Using online social networks to measure consumers' brand perception

Received (in revised form): 9th December, 2016

Jennifer Cutler
is an assistant professor of marketing at Northwestern University's Kellogg School of Management. Her research focuses on using social media to understand consumer behaviour and the effects of marketing communications. She received her PhD in business administration from Duke University with a specialisation in quantitative marketing.

Northwestern University, Kellogg School of Management, 2001 Sheridan Rd., Evanston, IL 60208, USA
Tel: +1 847 491 3522; E-mail: jennifer.cutler@kellogg.northwestern.edu

Aron Culotta
is an assistant professor of computer science at the Illinois Institute of Technology in Chicago, where he leads the Text Analysis in the Public Interest lab. His research focuses on extracting socially valuable insights from online social networks. He is a former Microsoft Live Labs Fellow with a PhD in computer science from the University of Massachusetts, Amherst. More than 40 of his research articles have been published, and his work has received best paper awards at the Association for the Advancement of Artificial Intelligence conference and the Conference on Computer-Supported Social Work and Social Computing.

Department of Computer Science, Illinois Institute of Technology, Chicago, IL 60616, USA
Tel: +1 312 567 5261; E-mail: aculotta@iit.edu

Abstract The ability to measure and monitor specific dimensions of brand image has a range of useful applications in marketing, from developing competitive strategy to identifying strength and weaknesses to evaluating the effectiveness of marketing initiatives. Nevertheless, obtaining reliable measurements is an ongoing challenge for marketers. Traditional methods such as administering surveys can be expensive and biased, and are limited in scale, both in terms of the number of brands and dimensions that can be tracked, and the frequency with which the measurements can be updated. The explosion of social media in recent years has created an enormous secondary data trail that is available for analysis. However, the most common analytics approaches, such as those that rely on user-generated text, are difficult to apply due to the scarcity of relevant conversations, as well as the ambiguity, variety, and often rapid changes in linguistic terms used by consumers. This paper describes a recent advance in marketing science that makes use of brand social network connections to make highly scalable inferences about brand image. This promising new approach provides many potential advantages, including the ability to fully automate monitoring for a large number of brands over a wide range of dimensions.

KEYWORDS: social media, brand image, social networks, Big Data, perceptual maps

INTRODUCTION

Understanding how consumers perceive different brands within a competitive set is integral to many marketing goals. However, brand image can comprise a wide range of dimensions relevant to sales. In addition to traditional attributes such as quality and price,[1] consumers consider aspects of a brand's personality, such as its sincerity and sophistication,[2] and, increasingly, the brand's alignment with social causes such as environmental friendliness, health and social responsibility.[3,4] For example, in a 2015 survey of 30,000 global consumers administered by the Nielsen Company, nearly three-quarters of millennials — and two-thirds of consumers overall — indicated a willingness to pay a premium for environmentally responsible products and services.[5]

Measuring and monitoring such increasingly dimensional aspects of brand image presents a real challenge for marketers. By and large, marketers have relied upon surveys or choice tasks administered to a sample of customers to measure brand perceptions. However, reliance on acquiring consumer responses hinders measurement capabilities: surveys are costly to administer; respondent pools are often sparse and deplete quickly, particularly for certain demographics; participants may be unable or unwilling to reveal their true beliefs; participant attention may wane in the face of too many questions; and results may become outdated quickly, particularly if there is a shock or campaign to shift brand image.[6–9]

The explosion of social media in the past decade has raised hope among many marketers that the 'Big Data' trail on platforms such as Twitter, Facebook, and Amazon could be mined to uncover richer and more scalable insights about consumer behaviour and perceptions. While much has been accomplished on this front, managers still frequently find themselves unable to extract meaningful, reliable and actionable insights from the sea of available data.

The goal of this paper is to describe a promising new methodological development in the space — the use of a brand's social media network connections to estimate the strength of specific dimensions of brand image, such as eco-friendliness.[10] The paper will provide both a general conceptual overview of this new approach, which can be flexibly adapted and implemented for a range of goals, as well as detailed instructions for how to utilise Twitter's application programming interface (API) to generate perceptual maps automatically.

WHY USE SOCIAL NETWORKS?

To many marketers, 'mining social media data' is synonymous with 'mining user-generated text'. Indeed, there is a lot of value to be gained from looking at what consumers are writing about brands in online spaces (for an overview, see Fader and Winer[11]). However, there are limitations to relying on user text. On many social media platforms, fewer than half the users write their own content;[12] fewer still write about the brands to be monitored; and even fewer write about brands in conjunction with topics or attributes of interest. Yet, every user who connects with a brand via an online platform (whether by following the brand on Twitter, liking the brand on Facebook, or even by liking or sharing a brand's post) provides information by their voluntary 'mere virtual presence'[13] in that online brand community. While liking or following a brand is not always indicative of affinity for the brand, it appears to be the case most of the time.[14,15] Furthermore, each member of a brand's online community is likely to be a part of many other online communities (ie to follow other accounts). By tracing network relationships to learn more about who a brand's fans are — what they value and are interested in — one can gain insights about brand image that remain invisible in user-generated text.

CONCEPTUAL APPROACH

The general approach to measuring the strength of association between a brand and a topic or cause of interest is to look at the communities forming online around the brand and the topic, and to measure the similarity, or overlap between these communities.

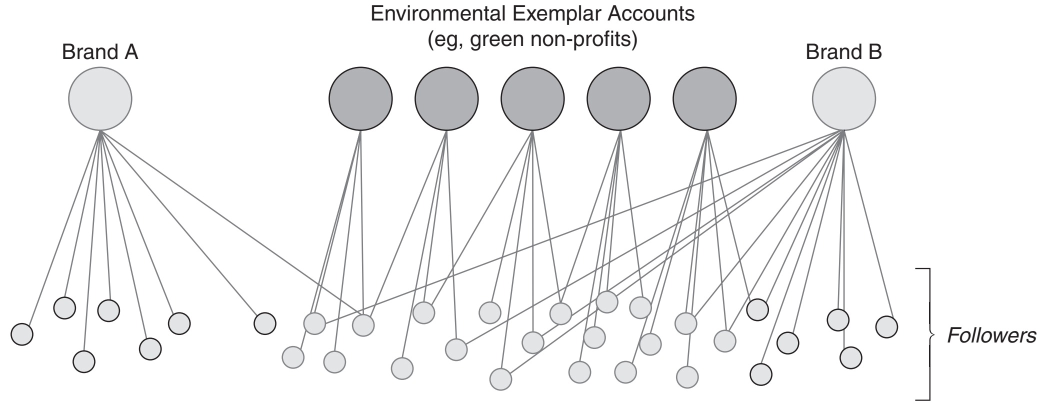

Figure 1: Here, the overlap between follower community and environmental community is greater for Brand B than for Brand A; one can therefore infer that Brand B is likely to have a stronger reputation for environmental friendliness than Brand A

For example, if a manager is interested in comparing the relative perceived eco-friendliness of two brands (Brand A and Brand B), they can:

- identify a sample of social media followers of Brand A and Brand B;
- identify a sample of social media followers of organisations, such as Greenpeace or the Conservation Fund, that exemplify interest in eco-friendliness (how to identify such accounts will be discussed in the next section); and
- measure which brand's community overlaps more with the eco-friendly community (a range of network similarity metrics can be used and will be discussed in the next section).

This approach is conceptually illustrated in Figure 1.

The approach rests on a simple assumption: that brand perceptions are reflected in the brand's followership. More specifically, it assumes that a specific dimension of brand image is reflected in the overlap between a brand's followership and the followership of accounts known to be perceived as strong in that dimension. There are many reasons to think that this might not work — that follower overlap would not reliably reflect perceptions. One objection is that not *all* follow relationships indicate affinity; a devoted environmentalist, for example, might follow the accounts of some organisations known to be environmentally unfriendly in order to keep abreast of news. Another is that there may not always be overlap in the types of organisations individuals follow; some environmentally conscious consumers, for example, might not follow any environmental 'exemplar' accounts. Yet despite the noise created by such cases, the results obtained through this method have been shown to have a high correlation with directly-elicited survey responses[16] — while being easier and less expensive to obtain. Thus, while individual motivations for following brands can be varied and complex, the aggregate signal obtained over millions of follower relationships appears to be quite informative.

METHOD

The social media platform Twitter is ideal for this kind of analysis because it is widely used by marketers and consumers to build brand communities, and because it has an open API that allows relevant information about social networks to be programmatically extracted for free using common scripting languages such as

Python. This section will describe a specific implementation of this approach using Twitter; however, the conceptual approach could be applied many ways, including on other platforms.

How to extract brand followers

To begin, a list of Twitter handles for the brands to be monitored is needed. Presumably, marketing managers are already aware of the handles for their brands and their primary competitors. Handles for additional brands can be easily searched for manually, either on Twitter itself, or on the brand's website. Once a list of brand handles has been created, the Twitter Search API can be invoked to create a list of user IDs (or screen names) for the *n* most recent followers of each brand.

As of the writing of this paper, access to the Twitter Search API is free and well documented.[17] However, Twitter limits the number of user IDs that can be extracted every 15 minutes. Many popular brands have millions of Twitter fans: for example, Pepsi, Whole Foods and McDonalds each have over 3 million followers, and celebrity brands such as Britney Spears and Kim Kardashian approach 50 million. Due to Twitter's rate limit, collecting *all* of a brand's followers may take a prohibitively long time for such popular icons; fortunately, studies have shown that accurate brand perception measures can be obtained using just samples in the range of few hundred thousand followers.[18] Capping follower collection also ensures that it is the most recent followers, whose follow connections are likely most indicative of current brand perceptions, that are factored into the measure, as Twitter returns account followers in reverse chronological order of follow data.

How to identify exemplar accounts

The next step is to identify a set of Twitter accounts for organisations that exemplify the attribute of brand image to be measured. The most straightforward way is to use domain knowledge to identify a list of exemplar organisations (potentially with the aid of a committee of experts or a written guide) and to look up each organisation's Twitter handle manually. Many aids exist for such tasks. For example, CharityNavigator.org maintains lists of nonprofit organisations, organised by the primary cause they support; these lists can be used to identify high-quality exemplars for causes such as environmental friendliness or social responsibility. In many cases, however, an automated approach to exemplar identification is feasible and preferable to such manual curation; details on automating the process are described next.

Fully automating the process

In many cases, manually curating high-quality exemplar lists may be a difficult and labour-intensive task (particularly if a manager is interested in tracking many perceptual attributes). Furthermore, the accounts that most exemplify an attribute of interest may change over time, requiring the lists to be updated. Fortunately, the crowd-organisation of Twitter can be leveraged to automatically identify accounts that users have already identified as relevant to an attribute of interest.

Since 2009, Twitter has maintained a 'Lists' feature[19] through which users can organise the accounts they follow into topic-cohesive lists. This allows them to create different newsfeeds that map to different interests, and to share lists with other users who are interested in the same topics. Managers can leverage this crowd-organisation by using a search engine to search through Twitter Lists for a keyword of interest (this search process can be automated through scripting). Keywords can range from a single word to a longer, more specific phrase; managers can vary the query according to their goals. Searching for a keyword will return a set of user-curated lists that map to that keyword, and each list will contain a set of accounts. Managers can take the top y lists returned for a keyword, and retain as exemplars accounts that appear on at least z lists. Requiring that accounts be on more than one relevant list reduces the risk of false positives, ie accounts idiosyncratic to a particular user, that do not accurately reflect a more general association between the account and the keyword-specified attribute. Finally, if any of the brand accounts (from the set of brands for which brand image is to be measured) happen to appear in the exemplar set, they should be eliminated from the exemplar set.

For example, a search through Twitter Lists for 'environment' returns hundreds of lists users have created that are relevant to that term. Retaining accounts that appear in at least two of the top 50 lists leads to a set of 74 exemplar accounts, with examples including @GreenPeace, @SierraClub and @Epa.

There are many benefits to this automated approach. First, it enables the identification of accounts that consumers see as relevant to a topic, but that might not be known to those doing the measurement (this is particularly relevant where consumers are of a different demographic from the marketing researchers). This is likely to improve the quality of resulting brand image measures, as small, esoteric exemplars — the kind perhaps less likely to be known by non-enthusiasts — tend to be the most informative. Secondly, it opens the door to scalability. With one script, exemplar accounts can be identified for a wide range of brand image dimensions, and exemplars can be automatically updated over time as the perceptual landscape changes. However, the approach has limitations as well. The quality of accounts returned for a given keyword query (that is, the extent to which those accounts have a followership that values the brand image dimension represented by the keyword) may vary and is *a priori* unknown. Studies have shown success with keywords such as 'environment', 'nutrition' and 'luxury' to identify exemplars of environmental friendliness, nutrition and luxury,[20] but have not yet probed the boundaries of when keyword-based matches break down (for example, for very broad terms such as 'quality') — or how they can be improved (for example, with sector-specific keyword additions). Managers are encouraged to experiment with the queries and selection algorithms to find implementations best suited to their specific goals.

How many exemplar accounts are needed?

For more specific keyword queries, or for manual exemplar curation, it may be difficult to obtain large numbers of exemplars. Of course, the number of exemplars needed is not independent from the dimension being tested or the quality of exemplars identified. Studies have shown success with exemplar sets ranging from 30 to 400, and the marginal increments in accuracy for adding additional exemplars seems to plateau quickly.

How to compute network similarity

Once exemplar accounts have been identified, a list of user IDs for the followers of each exemplar account can be collected following the same process as described for the brand accounts. At this point, each brand and each exemplar has a list of followers associated with it. The goal is now to create a single quantitative measure that indicates the 'similarity' between each individual brand's community and the full set of exemplar communities.

There are a wide range of standard network similarity metrics that could be employed at this point; a range of tests suggests that many different metrics will produce substantively similar results. This paper will describe in detail one implementation that has worked particularly well.

To quantify the perceived strength of association between a brand B_x and an attribute A_k (for which there are identified exemplar accounts E_{ki} to E_{kj}), the first step is to compute the similarity between B_x and each exemplar E_{ki} to E_{kj} individually. Because the brands in the set being monitored may vary a great deal in the number of followers they have, it is important not to rely simply on the raw number of followers that overlap — such counts are likely to be higher for extremely popular brands with millions of followers. Instead, one can employ a common measure of set similarity called the Jaccard index, which is defined as the size of the intersection of two sets divided by the size of the union of the same two sets. More formally:

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

Thus, $\mathrm{J}(B_x, E_{ki})$, is the ratio of the number of unique followers who follow both accounts to the full number of unique followers in the pooled communities, as follows:

$$\text{Similarity (Brand } B_x\text{, Exemplar } E_{ki}) = \frac{\#\text{ of unique users that follow both } B_x \text{ and } E_{ki}}{\#\text{ of unique users that follow either } B_x \text{ or } E_{ki}}$$

Because there are many exemplars ($E_{ki} \ldots E_{kj}$) associated with an attribute A_k, the next step is to roll up these individual similarities between the brand and each exemplar into a single quantification for the attribute. A simple and direct approach would be to calculate the mean of the similarities between B_x and each exemplar. However, such an approach treats all exemplars as equally informative, while both theory and prior research suggest that accounts with fewer followers are likely to be more informative for topic associations (see Culotta and Cutler[21] and Manning *et al.*[22]). For example, @AlGore appears on many user-generated 'environment' lists, and is clearly an account that is relevant to environmental issues — but his count of nearly 3 million followers reflects that his appeal is likely broader than just environmental. As many of his followers may follow him for reasons other than the environment, the environmental signal of his followership is diluted. In contrast, @DarrenGoode, an environmental reporter, also appears on multiple user-generate 'environment' lists — and his count of only 8,000 followers signals that his draw may be more targeted. Thus, it is recommended to use a weighted average of Jaccard similarities, where the similarities are weighted by the inverse of the exemplar's follower count. It is further recommended to take the square root of the final sum, to reduce the skew of the resulting distribution of similarity measures.

Specifically:

$$\mathrm{SPS}(B_x, A_k) = \sqrt{\frac{\sum_{n=i}^{j} \frac{1}{|E_{kn}|} * \mathrm{J}(B_x, E_{kn})}{\sum_{n=i}^{j} \frac{1}{|E_{kn}|}}}$$

Where $|E_{kn}|$ is the number of followers of exemplar E_{kn}, and $\mathrm{SPS}(B_x, A_k)$ is the 'social perception score' of B_x for attribute A_k, ie the estimated strength of B_x's brand image along the specific attribute of A_k. Each brand's social perception score will range from zero to one, with greater numbers indicating stronger brand image for that attribute. While the raw score returned for a single brand is, on its own, generally difficult to interpret, the relative scores of different brands provide meaningful information about relative associations, and can be used to populate perceptual maps. Although the above equation is presented as an example of a similarity calculation that has worked well, it is not a prescription that must be followed in all cases; the primary goal is to use a consistent method of quantifying the similarity between the follower base of a brand, and that of a group of exemplar accounts — and this general method appears to be robust to a range of different similarity metrics.

Extending to other platforms

While tested primarily on Twitter, the approach could be applied in any context in which brand network information can be obtained. For example, Facebook also has an open API that can be used to extract information about user activity on brand fan pages. While the API does not currently provide access to the list of users that 'like' a fan page (which would be the most direct corollary to the 'follow' relationship on Twitter), networks can be constructed in other, accessible ways — for example, by using lists of users who have liked or commented on marketing-generated posts that appear on the brand's fan page (such lists are easily retrievable for all fan pages through the API). In cases where brands do not maintain centralised social accounts, networks could potentially be constructed based on the authoring of user-generated posts that mention the brand, or relevant forum participation.

EXAMPLES

Implementing the above process for a set of brands and a keyword representing a dimension of brand image will result in a series of scores for the brands, indicating estimates of the relative strength of each brand's image in that dimension. These scores can then be used to generate perceptual maps or other market structure visualisations. Below, this method is applied to two examples (eco-friendliness perceptions of personal care brands and nutrition perceptions of food brands) and the resulting estimates compared against consumer surveys, which are a more traditional method of estimating brand perceptions.

Eco-friendliness perceptions of personal care brands

For the first example, the perceived eco-friendliness of 20 personal care brands is measured (using the keyword 'environment' to automatically identify exemplar accounts), and compared against the average responses of a questionnaire administered to 500 people (via Amazon Mechanical Turk) asking them to directly rate the eco-friendliness of each brand on a scale of 1 to 5. The resulting scatter plot is shown in Figure 2.

Overall, the correlation between the Twitter-based estimates and the survey-based estimates is 0.79 (Pearson method, $p<0.0001$), indicating a strong — though not perfect — similarity between the two measures. Survey-based methods are generally considered to be the gold standard for measuring consumer perceptions, although even these are not perfect: first, there is often wide variance in consumer beliefs, and summary statistics such as mean and median values may not always represent general perceptions well; and secondly, consumers do (or can) not always accurately indicate their beliefs through the questionnaire (they may not be aware of their own beliefs, or may falter in translating them to a linear scale). Thus, it is not entirely clear that the Twitter method is fully in error when it does not align perfectly with the survey scores; it is likely that both the survey and the Twitter method are approximations. Nonetheless, the strong similarity between the measures is encouraging that there is meaningful signal in the Twitter-based estimates. Looking at the scatter plot in Figure 2, one can observe that both the Twitter and survey methods reveal similar market structures: Burt's Bees stands out as having a very strong green image, followed by Aveda (which is perceived as green, but not as strongly), followed by the rest of the brands in a cluster, which all receive near-average green ratings. Within this cluster of relatively

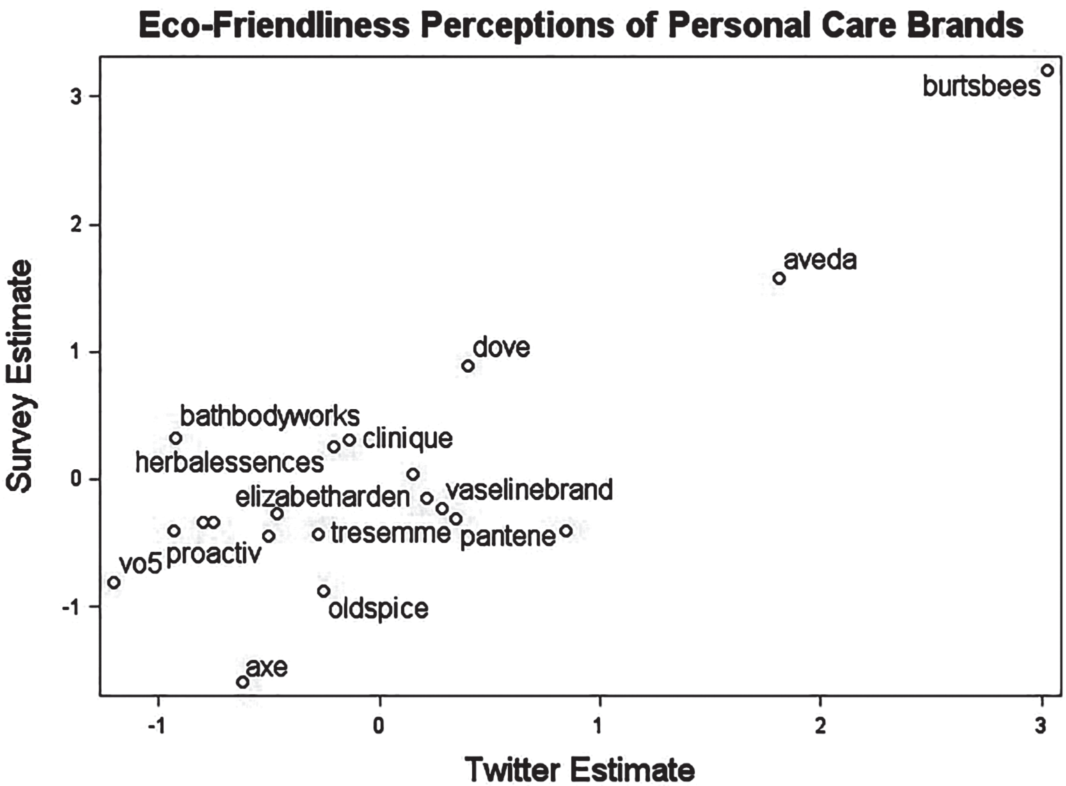

Figure 2: Comparing survey vs Twitter network based estimates of the perceived eco-friendliness of twenty personal care brands. Scores have been standardised for easier comparison (mean = 0, standard deviation = 1). Each point represents the scores of a single brand, labelled with its Twitter handle. Some labels are omitted for readability.

neutral perceptions, the survey and Twitter methods do not always align on the precise rankings of the brands; this may be due to error in one or both measures, or it may simply be that the differences in green image among the brands in this cluster are not clear in the population. Either way, the overall market structure and the key outliers are clear via both methods.

Nutrition perceptions of food brands

As a second example, the nutrition perceptions of 43 food brands is measured and compared against survey results obtained in a manner similar to that described above, using the keyword 'nutrition' to automatically identify exemplar accounts. This scatter plot is shown in Figure 3.

Here, the correlation between Twitter and survey based estimates is also 0.79 (Pearson method, $p<0.0001$). Looking broadly, one can see that both measures predictably show health-oriented brands such as Organic Valley, Nature's Path and Green Giant at the top, and candy/snack food brands such as Oreo, Doritos and Snickers at the bottom. One can also make out some sub-structures. For example, looking specifically at cereal brands, one can see through both measures that while none of the brands appear on the extreme ends of the spectrum, brands such as Special K, Cheerios and Kellogg's are substantially higher up the nutrition scale than, for example, Cap'n Crunch, which is rated more closely to candy and junk food. As in the prior example, disagreements between the Twitter and survey measures appear more pronounced in the middle of the scale. Again, some of this seeming error

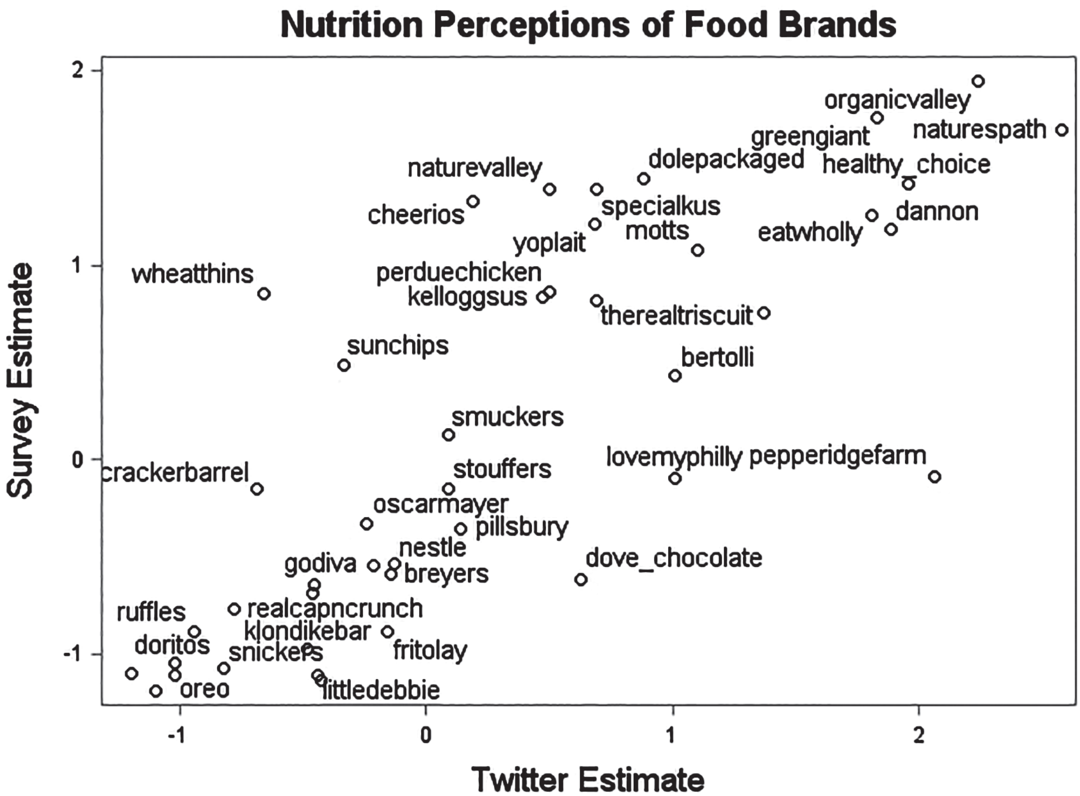

Figure 3: Comparing survey vs Twitter network based estimates of the perceived nutritional value of 43 brands. Scores have been standardised for easier comparison (mean = 0, standard deviation = 1). Each point represents the scores of a single brand, labelled with its Twitter handle. Some labels are omitted for readability.

may be a by-product of more inconsistent perceptions/ratings for brands that are not strongly differentiated by the perceptual attribute at hand, rather than a systematic difference between measurement techniques. However, there are also individual cases where the Twitter method appears to substantially over or underestimate nutrition perceptions, such as for Wheat Thins (where it is underestimated) and Pepperidge Farms (where it is overestimated). Such instances may reflect strategic marketing campaigns designed to change/amend prevailing image (the effects of which may be more apparent through Twitter followership than surveys of the general population); may reflect idiosyncrasies in follower motivations (ie users may follow one account for substantially different reasons than other accounts in the sector, such as for news updates); or they could simply reflect noise. It may be possible to improve accuracy by adjusting the search keyword (here, the single word 'nutrition' was used), by manually reviewing and filtering the returned exemplar accounts, or otherwise adjusting the implementation. While the overall correlation between Twitter-based and survey-based estimates is high in the default implementation, the disconnected examples serve as a reminder that this method is nascent — a tool to add to the marketing analyst's arsenal, but not to be applied blindly.

CONCLUSION

A large amount of information is revealed by who follows a brand on social media — and whoever else they in turn follow. Harnessing this information provides exciting new opportunities for monitoring brand image. Compared with traditional methods for measuring brand image, such as administering consumer surveys, this approach allows for unprecedented scale; full automation over freely available data allows for a large number of brand image dimensions to be tracked over time for a large number of brands. Compared with more common social media-based approaches, such as analysing the text of user-generated posts, this approach draws inferences from a wider range of brand fans (the majority of whom generally do not contribute content) and can measure associations unlikely to be discussed directly on public platforms.

A specific, tested implementation was described in detail to help marketers hit the ground running when trying out this new approach. But this is just a starting point. The main takeaway is the general addition of social media community-based measures into the marketer's broader analytics toolkit. Community membership and community similarity can be measured a number of different ways, and marketers are encouraged to explore and test implementation decisions that work best in different contexts.

ACKNOWLEDGMENTS

This research was funded in part by support from the National Science Foundation under grant #IIS-1618244. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsor.

References

1. Zeithaml, V. A. (1988) 'Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence', *Journal of Marketing*, Vol. 52, No. 3, pp. 2–22.
2. Aaker, J.L. (1997) 'Dimensions of brand personality', *Journal of Marketing Research*, Vol. 34, No. 3, pp. 347–356.
3. Jahdi, K.S. and Acikdilli, G. (2009) 'Marketing communications and corporate social responsibility (CSR): marriage of convenience or shotgun wedding?' *Journal of Business Ethics*, Vol. 88, No. 1, pp. 103–113.
4. Sen, S. and Bhattacharya, C.B. (2001) 'Does doing good always lead to doing better? Consumer reactions to corporate social responsibility', *Journal of Marketing Research*, Vol. 38, No. 2, pp. 225–243.
5. Nielsen Company (2015) 'The sustainability imperative: new insights on consumer expectations', available at: http://www.nielsen.com/us/en/insights/reports/2015/the-sustainability-imperative.html (accessed 28th October, 2016).
6. Aaker D. (1996) 'Measuring brand equity across products and markets', *California Management Review*, Vol. 38, No. 3, pp. 102–120.
7. Day, G.S. (1975) 'The threats to marketing research', *Journal of Marketing Research*, Vol. 12, No. 4, pp. 462–467.
8. McDaniel, S.W., Verille, P. and Madden, C.S. (1985) 'The threats to marketing research: an empirical reappraisal', *Journal of Marketing Research*, Vol. 22, No. 1, pp. 74–80.
9. Steenkamp, J.B. and van Trijp, H. (1997) 'Attribute elicitation in marketing research: a comparison of three procedures', *Marketing Letters*, Vol. 8, No. 2, pp.153–165.
10. Culotta, A. and Cutler, J. (2016) 'Mining brand perceptions from twitter social networks', *Marketing Science*, Vol. 35, No. 3, pp. 343–362.
11. Fader, P.S. and Winer, R.S. (2012) 'Introduction to the special issue on the emergence and impact of user-generated content', *Marketing Science*, Vol. 31, No. 3, pp. 369–371.
12. Toubia, O. and Stephen, A.T. (2013) 'Intrinsic vs image-related utility in social media: why do people contribute content to Twitter?' *Marketing Science*, Vol. 32, No. 3, pp. 368–392.
13. Naylor, R.W., Lamberton, C.P. and West, P.M. (2012) 'Beyond the "like" button: the impact of mere virtual presence on brand evaluations and purchase intentions in social media settings', *Journal of Marketing*, Vol. 76, No. 6, pp. 105–120.
14. Kuksov, D., Shachar, R. and Wang, K. (2013) 'Advertising and consumers' communications', *Marketing Science*, Vol. 32, No. 2, pp. 294–309.
15. Pereira, H.G., de Fátima Salgueiro, M. and Mateus, I. (2014) 'Say yes to Facebook and get your customers involved! Relationships in a world of social networks', *Business Horizons*, Vol. 57, No. 6, pp. 695–702.
16. Culotta and Cutler, ref. 10 above.
17. Twitter, Inc. (2016) 'Twitter developer documentation: the search API', available at: https://dev.twitter.com/rest/public/search (accessed 7th December, 2016).
18. Culotta and Cutler, ref. 10 above.
19. Stone, B. (2009) 'There's a list for that', available at: https://blog.twitter.com/2009/theres-a-list-for-that (accessed 28th October, 2016).
20. Culotta and Cutler, ref. 10 above.
21. Culotta and Cutler, ref. 10 above.
22. Manning, C.D., Raghavan, P. and Schütze, H. (2008) 'Introduction to Information Retrieval, Vol. 1', Cambridge University Press, New York, NY.