Social Measurement Depends on Data Quantity *and* Quality

As social platforms proliferate, enthusiastic users are generating more data than ever. Social media data are fast becoming the hottest commodity in market research.

Anne Czernek
Senior Research Analyst
Emerging Media Lab
Millward Brown/Dynamic Logic

But can social data yield measurements that are comparable to those from other, more established forms of research? Is it really possible for brand managers to tap into these data streams to gain insight into brand equity?

We believe it is too early to say for sure, even though social data have been used effectively by PR and marketing departments for years. For crisis management and on-the-fly campaign assessments, social monitoring involves watching a wide stream of updates in real time and using those to gauge immediate next steps. For these purposes, a qualitative sense of the consumer mood is adequate; the precision of quantitative research is not required.

But increasingly, insight and brand strategy teams are interested in using social data, and they would like to place social measurement alongside other types of brand metrics (attitudinal, behavioral, and so on). In this context, social data must be treated with the same rigor we expect of more traditional forms of measurement. Therefore, Millward Brown's Emerging Media Lab has conducted tests across 60

The future of actionable social media measurement is only as strong as its standards for data quality.

brands and more than 30 million online conversations to determine the most appropriate methodologies for working with social measurement from a brand perspective. Our conclusion? The future of actionable social media measurement is only as strong as its standards for data quality.

THE "SOCIAL" VOICE: HOW IS IT DIFFERENT?

Social data are fundamentally different from traditional brand measurement data. We can think of consumers speaking in two different voices. As Figure 1 shows, the "survey" voice of consumers is captured under structured and replicable conditions, while their "social" voice is observed in a fluid state.

FIGURE 1: SOCIAL VOICE VS. SURVEY VOICE

MillwardBrown
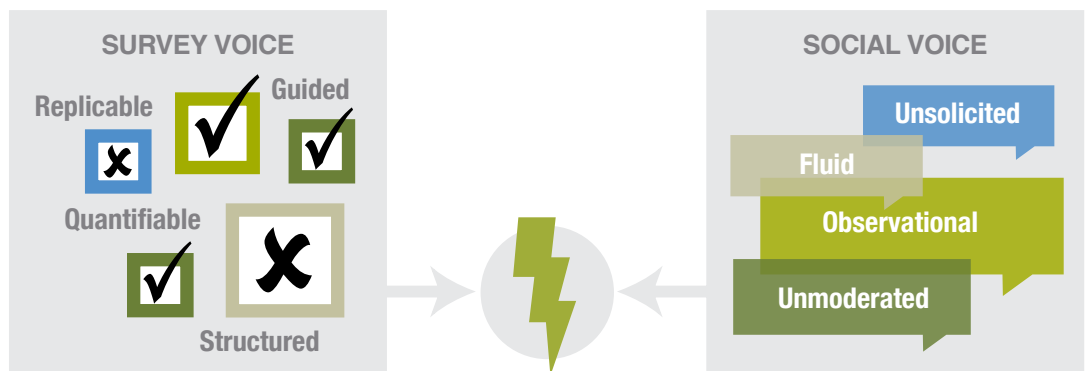
The challenge is first making these two types of data structurally comparable and then establishing linkages across them. Methodological issues arise out of the nature of the samples available to us in these two data sources. Traditional brand tracking and brand equity measurement rely on observing a statistically similar group of people over time. Individuals in a quantitative dataset may be treated differently—their opinions weighted more or less heavily—to ensure that the sample is representative and consistent.

> The only way to address the issues of user duplication and lack of profiling information is to put boundaries on our dataset. We need to make some choices.

But the social sample is a rolling sample of active, not necessarily representative, voices. Activists may overamplify topics they care deeply about, while people having positive but ordinary experiences with a brand may not feel compelled to speak up at all. It is not possible to weight these responses effectively because the information needed to profile respondents is not consistently available. Thus it is difficult, if not impossible, to ascertain whether those who post, tweet, or comment on a brand are representative of the population of interest.

We expect that social profiling will improve over time as we're able to derive more of these attributes from implicit relationships. But even with improved profiling, we anticipate ongoing concerns as basic as user duplication—e.g., how can we ascertain whether one individual has posted five times—once each on Facebook, Twitter, Tumblr, Blogger, and a forum—or whether five different individuals posted similar content?

DEFINING A SOCIAL UNIVERSE TO MEASURE

In the current environment, the only way to address the issues of user duplication and lack of profiling information is to put boundaries on our dataset. We need to make some choices.

Our collective eagerness to be perpetually connected has spawned an ever-expanding ecosystem of technology platforms, and as a result, the prevailing methodology for social media listening is to capture as much social data as possible—from Facebook, Twitter, Instagram, Pinterest, Tumblr, etc. But should comments on news articles be considered? What about reviews on sites like Yelp? How do we define the borders of the social universe?

It seems that the edges of the social universe are as murky as those of the real universe; in both cases, the boundaries are expanding at an apparently increasing rate. Without an agreed-upon definition of the limits of the universe—or even what "social" means—it's difficult to know whether we are truly capturing all of the relevant data.

Of course, there are also significant differences in the type of data available across platforms. Blog and forum discussion tends to be more composed and "conversationally" oriented than Twitter and Facebook updates, which tend to be short and posted on the fly. And while, to consumers, it may seem like social data points are easily accessible and browsable, from a research perspective, platforms control how those data streams are syndicated at scale. Twitter has monetized its "firehose" and charges for full access to it; Yelp, in contrast, prohibits collection of its reviews for research purposes. Table 1 shows the data fields that are available (or generally inferable) across different types of platforms.

TABLE 1: TYPES OF INFORMATION AVAILABLE ACROSS SOCIAL MEDIA PLATFORMS

	TWITTER (TWEETS)	FACEBOOK PERSONAL PAGES*	BLOGS/ FORUMS (POSTS)
Full data stream	✓	–	
Full text availability	✓	–	✓
Majority public profiles	✓	–	✓
Historical data	✓	–	✓
User-level data	✓	–	✓
User profiles	✓	–	
Influence	✓	–	
Location	✓	–	
Gender	✓	–	

*Facebook is a special case: Because data is only released in aggregate, none of the measures above are publicly available for research at a user level.

TWITTER, THE BEST SOCIAL RESEARCH SOURCE — FOR NOW

Because of all the issues we have mentioned thus far, we believe it is necessary to rely on a single source for social data, and to us, Twitter seems to be the strongest candidate. It's open, it's mobile, and it's the world's largest information platform. And because the Twitter firehose flows with a good deal of tweet-level information (including the full text, user ID, and timestamp), we're able to ascertain more not only about the text, but also the users. Thus Twitter meets our requirement for respondent-level information in a bounded dataset.

Moreover, because of the way it is used and perceived by users, Twitter seems to us to be most representative of the broader social sphere. Dynamic Logic's 2010 AdReaction study found that consumers viewed Facebook as being about connecting to friends and family, whereas Twitter was seen as an information platform for discovering, sharing, and learning. Though much has changed in social media in the intervening years, those observations still hold true. Facebook is still the central connection platform, even though consumers are also using smaller, more interest-based social networks to share content related to cooking, photography, sports, or other specialized interests. When the wealth of data from these sites is shared into the broader social stream, it usually comes through Twitter. Thus Twitter seems to us the best aggregation of discovery and sharing.

DATA ISSUES

But having chosen a data source, our work is not done. Further processing of social data is needed. Even if Twitter is the best source of data for social measurement, not all of the data within it is clean and useable—far from it. Our tests of over 30 million tweets show that up to 60 percent of Twitter data must be removed from the dataset before it is ready for analysis.

Twitter tends to have two main issues that can degrade data quality:

1. KEYWORD AMBIGUITY

Social data are generally collected through keyword searches, so when a brand name is also a common word, a large proportion of content returned will not be about the brand. For example, collecting data for the sandwich chain "Subway" returns many mentions of the New York City transit system. This not only distorts the themes of conversation topics tracked, but can also wreak havoc on other metrics, like sentiment. (In this case, a delayed subway train might generate many complaints that are irrelevant to $5 foot-long sandwiches.)

2. SPAM

There is widespread proliferation of spam content on Twitter. As Figure 2 shows, half of the social media mentions of a particular CPG brand were spam. When spam comments are removed,

FIGURE 2: SPAM PROPORTION OF TOTAL BRAND KEYWORD MENTIONS
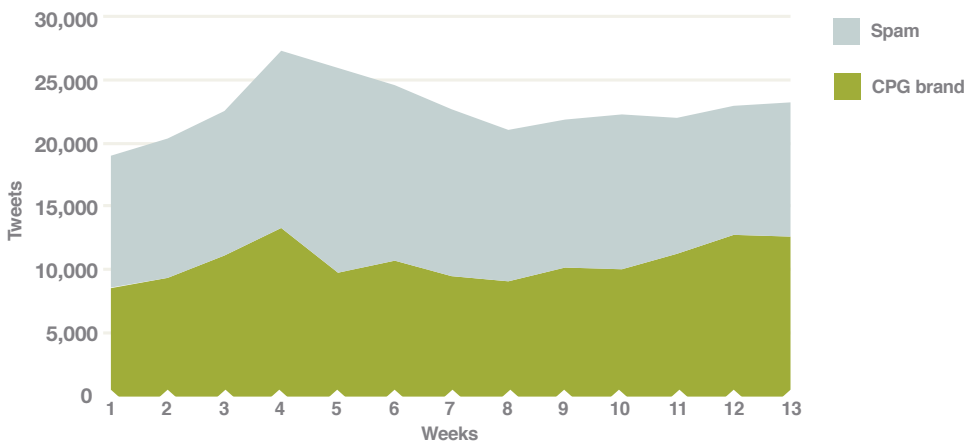
"net sentiment" tends to go down. (We obtain our measure of net sentiment by adding the percentages of positive and neutral content, and from that sum subtracting the percentage of negative comments.) This happens because most spam is neutral in tone; e.g., "Check out this coupon"—so removing it leaves negative sentiment proportionally higher, as illustrated by Figure 3.

FIGURE 3: EFFECT OF CLEANING ON SENTIMENT

Uncleaned   Cleaned

Net Sentiment* for CPG Brand
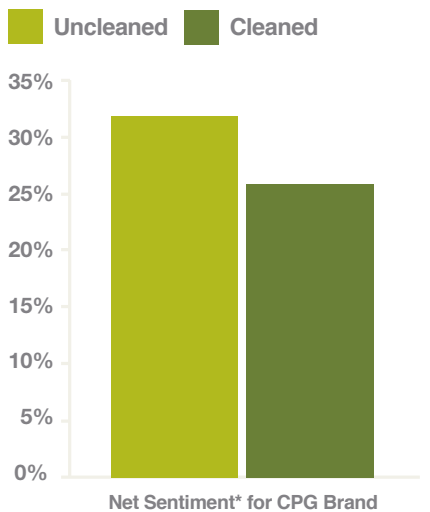
*Net Sentiment: (Positive + Neutral) - Negative

THE FUTURE OF SOCIAL MEASUREMENT

Using social measurement effectively will require unstinting attention to the quality of data we consider. The current generation of technology can help aggregate the dataset (through techniques like Natural Language Processing and applying Bayesian rules to cleaning), but human discretion is still needed to evaluate its source, quality, and worth.

While we believe social data have value for measuring brand performance, further work is needed to understand the exact relationship—if any—to brand equity. Looking forward, Millward Brown and Dynamic Logic are examining how our proprietary measure of brand performance in social media—which we call "social vitality"—relates to brand equity. We're assessing its value as a leading indicator of brand performance, as well as its usefulness in understanding the influence of events and media on brand perceptions.

This research is under way now, and we anticipate sharing results in the coming months.

Using social measurement effectively will require unstinting attention to the quality of data we consider.

The first step in harnessing the brand insights contained in social data is to create a dataset that can be managed by the same principles that govern established, trusted methodologies. We need to be assured that we are working with respondent-level data from a platform that has both breadth and depth, and we need to cleanse the dataset of spam as well as irrelevant references. Only when these steps have been accomplished can brands be confident that they are making defensible decisions based on reliable data.

The Emerging Media Lab is Dynamic Logic's specialty practice dedicated to research innovation across new media platforms — namely mobile, gaming and social media.

To read more about social media, please visit www.mb-blog.com.

If you enjoyed "Social Measurement Depends on Data Quantity and Quality," you might also be interested in:

"Facebook: Not an Ad Platform but an Ecosystem"

"Are You Getting Your Fair Digital Share?"

"Social Media: Fans and Followers are an End, not a Means"

"How Social Technologies Drive Business Success"