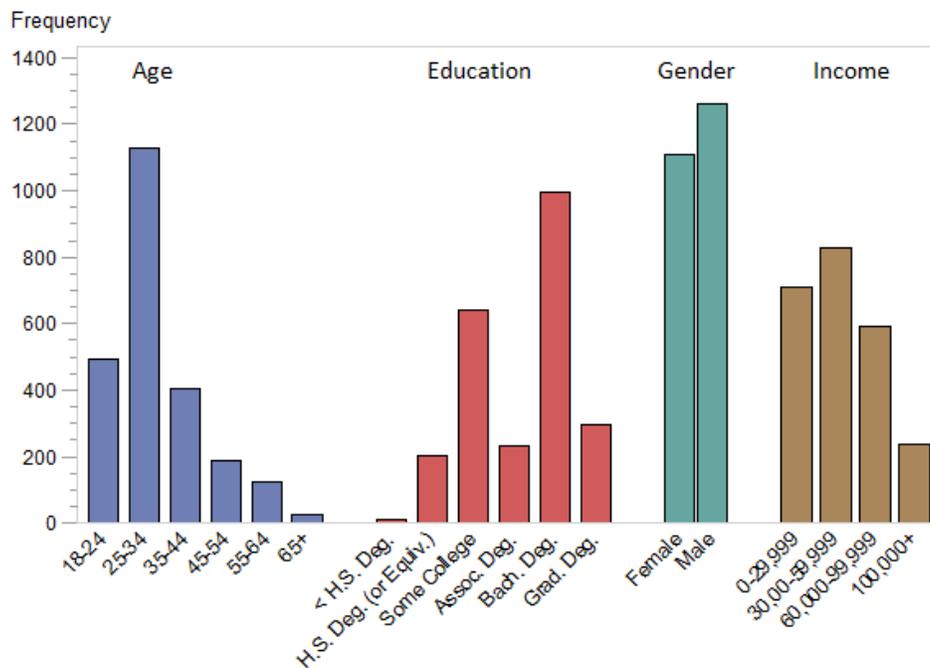
Figure 10: Demographics of survey respondents.

We computed the correlations between ratings by different demographic categories rather than constructing a regression problem to identify main effects of demographic categories on rating level because we are interested in detecting major shifts in the relative ratings of different brands within the set, rather than linear shifts in baseline rating levels. We find strong correlations across all demographic categories, with an average $r = 0.91$, and $p < .0001$ in every case. For 90% of the demographic category pairs, $r > 0.8$. The smallest correlations are for the Eco-Friendliness attribute between $Age = 65+$ and the other Age categories, with a still strong but lower $r = 0.65 - 0.73$. We note that this category has a particularly small sample size, which may be affecting these results.

This analysis provides evidence that attribute-specific brand perceptions are similar across standard marketing demographic categories. Thus, we conclude that moderate demographic bias in our proposed methodology is unlikely to notably bias perception estimates, though we encourage further investigation for those particularly interested in the perceptions of the senior community. We also note that there may be other (non-demographic) sources of bias in our sample; both Twitter users and AMT workers may, for example, spend more time online than the average U.S. consumer. We suspect that any such biases are no stronger than the biases inherent to any survey used for marketing research, but we encourage future researchers to explore this further.

	Luxury					Eco-Friendliness					Nutrition				
Age	18-24	25-34	35-44	45-54	55-64	18-24	25-34	35-44	45-54	55-64	18-24	25-34	35-44	45-54	55-64
25-34	0.98					0.91					0.97				
35-44	0.97	0.99				0.86	0.93				0.98	0.98			
45-54	0.95	0.95	0.95			0.82	0.85	0.88			0.94	0.94	0.95		
55-64	0.92	0.92	0.92	0.95		0.83	0.82	0.81	0.74		0.92	0.93	0.94	0.96	
65+	0.87	0.89	0.88	0.9	0.89	0.65	0.65	0.66	0.73	0.69	0.82	0.82	0.83	0.82	0.83

Edu.	No Col.	Sm. Col.	Assoc.	Bach.		No Col.	Sm. Col.	Assoc.	Bach.		No Col.	Sm. Col.	Assoc.	Bach.	
Some Col.	0.96					0.74					0.97				
Associate	0.94	0.93				0.74	0.81				0.95	0.97			
Bachelor	0.96	0.99	0.95			0.74	0.93	0.8			0.96	0.98	0.97		
Grad	0.96	0.97	0.97	0.98		0.76	0.91	0.82	0.91		0.95	0.97	0.97	0.98	

Income	0-30k	30-60k	60-100k			0-30k	30-60k	60-100k			0-30k	30-60k	60-100k		
30k-59,999	0.98					0.94					0.99				
60k-99,999	0.98	0.99				0.92	0.93				0.98	0.98			
100k+	0.98	0.99	0.98			0.91	0.92	0.89			0.98	0.98	0.97		

Gender	Male					Male					Male				
Female	0.99					0.94					0.99				

Table 5: Pearson correlation matrices (compressed for readability) for the mean brand ratings by attribute per demographic category.

References

Aaker, David A. 1996. Measuring brand equity across products and markets. *California Management Review* **38**(3) 103.

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

Baird, Carolyn Heller, Gautam Parasnis. 2011. From social media to social CRM: reinventing the customer relationship. *Strategy & Leadership* **39**(6) 27–34.

Barnes, NG, AM Lescault, S Wright. 2013. Fortune 500 are bullish on social media. *Charlton College of Business Center for Marketing Research, University of Massachusetts Dartmouth* URL `http://www.umassd.edu/cmr/socialmediaresearch/2013fortune500/`. Accessed: 2013-12-07.

Bearden, William O., Michael J. Etzel. 1982. Reference group influence on product and brand purchase decision. *Journal of Consumer Research* **9**(2) 183–94.

Berens, G., C. B. M. Van Riel, G. H. Van Bruggen. 2005. Corporate associations and consumer product responses: the moderating role of corporate brand dominance. *Journal of Marketing* **69**(3) 35–48.

Berger, Jonah, Chip Heath. 2007. Where consumers diverge from others: Identity signaling and product domain. *Journal of Consumer Research* **34**(2) 121–134.

Bijmolt, Tammo HA, Michel van de Velden. 2012. Multiattribute perceptual mapping with idiosyncratic brand and attribute sets. *Marketing Letters* **23**(3) 585–601.

Bottomley, Paul A., John R. Doyle, Rodney H. Green. 2000. Testing the reliability of weight elicitation methods: Direct rating versus point allocation. *Journal of Marketing Research* **37**(4) 508–513.

Buhrmester, Michael, Tracy Kwang, Samuel D Gosling. 2011. Amazon's mechanical turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science* **6**(1) 3–5.

Cha, Meeyoung, Hamed Haddadi, Fabricio Benevenuto, Krishna P. Gummadi. 2010. Measuring user influence in Twitter: The million follower fallacy. *Proceedings of the fourth international AAAI conference on weblogs and social media* 10–17.

Childers, Terry L., Akshay R. Rao. 1992. The influence of familial and peer-based reference groups on consumer decisions. *Journal of Consumer Research* **19**(2) 198–211.

Crystal, David. 2001. *Language and the Internet*. Cambridge University Press, New York.

Danaher, Peter J, Isaac W Wilson, Robert A Davis. 2003. A comparison of online and offline consumer brand loyalty. *Marketing Science* **22**(4) 461–476.

Das, Sanjiv R, Mike Y Chen. 2007. Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science* **53**(9) 1375–1388.

Day, George S. 1975. The threats to marketing research. *Journal of Marketing Research* **22** 462–467.

36

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

Day, George S, Allan D Shocker, Rajendra K Srivastava. 1979. Customer-oriented approaches to identifying product-markets. *The Journal of Marketing* **43**(4) 8–19.

Dillon, William R, Donald G Frederick, Vanchai Tangpanichdee. 1985. Decision issues in building perceptual product spaces with multi-attribute rating data. *Journal of Consumer Research* **12**(1) 47–63.

Duggan, Maeve, Nicole B. Ellison, Cliff Lampe, Amanda Lenhart, Mary Madden. 2015. Social media update 2014. *Pew Reseach Center* URL `http://www.pewinternet.org/2015/01/09/social-media-update-2014`.

El Gazzar, N., M. Mourad. 2012. The effect of online communication on corporate brand image. *International Journal of Online Marketing* **2**(1) 1–15.

Escalas, Jennifer Edson, James R. Bettman. 2003. You are what they eat: The influence of reference groups on consumers connections to brands. *Journal of Consumer Psychology* **13**(3) 339–348.

Etter, Michael, Thomas Plotkowiak. 2011. CSR communication strategies for Twitter. URL `http://hdl.handle.net/10398/8529`. Accessed 2013-12-06.

Fader, Peter S., Russell S. Winer. 2012. Introduction to the special issue on the emergence and impact of user-generated content. *Marketing Science* **31**(3) 369–371.

Goel, Sharad, Daniel G Goldstein. 2013. Predicting individual behavior with social networks. *Marketing Science* **33**(1) 82–93.

Goel, Sharad, Duncan J Watts, Daniel G Goldstein. 2012. The structure of online diffusion networks. *Proceedings of the 13th ACM Conference on Electronic Commerce*. ACM, 623–638.

Grabowicz, Przemyslaw A., Jos J. Ramasco, Esteban Moro, Josep M. Pujol, Victor M. Eguiluz. 2012. Social features of online networks: The strength of intermediary ties in online social media. *PLoS ONE* **7**(1) e29358.

Green, Paul E. 1975. Marketing applications of mds: Assessment and outlook. *Journal of Marketing* 24–31.

Green, Paul E, Frank J Carmone, Scott M Smith. 1989. *Multidimensional scaling: concepts and applications*, vol. 18. Boston: Allyn and Bacon.

Hauser, John R, Frank S Koppelman. 1979. Alternative perceptual mapping techniques: Relative accuracy and usefulness. *Journal of Marketing Research* **14**(4) 495–506.

Hauser, John R, Patricia Simmie. 1981. Profit maximizing perceptual positions: An integrated theory for the selection of product features and price. *Management Science* **27**(1) 33–56.

Henderson, Geraldine R, Dawn Iacobucci, Bobby J Calder. 1998. Brand diagnostics: Mapping branding effects using consumer associative networks. *European Journal of Operational Research* **111**(2) 306–327.

Huber, Joel, Morris B Holbrook. 1979. Using attribute ratings for product positioning: some distinctions among compositional approaches. *Journal of Marketing Research* **16**(4) 507–516.

John, Deborah Roedder, Barbara Loken, Kyeongheui Kim, Alokparna Basu Monga. 2006. Brand concept maps: a methodology for identifying brand association networks. *Journal of Marketing Research* **43**(4) 549–563.

Johnson, Michael D, Elania J Hudson. 1996. On the perceived usefulness of scaling techniques in market analysis. *Psychology & Marketing* **13**(7) 653–675.

Kaul, Anil, Vithala R Rao. 1995. Research for product positioning and design decisions: An integrative review. *International Journal of Research in Marketing* **12**(4) 293–320.

Kim, Angella J, Eunju Ko. 2012. Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand. *Journal of Business Research* **65**(10) 1480–1486.

Kuksov, D., R. Shachar, K. Wang. 2013. Advertising and consumers' communications. *Marketing Science* **32**(2) 294–309.

Kwon, Eun Sook, Yongjun Sung. 2011. Follow me! Global marketers' Twitter use. *Journal of Interactive Advertising* **12**(1) 4–16.

Lancaster, Kelvin. 1971. *Consumer demand: A new approach*. New York: Columbia University Press.

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

Lehmann, Donald R, Kevin Lane Keller, John U Farley. 2008. The structure of survey-based brand metrics. *Journal of International Marketing* **16**(4) 29–56.

Ludwig, Stephan, Ko de Ruyter, Mike Friedman, Elisabeth C Brüggen, Martin Wetzels, Gerard Pfann. 2013. More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing* **77**(1) 87–103.

Lydon, John E., David W. Jamieson, Mark P. Zanna. 1988. Interpersonal similarity and the social and intellectual dimensions of first impressions. *Social Cognition* **6**(4) 269–86.

Manning, Christopher D, Prabhakar Raghavan, Hinrich Schütze. 2008. *Introduction to information retrieval*, vol. 1. Cambridge: Cambridge University Press.

Mason, Winter, Siddharth Suri. 2012. Conducting behavioral research on Amazons Mechanical Turk. *Behavior Research Methods* **44**(1) 1–23.

McDaniel, Stephen W, Perry Verille, Charles S Madden. 1985. The threats to marketing research: An empirical reappraisal. *Journal of Marketing Research* **22**(1) 74–80.

McPherson, Miller, Lynn Smith-Lovin, James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* **27**(1) 415–444.

Morry, Marian M. 2007. Relationship satisfaction as a predictor of perceived similarity among cross-sex friends: A test of the attraction-similarity model. *Journal of Social and Personal Relationships* **24**(1) 117–38.

38

Culotta and Cutler: *Mining Brand Perceptions from Twitter Social Networks*
Article submitted to *Marketing Science*; manuscript no. (Please, provide the manuscript number!)

Naylor, Rebecca Walker, Cait Poynor Lamberton, Patricia M West. 2012. Beyond the "like" button: The impact of mere virtual presence on brand evaluations and purchase intentions in social media settings. *Journal of Marketing* **76**(6) 105–120.

Netzer, Oded, Roneb Feldman, Jacod Goldenberg, Moshe Fresko. 2012. Mine your own business: Market-structure surveillance through text mining. *Marketing Science* **31**(3) 521–543.

Pan, Ying, De-Hua Li, Jian-Guo Liu, Jing-Zhang Liang. 2010. Detecting community structure in complex networks via node similarity. *Physica A: Statistical Mechanics and its Applications* **389**(14) 2849–2857.

Pennebaker, James W, Cindy K Chung, Molly Ireland, Amy Gonzales, Roger J Booth. 2007. *The development and psychometric properties of LIWC2007*. LIWC.Net, Austin, TX.

Pereira, Hélia Gonçalves, Maria de Fátima Salgueiro, Inês Mateus. 2014. Say yes to Facebook and get your customers involved! Relationships in a world of social networks. *Business Horizons* **57**(6) 695–702.

Peters, Isabella. 2009. *Folksonomies: Indexing and retrieval in Web 2.0*, vol. 1. Walter de Gruyter.

Schmalensee, Richard, Jacques-Frangois Thisse. 1988. Perceptual maps and the optimal location of new products: An integrative essay. *International Journal of Research in Marketing* **5**(4) 225–249.

Shocker, Allan D, Venkataraman Srinivasan. 1979. Multiattribute approaches for product concept evaluation and generation: A critical review. *Journal of Marketing Research* **16** 159–180.

Sonnier, Garrett P, Leigh McAlister, Oliver J Rutz. 2011. A dynamic model of the effect of online communications on firm sales. *Marketing Science* **30**(4) 702–716.

Sprouse, Jon. 2011. A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods* **43**(1) 155–167.

Steenkamp, Jan-Benedict, Hans Van Trijp. 1997. Attribute elicitation in marketing research: A comparison of three procedures. *Marketing Letters* **8**(2) 153–165.

Steenkamp, Jan-Benedict E. M., Hans C. M. Van Trijp, Jos M. F. Ten Berge. 1994. Perceptual mapping based on idiosyncratic sets of attributes. *Journal of Marketing Research* **31**(1) 15–27.

Stephen, Andrew, Yaniv Dover, Jacob Goldenberg. 2010. A comparison of the effects of transmitter activity and connectivity on the diffusion of information over online social networks. *INSEAD Working Paper* URL http://ssrn.com/abstract=1609611. Accessed: 2013-12-06.

Tang, Chuanyi, Lin Guo. 2013. Digging for gold with a simple tool: Validating text mining in studying electronic word-of-mouth (eWOM) communication. *Marketing Letters* **26**(1) 67–80.

Tirunillai, Seshadri, Gerard J Tellis. 2012. Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science* **31**(2) 198–215.

Toubia, Olivier, Andrew T Stephen. 2013. Intrinsic vs. image-related utility in social media: Why do people contribute content to Twitter? *Marketing Science* **32**(3) 368–392.

Urban, Glen L., P. L. Johnson, John R. Hauser. 1984. Testing competitive market structures. *Marketing Science* **3**(2) 83–112.

Wu, Shaomei, Jake M Hofman, Winter A Mason, Duncan J Watts. 2011. Who says what to whom on Twitter. *Proceedings of the 20th international conference on World Wide Web*. ACM, 705–714.