Collective Smile: Measuring Societal Happiness from Geolocated Images

Saeed Abdullah
Information Science
Cornell University
sma249@cornell.edu

Elizabeth L. Murnane
Information Science
Cornell University
elm236@cornell.edu

Jean M. R. Costa
Information Science
Cornell University
jmd487@cornell.edu

Tanzeem Choudhury
Information Science
Cornell University
tanzeem.choudhury@cornell.edu

ABSTRACT
The increasing adoption of social media provides unprecedented opportunities to gain insight into human nature at vastly broader scales. Regarding the study of population-wide sentiment, prior research commonly focuses on text-based analyses and ignores a treasure trove of sentiment-laden content: *images*. In this paper, we make methodological and computational contributions by introducing the *Smile Index* as a formalized measure of societal happiness. Detecting smiles in 9 million geo-located tweets over 16 months, we validate our Smile Index against both text-based techniques and self-reported happiness. We further make observational contributions by applying our metric to explore temporal trends in sentiment, relate public mood to societal events, and predict economic indicators. Reflecting upon the innate, language-independent aspects of facial expressions, we recommend future improvements and applications to enable robust, global-level analyses. We conclude with implications for researchers studying and facilitating the expression of collective emotion through socio-technical systems.

Author Keywords
Sentiment Analysis; Image Processing; Twitter

ACM Classification Keywords
H.5.m Information Interfaces and Presentation: Misc.

INTRODUCTION
The science of observing and interpreting human behaviors, thoughts, and feelings has been fundamentally changed by computer-mediated communication technologies and specifically social media platforms. Of particular interest to a wide array of researchers from psychology, sociology, HCI, economics, healthcare, and politics is the socially embedded affective and cognitive experience of how people perceive the quality of their lives. In the mid 1980s, Diener formalized this notion comprising positive affect, negative affect, and life satisfaction as *subjective well-being* (SBL) [15]. The corresponding Satisfaction with Life (SWL) score has since been developed to quantify happiness and compare societies of people in terms of fulfillment with life on the whole [16].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CSCW 2015, March 14 - 18 2015, Vancouver, BC, Canada
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-2922-4/15/03...$15.00
http://dx.doi.org/10.1145/2675133.2675186

In recent years, continued efforts have been made by scholars and policymakers to measure and promote subjective perceptions of well-being for individuals and groups at community, regional, national, and global levels. The term Gross National Happiness (GNH) originated as a holistic measure of well-being [68, 72], and GNH principles have inspired initiatives to measure national happiness from governments and organizations across multiple continents [80, 85, 74, 73, 83]. In 2012, the United Nations General Assembly resolved to design well-being indicators, launching the annual World Happiness Report [86] to rank countries by happiness. Similar agendas to calculate happiness for countries around the globe continue to grow each year [84, 79]. Measurement typically relies on polls, surveys, and other forms of self-report from individuals and households. However, this approach is costly in terms of time, money, and manpower and therefore is typically administered with limited samples of the population and only annually or less frequently [26, 54, 53].

Opportunely, a growing body of research demonstrates that an unobtrusive, inexpensive, and scalable solution to such challenges is implicit sentiment analysis using public social media data such as blog posts [9, 49], Facebook status updates [41, 53], and tweets [2, 18, 54]. This computational form of well-being monitoring aligns well with the aforementioned indices for both individuals and entire communities. However, approaches focus on text analysis, which has its own limitations, especially when input has complex semantics (e.g., sarcasm or humor) or syntax (e.g., unrecognized characters, non-English words, or typographical errors).

At the same time, the use of internet- and camera-capable smart devices is becoming increasingly prevalent. 1.8 billion new smartphones were sold in 2013 alone [71], and there are an estimated 7 million mobile subscriptions worldwide as of May 2014 [76]. Furthermore, numerous tools and sites have appeared that revolve around photo sharing and image curation, activities in which 62% of internet users take part [81]. This has produced a shift in online user-generated content from predominantly text-based data to richer forms of image-based media. 300 million images are uploaded daily on Facebook alone (7 petabytes worth of data per month) [69]; and between Facebook, Instagram, Flickr, Snapchat, and WhatsApp, 1.8 billion images are uploaded every day [82].

Recognizing this opportunity and necessity to design next-generation methodologies, we explore images on social media as a resource for sensing societal characteristics, focusing in this study on collective happiness. We apply our approach to 9 million geo-located images on Twitter and analyze aggregated emotional trends across multiple timescales. Interpreting uncovered mood patterns in light of culturally and politically significant events that took place during the captured window, we relate our findings to research from quantitative psychology and discuss important considerations such as the variety of cultural contexts in which users may be embedded and the implications of predicting happiness on a mass scale.

Our contributions are thus *methodological*, *computational*, and *observational*:

- We provide a novel technique for large-scale sentiment assessment based on smiles detected in images shared on social media.
- We present our analyses, which demonstrate our approach's ability to: i) track how societal happiness changes over time and in response to emotionally significant events; ii) characterize communities from a socio-economic perspective; and iii) predict consumer confidence.
- Lastly, we discuss theoretical and practical implications of our work, including the potential that images hold as a far-reaching, fast, and low cost alternative to existing methods for capturing collective characteristics and social behavior.

RELATED WORK

Facial Expression, Emotion, and Photography

The relationship between facial expressions and emotions has been extensively studied, leading to central theories from various fields that together suggest these brief non-verbal cues reveal a great deal about individuals' emotional states [60, 42]. Facial expressions can be annotated according to a number of core emotions [19], including 6 initially identified by Darwin: anger, disgust, fear, happiness, sadness, and surprise. Ekman's Facial Action Coding System (FACS) has been used to characterize facial expressions such as smiles according to more fine-grained features [20]. Multiple studies have used FACS to code smile intensity [28, 57].

When it comes to *conveying* emotions and opinions, visual imagery proves a highly effective medium. Recent studies show that expression of positive affect in photographs can be indicative of an individual's long term well-being and happiness as well as provide insight into underlying personality traits that shape social, cognitive, and behavioral tendencies and responses [36]. Notably, two longitudinal studies by Seder et al. [57] show that smile intensity in Facebook profile pictures is positively associated with social relationships and is a robust predictor of self-reported life satisfaction after 3.5 years for both male and female participants.

That being said, a complete and overarching theoretical framework explaining the causal relationship between happiness and smiling in photographs has yet to be developed. Specifically, it is important to consider individual differences in the expression of emotion in photographs as well as the authenticity of emotion displayed. Prior research on posing and smiling in photos suggests that high-status and dominant individuals tend to smile less in photos than submissive types and that men smile less than women [27]. Smiling in spontaneous, candid photos is regarded as an indicator of authentic happiness [40], and though some photographed smiles may not reflect true happiness or may even mask negative emotions, the proportion of such smiles is generally small compared to the ones displaying genuine emotion, for people from both Eastern and Western cultures [30]. Regarding this last point, the role of culture is another factor requiring further attention when measuring emotion from photos. We address the feasibility of our framework for cross-cultural sentiment analysis in the latter portion of our Discussion section.

Assessing Sentiment with Social Media

Today's unprecedented access to records of self-expression and social interaction archived on social media has empowered the CSCW community to explore data-driven social science questions. Significant research has leveraged this resource both to extend our understanding of human behaviors and contexts as well as to contemplate practical implications such insights can have on sociotechnical research happening across diverse fields.

Considerable recent attention has been given in particular to leveraging online social networks and content-sharing platforms to assess the affective states of individuals and groups, for example to study mood sharing and emotional expression [38, 63]; population-wide assessment and diagnosis of mood and mental health [11, 12]; and public emotional responses to political events such as debates [45] and armed conflicts [13].

Prior work has also shown social influence to be a major factor impacting subjective well-being [24], further motivating our use of data contributed within the social setting of a platform like Twitter. With respect to societal happiness in particular, a growing body of work related to large-scale sentiment analysis has recognized the opportunity social media affords to study trends of collective mood. This research typically applies natural language processing (NLP) techniques and uses tools such as the Linguistic Inquiry Word Count [52] to quantify sentiment. Kim et al. [37] use Affective Norms for English Words (ANEW) [4], a list of terms with emotional ratings, to track mood on Twitter. Bollen et al. [2] infer sentiment from tweets by extending the psychometric measurement Profile of Mood States (POMS) [47], which consists of mood adjectives. Dodds et al. [18] compile and apply an emotion-annotated corpus of words to measure the happiness expressed in tweets.

Computational Emotion Recognition

In parallel to the psychological and sociological studies on emotion, systems-focused researchers have also undertaken work to detect as well as foster emotional expression.

Much work from Computer Vision and Processing has pursued the development of hardware and algorithms for facial expression recognition, including for smile detection [64]. There has also been increased interest in the development of computational frameworks capable of automatically inferring about the emotion associated with digital images [35].

Considering the beneficial influence smiling can have on overall affective state, a number of systems developed by the HCI community focus on encouraging smiles. Tsujita et al.

[61] use household cameras to detect smiles and provide feedback to users in order to increase self-awareness about smiling behavior and mood. Hernandez et al. [29] station cameras on a college campus to detect smiles in real time and reflect upon the community's emotional reactions. Visualizing patterns in collective sentiment over time and in relation to campus events, their experiment finds that trends in smiling strongly reflect events such as exam periods and graduation.

Thus to summarize, by building on findings from psychology, leveraging large-scale emotionally-rich image data from social media, and applying facial recognition techniques from computer vision, we pursue a novel methodology for detecting smiles in online user-contributed content that allows for the assessment of collective emotion on a global scale and over fine-grained time intervals. In the sections that follow we describe our technique, the results of our validation experiments, and the insights we derive from these applications.

METHODOLOGY

Data

In this study, we use a dataset of nine million geo-tagged tweets with images that were posted between January 1, 2012 and April 30, 2013, a time window that gives us access to content posted during emotionally significant political and cultural events in the United States such as the presidential inauguration of Barack Obama and the Boston Marathon bombings. We obtain the tweets via Twitter's "garden hose" — the service's official sample of approximately 10% of all tweets. Our dataset is thus the subset of garden hose tweets containing images that were posted during our collection period.

We determine the geographic locations of tweets from the latitude and longitude coordinates provided in the *geo* metadata field. Figure 1 illustrates the distribution of the tweets' counties of origin across the United States normalized by population. To handle timezone differences across geographical regions, we use the same *geo* metadata to convert a tweet's Coordinated Universal Time (UTC) formatted creation-time to local timezone for fine-grained temporal analysis.

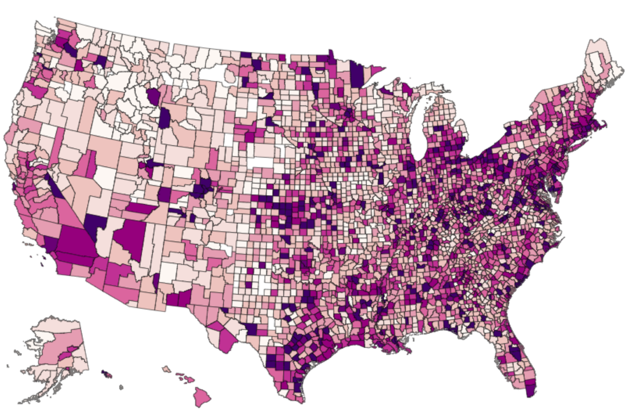

Figure 1. Distribution of tweets per-capita across counties in our dataset

It is worth noting here that Twitter (along with other popular social media sites including Facebook) remove metadata (e.g., Exif data) from images. As a result, it is impossible to differentiate when and where an image is captured vs. when and where it is posted, based on information available in Twitter's datastream. Researchers whose work depends on the assumption that posting happens at the same time and location as capture should thus be aware to either seek data from sources that do not strip this metadata or take steps such as survey follow-ups to ensure geography and recency of content. Photos posted by tourists vs. locals may be the main source of discrepancy [25, 77]. That being said, prior work has shown that in general, users do share information on social media in real-time and post geotagged tweets from their current locations (e.g., [67]). Particularly relevant is the CityBeat project [66], which uses geotagged social media photos to classify diverse events such as car accidents, fires, and music concerts. The high accuracy they achieve in detecting these events demonstrates the geographical and temporal reliability of such image data at the hyper-local level. Relatedly, the transient nature of the Twitter medium encourages users to focus on current as opposed to past events, and this in-situ nature of tweets has been confirmed by a number of projects with focuses ranging from real-time news detection to emergency support during events such as earthquakes [56]. Given the large-scale nature of our data, we therefore expect that any potential noise-to-signal ratio resulting from individual sharing habits would be small.

The images in our tweets were uploaded using the official photo-sharing service provided by Twitter. Around 12% of these images were inaccessible at the time of crawling, presumably due to user deletion. Of all the accessible images, around 40% contain a human face according to our smile detection framework, which we describe next.

Smile Index Framework

Figure 2 shows the phases of our happiness assessment. One of our key aims is building a system that is both fast and reliable given the large size of current and future data sources.

Extracting images

The first step is processing our Twitter dataset to extract images. To determine whether a tweet contains any image, we use the post's associated metadata. Specifically, we check the *media* field in the *Entity* metadata[1] and retain a tweet for further processing if the media element has a type *photo* associated with it. An exploratory look into a random sample from the API shows that more than 72% of all images uploaded in Twitter use the site's official photo-uploading service; all images in our dataset have been uploaded using this service. After identifying image-containing tweets, we parse the metadata to build a database of links for crawling the images directly from Twitter. Since the images from Twitter can have different encoding formats, we perform image conversion to ensure all the images are in a standard JPEG format. We also convert RGB images to grayscale since subsequent phases in our framework do not use color information.

[1] `https://dev.twitter.com/docs/platform-objects/entities`

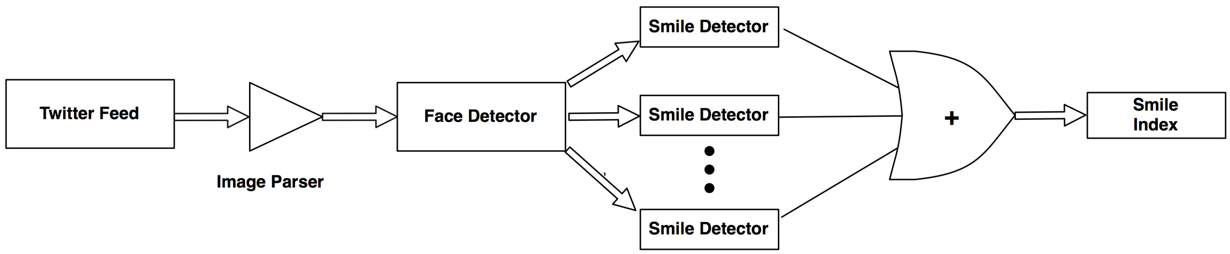

Figure 2. Framework for assessing happiness through Smile Index

Face and smile detection

To detect faces within images, we use a cascade of boosted classifiers with Haar-like features [62]. The classifier uses the change in contrast between adjacent rectangular regions instead of intensity of individual pixels to determine relative dark and light areas. These features can be easily scaled, which allows efficient detection of objects of various sizes. The final classifier is a cascade of several stages, where each individual stage consists of boosted classifiers. We use a trained cascade consisting of twenty stages from Lienhart et al. [43]. The processing in these stages allows for the efficient elimination of false positives. Ultimately, we achieve 100% accuracy in detecting faces in our test dataset. Such high accuracy is expected since we are leveraging well-established and well-tested facial detection techniques [62]. After this phase, we retain only images containing human faces.

Next, we detect smiles in these images by using a Haar feature based classifier trained by Hormada et al. [31], who manually labeled images from multiple sources including the photo sharing site Flickr. The classifier achieves an AUC (area under the ROC curve) score of 0.9, a strong prediction rate. If an image contains multiple human faces, we run the smile detection algorithm for each detected face and combine the results by performing a logical OR since individuals are likely to show similar sentiment in a photo. This means that if any face detected in an image is smiling, we consider that a positive instance of smiling.

To test our system, we use the public GENKI [64] database, which consists of images of people from publicly accessible personal web pages. The people in the images are diverse in gender, ethnicity, age, geographical location, facial features (e.g., eyeglasses, facial hair), and photo setting (e.g., indoor and outdoor). Our system achieves 0.85 accuracy and F-Score of 0.71 on GENKI-4K[2], a subset of 4000 images manually labeled for smile presence.

Quantifying happiness

Our final step is building our Smile Index of happiness. A possibility for computing this measure over a time period t for a given geographical area is using the total count of images in which smiles are detected:

$$S_t = \sum_{i \in I_t} f_{smile}(i)$$

I_t is the total number of images in time period t from that geographical area, and f_{smile} represents our smile-detection output (1 if the image i contains any smiling human face and 0 otherwise). However, Twitter's growth in popularity and contributed content over time means that the volume of tweets in our dataset increases significantly over the collection period. Simply using a raw smile count as seen in Figure 3 would therefore lead to a systematic bias where the Smile Index would merely reflect the increasing contribution levels in Twitter rather than the happiness levels of those contributors.

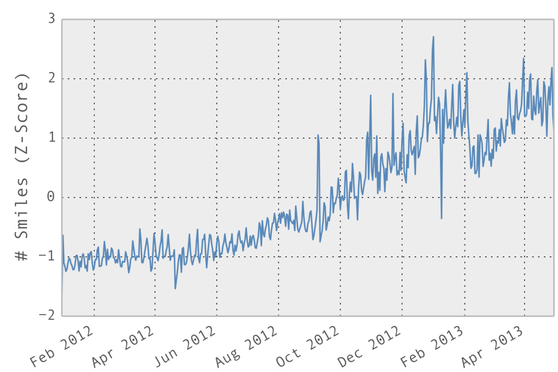

Figure 3. Smile count (normalized using z-scores) in tweets over time

For this reason we use another metric — the ratio of smile-containing images to the total number of images posted during a given time period:

$$R_t = \frac{S_t}{I_t}$$

Our Smile Index is thus a normalized smile frequency — that is, the number of images containing smiles normalized by the total number of images posted in the same time period. Figure 4 shows the resulting distribution of this ratio for U.S. tweets.

Not only does this ratio prevent the bias just described, but it also allows for the comparison of Smile Index measurements arising from datasets of different sizes. It should be noted that we do not use the ratio of smiling faces to the number of faces because images that do not contain faces might still convey important information about external events. This is particularly true because given the nature of social media, if an image contains a face it is more likely to be a smiling one, causing the ratio of smiling faces to total face count to be relatively constant over time. Instead, setting the total count

[2]MPLab GENKI-4K database `http://mplab.ucsd.edu`

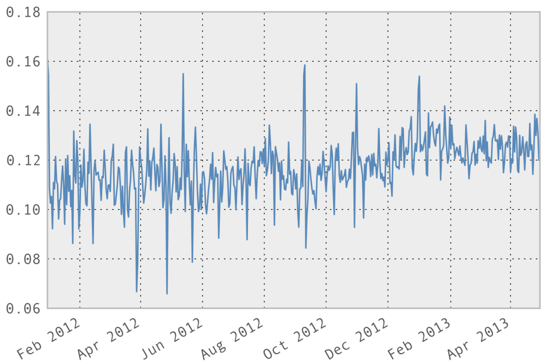

Figure 4. Ratio of images that contain smile(s)

of images as the denominator lets us quantify the appearance of smiles within a given time-period in a broader context. For example, during Hurricane Sandy, we see a discernible rise in the number of images that do not contain faces but that do capture the surroundings and communicate about the event in that way. This lower ratio of smiles to the total number of images allows us to correctly detect the negative sentiment associated with this event.

A key question is whether the image-based methodology we propose is reliable and enables the robust study of emotional and sociological phenomena. To address this, we now compare our approach with a state-of-the-art text-based analysis as well as with ground-truth self-reported happiness data.

Dodds et al. [18] publish a daily happiness score[3] computed over 100 million words per day from Twitter. Given that our and Dodds et al.'s indices use different scales, we normalize the time series to a $[0,1]$ scale using the following equation:

$$h_{\text{norm}} = \frac{h - h_{\text{min}}}{h_{\text{max}} - h_{\text{min}}}$$

To compare the happiness scores computed from smiles and from text, we perform statistical equivalence testing. Using two one-sided tests (TOST) with interval margin $\delta = 0.3$, the null hypothesis that the difference between the samples' means is larger than the specified margin threshold can be rejected with p-value < 0.001. In other words, we find statistical equivalence between happiness computed via traditional text-based and our image-based analyses. It should be noted though that our paper's focus is not to compare and contrast our alternative methodology with existing text-based techniques or to argue that our approach is superior. Rather, our intention is more primarily about mining, analyzing, and exploring a previously unplumbed datastream. In fact, it would be desirable to combine both textual and image based sentiment analyses together to play on the strengths of each. For instance, text-based analysis is computationally simpler; but in many cases text is malformed, not in a processable language, or simply non-existent, which motivates the use of an alternate, text-independent approach such as ours.

[3] `http://hedonometer.org/`

We also compare our Smile Index with the Gallup-Healthways Well-Being Index [70], which tracks the ratio of Americans experiencing "lots of happiness and enjoyment" and "lots of stress" using daily telephone surveys. While Gallup's index is limited in the sense that it tracks the extreme ends rather than the full happiness spectrum, we still find expected correlations with our Smile Index, as seen in Table 1. Day-to-day changes in happiness and stress are also strongly correlated with detected smiles: 0.5^{***} and -0.43^{***}, respectively. To further compare the Gallup happiness index and our Smile Index, we again perform statistical equivalence testing. Since our Smile Index is a ratio (of images with smiles to all images), we normalize the Gallup happiness score to the same $[0,1]$ scale as we did when making the comparison with text-analysis. The result shows that the happiness time series of Gallup and of our Smile Index are statistically equivalent with an interval margin of $\delta = 0.3$ and $p < 0.001$.

	Happiness/Enjoyment	Stress/Worry
Smile Index	0.26^{***}	-0.22^{***}

Table 1. Smile Index compared with Gallup Well-Being Index. (Throughout the paper, we use $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$ to mark level of statistical significance)

EXPERIMENT AND RESULTS

Temporal Smile Distribution

Here we examine temporal trends in the Smile Index across different timescales (hours, days, and weeks).

Daily patterns of happiness

As the term "Blue Monday" indicates, there is a commonly held cultural belief about the relationship between affective state and the day of the week. To see if we observe such an effect, we compute the distribution of smile-containing images across days of the week, aggregated over the United States. The result is shown in Figure 5.

There is a clear peak during the weekend, which makes sense intuitively and also conforms with prior research that observes the expression of increased positive sentiment and lower negative sentiment on weekends [44]. A sharp decay occurs after the weekend, but we see the minimum reached mid-week on Wednesday. A number of previous studies [59, 48, 18] have similarly reported the absence of a Blue Monday

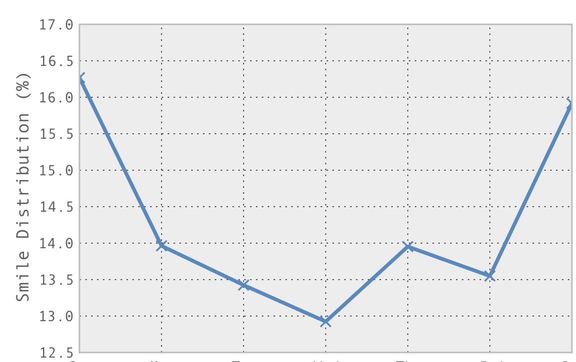

Figure 5. Distribution of smiles in U.S. tweets over days of the week

phenomenon in their findings, and the mid-week dip in happiness has also been observed in prior work based on blog-post data [48] and tweets [17]. We conjecture that relatively higher Monday happiness is a result of inertia from the preceding weekend and note the analogy of the Wednesday mid-week mood dip with the known mid-day dip [50].

Hourly patterns of happiness
Examining the hourly pattern in the appearance of smiles, we find a diurnal rhythm as shown in Figure 6. Our data shows a gradual increase in happiness during the morning, which then levels off and reveals the mid-afternoon dip. The patterns we observe are consistent across workdays and weekends as shown in Figure 7.

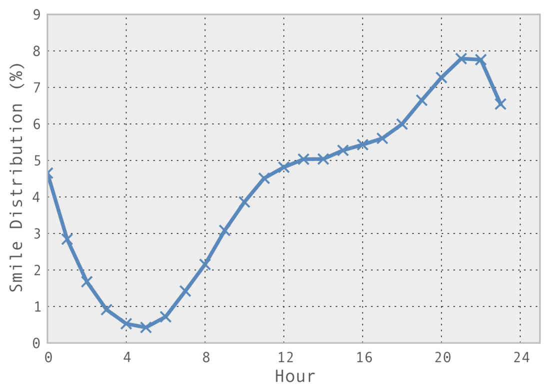

Figure 6. Distribution of smiles detected in U.S. tweets over hours

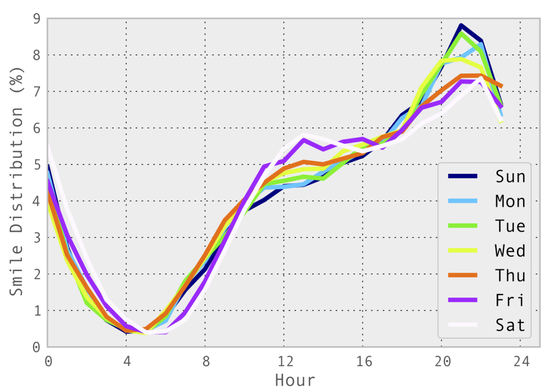

Figure 7. Daily breakdown of hourly smile distribution

Clark et al. [8] in a study on undergraduate students note a similar plateau effect in the middle of the day, albeit their reported dip lasts from noon to 9 pm. In the evening, we see smiles again increasing steadily with a peak around 9-10 pm, then dropping to a minimum in the middle of the night, which is when research shows depressed individuals tend to be more active on Twitter [11]. A recent study by Golder et al. [26] uses text-based content from Twitter to analyze patterns in positive and negative affect and finds trends similar to ours — a peak in Positive Affect (PA) in the morning and another at midnight as well as a minimum in PA around 5 am. However, the peaks from Golder et al. have similar PA values to each other, while we see much higher levels in the evening.

Public mood surrounding emotionally significant events
To investigate whether public mood is reflected by smiles in uploaded images, we look closely at a variety of societal events. While Twitter is a global service, the majority of its users are from the United States, so for reasons of scope in this study we consider only U.S. events: the Boston Marathon bombings, the inauguration of President Obama, the period during Hurricane Sandy and its aftermath, and New Year's Eve. To keep the dataset diverse, we focus on particular geographical regions of the country during our analyses. Figure 8 summarizes our findings from these case studies.

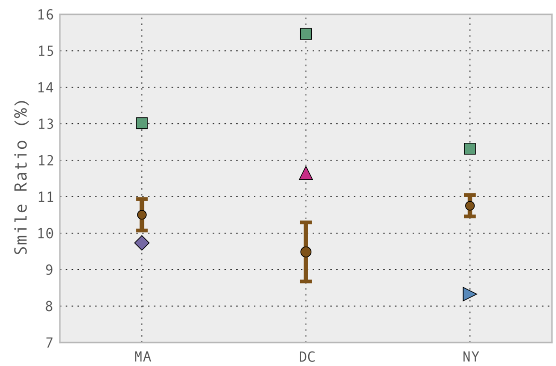

Figure 8. Smile Index (SI) of indicated events. ● shows the mean for each of the three geographical regions, and 99% confidence interval is included. ◆ indicates SI in Massachusetts during Boston Marathon Bombings and subsequent manhunt period; ▲ marks SI during Presidential Inauguration in DC; ► indicates SI during Hurricane Sandy in NY; ■ represents SI in the corresponding areas during New Year's Eve

First, we analyze images posted April 15-19, 2013 from Massachusetts during the tragic Boston Marathon bombings and subsequent manhunt. The mean of daily smiles detected during this period falls with a margin of more than 7% compared to the total time period in our dataset. In other words, even though the number of images increased during this time period, the Smile Index was significantly lower.

Second, we measure smiles detected in images from Washington D.C. during the second inauguration of Barack Obama as President of the United States. The event included concerts, a swearing-in ceremony, a parade, and other activities spread over January 19-21, 2013; so we use images posted during these dates. For this time frame, we see a 22% increase in the Smile Index compared to the overall mean. Figure 9 shows an example of a smile-containing image posted during this period and relevant to the event.

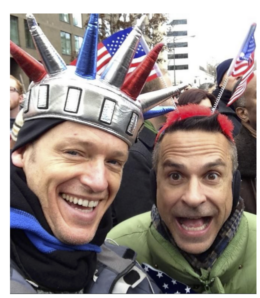

Figure 9. Example of a smile-containing image contributing positively to the Smile Index for the 2013 U.S. Presidential Inauguration

Third, we study images from NY posted during and after Hurricane Sandy, a Category-3 storm that caused severe devastation, including 147 deaths and a damage estimation exceeding $\$50$ billion [75]. Its storm-force winds at one point had a diameter of 1000 miles. Hurricane Sandy made landfall in NY on October 20, 2012, so we use images posted from this date until November 1, 2012. Comparing to the overall mean, the Smile Index during this period drops 22.5%.

Finally, we also analyze images posted from each of these geographical areas on New Year's Eve (on December 31, 2012 and Jan 1, 2013 to be specific). As shown in Figure 8, there is a substantial increase in the Smile Index during this holiday.

Hence for all four events, there is a noticeable increase or decrease in our Smile Index. The ratio of images containing smiling faces increases during positively affective celebrations and decreases during negatively affective calamities, and the difference in each case is significant with $p < 0.01$. In these case studies, our simple smile metric thus proves an effective barometer of public mood during societal events.

Happiness and Economy

Understanding the potential link between well-being and various economic phenomena is arguably one of the most mainstream topics in Economics. This is evidenced by Clark's [6] recent comparison of article titles between 1960 and 2000 in the electronic bibliography of The American Economic Association — the publishing rate of articles on well-being and happiness increased more than 8 times after 2000. Researchers from the discipline commonly assume that happiness is a representation of true internal utility with some noise. Rafael et al. [14] note that if this signal-to-noise ratio in the available data is large enough, a measurement of happiness can enable study of numerous topics ranging from determinants of political economy to the effects public policies have on social welfare. In the following sections, we draw on this expanding literature to look into the relation between economic phenomena and our Smile Index of happiness.

Median household income

The connection between income and happiness has been well-studied. Generally, the relationship is suggested to be positive but with diminishing returns [7], and a number of studies argue in favor of relative rather than absolute income as a determinant of well-being measurement [23].

Using the U.S Census Bureau's 2007-2011 American Community Survey 5-year estimate data[4], we compare the distribution of median household income with per-capita smile count across New York City zip areas. We use per-capita smile count since this analysis focuses on community characteristics rather than temporal trends. We see a high degree of variation in median household income — the maximum is $\$232,031$ in Manhattan and the minimum is $\$19,840$ in the Bronx.

Table 2 shows the correlation between per capita smile count and median household income. Income does positively correlate with the smile metric, though perhaps more interesting

	Correlation
Overall	0.173074^{*}
Black Majority Area ($> 50\%$)	-0.586325^{**}
Black Majority Area ($> 75\%$)	-0.726471^{**}
White Majority Area ($> 50\%$)	0.312572^{*}
White Majority Area ($> 75\%$)	0.536020^{**}

Table 2. Spearman correlation between median household income and smile count per-capita across NYC areas

is the clear pattern that emerges when we take racial demographics into consideration. Specifically, the correlation goes from positive to negative in black majority areas (i.e., higher income is associated with lower positive sentiment and vice versa), while the correlation is more positive in white majority areas. The correlations for other races are not statistically significant in our dataset. Note that we follow the 2010 Census in treating Hispanic as an ethnicity but not a race. To better isolate the relationship between racial distribution, income, and smile data, we use the following regression model:

$$L1: \ C_z = \alpha + \beta_i I_z + \beta_b B_z + \beta_w W_z + \beta_o O_z + \epsilon_z$$

where C_z is the number of smile-images per capita, I_z is the median income, B_z is the fraction of the population identifying as black in zipcode z, and W_z is the fraction identifying as white. O_z represents the other racial proportions in that region. Table 3 presents the result estimated by Ordinary Least Squares regression. As the summary statistics show, there is no serious auto-correlation problem with the model. The strong positive coefficient for W_z indicates that an increased white population proportion in a given zip code corresponds to an increase in the number of images with detected smiles.

	Coeff	Std. Err.	t
I_z	-0.001	0.00	-0.731
W_z	0.774	0.2	3.8^{***}
B_z	-0.224	0.4	-0.5
O_z	0.186	0.17	1.06
Summary	DW	Adj-R	F-Stat
	1.4	0.059	4.101^{**}

Table 3. Summary and coefficient statistics of $L1$. DW stands for Durbin-Watson statistic, a measure of residual auto-correlation

Given the relatively small coefficient associated with income in $L1$, we also look into the correlation between racial distribution and images with smiles across the NYC area. The results are shown in Table 4. For this analysis, we use the percentage of black populations and white populations in every zip code and compare against the number of per capita images with smiles. We find a strong positive correlation for the white populations, while the correlation for black populations is negative.

Race	Correlation
White	0.464114^{***}
Black	-0.371355^{***}

Table 4. Statistically significant correlation between race and number of images with smiles per-capita across zip codes in NYC

[4] `http://www.census.gov/acs/www/`

This intriguing finding is consistent with extant research on happiness and well-being. For instance, using surveys between 1957 and 1972, Campbell et al. [5] found 32% of whites to be "very happy" compared to only 18% of blacks. More recent studies [33, 10] report that although declining, there still remains a discernible black-white gap in terms of well-being and happiness.

Public Opinion Polls

Following the observed correlations between our Smile Index and economic factors, we investigate to what extent our metric may also capture trends in public opinion polls. In particular, we are interested in the consumer confidence index, which captures the collective degree of optimism that people express about the overall status of the economy and their personal financial situations. Such measurements are of substantial value to both governments and business entities since high consumer confidence indicates higher spending and often results in economic growth.

A methodology based on surveys and polling has been extensively utilized to understand such attitudes of population samples, and data collected with these methods are often considered a gold standard when it comes to public opinion. Thus, comparing against these indicators provides an additional validation of the reliability of our approach. A number of surveys attempt to measure consumer confidence in the U.S. For instance, Reuters/University of Michigan Surveys of Consumers maintains the monthly-published Index of Consumer Sentiment (ICS)[5]. However, since we have temporally fine-grained data from Twitter, we prefer a consumer dataset with more frequent observations and use the daily-administered economic confidence index from Gallup[6], which is calculated from a 3-day rolling average where each data point is based on interviews with 1,500 Americans aged 18 and older. The aggregated Gallup economic index is known to correlate well with ICS, and both indices do correlate strongly for the duration of our study ($r = 0.59$, $p = 0.01$).

We also find that the consumer confidence index from Gallup correlates strongly with the normalized frequency of smile-containing images ($r = 0.608$, $p < 0.0001$). However, we are more interested in whether variations in detected smiling is a reflection of consumer confidence. To formally test the relationship between these two time-series, we use the well-known econometric technique of Granger causality analysis.

Granger causality analysis

Granger causality tests for the ability of one series to predict another one — in our case, whether our Smile Index provides information to forecast the economic index. More formally, the analysis uses two linear models G_1 and G_2 to predict E_t as shown in equations (1) and (2). The first model, G_1, uses endogenous information — only the lagged value of E_t. G_2 incorporates another variable, H_t, combining the lagged value of both E_t and H_t. If the variance of the regression model G_2 decreases by including G_t, then H_t is said to Granger-cause E_t. The assumption here is that if H_t causes E_t, then the lagged value of H_t will strongly correlate with E_t. We note that this analysis is not asserting true causation but rather the presence of useful predictive information.

$$\mathbf{G_1}: \quad E_t = \alpha + \sum_{i=1}^{k} \beta_i E_{t-i} + \epsilon_t \tag{1}$$

$$\mathbf{G_2}: \quad E_t = \alpha + \sum_{i=1}^{k} \beta_i E_{t-i} + \sum_{i=1}^{k} \gamma_i H_{t-i} + \epsilon_t \tag{2}$$

In the equations, ϵ_t represents noise and k is the lag parameter, which is free in Granger causality analysis. In our case, we are interested in the trend of daily change, and so we use the difference in values across subsequent days:

$$E_t = EI_t - EI_{t-1}$$
$$H_t = S_t - S_{t-1},$$

EI_t and S_t are respectively the values of Gallup's economic index and the frequency of images with smiles, on day t. The Granger causality analysis assumes that the data is stationary, which the differencing measurement ensures.

Lag	F-statistic	p-value
1 day	4.74	0.03*
2 days	4.64	0.01*
3 days	2.5047	0.058

Table 5. Summary of bivariate Granger causality analysis between Gallup Economic Confidence Index and frequency of images with smiles

We estimate equations (1) and (2) using an Ordinary Least Squares regression model. Table 5 shows the result of the causality analysis. In this case, the model G_2 performs significantly better than G_1. The F-statistic here is proportional to the relative change in the residual sum of squares between models. In other words, we can reject the null hypothesis that our Smile Index does not predict consumer confidence with a high level of confidence.

Forecasting analysis

The result of our preceding Granger analysis indicates that our image based Smile Index captures predicative information about consumer economic confidence. To further validate, we train and evaluate our model in a rolling forecast setting. Similar to Brendan et al. [51], we use a one variable least-square model:

$$EI_t = \alpha + \beta \sum_{i=0}^{k-1} S_{t-i} + \epsilon_t \tag{3}$$

For a given hyperparameter k, the model uses a k-day window for training. In other words, to predict the forecast on day t, the slope and bias parameter of the model β and α are fitted using data from $t - k$ to $t - 1$ days. The fitted model is then used for prediction. For evaluating in a rolling forecasting setup, we introduce another hyperparameter L:

$$EI_{t+L} = \alpha + \beta \sum_{i=0}^{k-1} S_{t-i} + \epsilon_t \tag{4}$$

[5] `http://www.sca.isr.umich.edu/`

[6] `http://www.gallup.com/poll/122840/Gallup-Daily-Economic-Indexes.aspx`

The parameter L determines how many days in the future we are forecasting. In other words, for a target forecast date $t+L$, we train the model through $t-k$ to $t-1$ days and predict using the value from the last day of the window, day t.

To evaluate performance, we use Mean Absolute Error (MAE). We find that the best model for forecasting 7 days ahead uses 13 days of historical data. That is, we use the data from days $t-13$ to $t-1$ to learn the parameters of the model, and we then ask the model to use the value from day t to predict day $t+7$. We then measure the error in this predicted value compared to the actual eventual outcome. Figure 10 illustrates the alignment between this forecast and the actual Gallup economic confidence index.

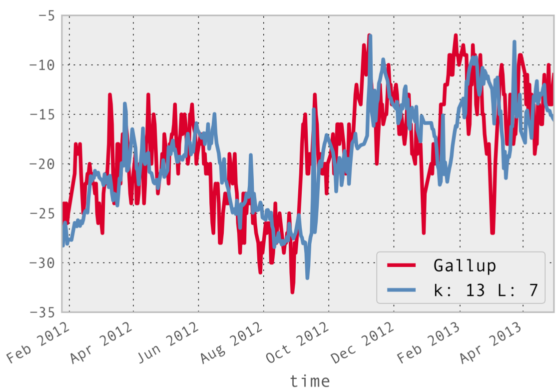

Figure 10. Rolling forecast of Gallup economic confidence index using Smile Index

Performance for alternative lag parameters is shown in Table 6. For $k = 13$, the 7 days ahead forecast has a mean absolute error of 3.53. We compare that result with a baseline model that uses endogenous values for prediction. In other words, the baseline model uses the lagged self-value for prediction ($EI_{t+L} = \alpha + \beta \sum_{i=0}^{k-1} EI_{t-i} + \epsilon_t$). We find that this baseline performs slightly worse with a mean absolute error of 3.55.

	$L = 7$	$L = 10$	$L = 14$
$k = 7$	3.70	3.95	4.19
$k = 13$	3.53	3.68	4.10
$k = 10$	3.61	3.75	4.05

Table 6. Mean Absolute Error (MAE) for rolling forecast of economic confidence (L = # of days in the future we forecast, k = # of days of historical data used to train model)

Overall, the performance of this simple model thus validates our finding from the Granger analysis: smile-containing images hold predictive information about the economic confidence index. The performance of the model is also encouraging since the rich content of social media could be used as a faster and lower-cost supplement for expensive traditional polling. Given the deep reach and broader scope of social media, being able to combine such signals in this way would not only allow us to measure well-being but also to analyze factors that may contribute to it.

DISCUSSION

Methodological Implications

In this study, we set out to explore the potential of using pictures shared on social media to assess collective well-being. By applying computer vision techniques and calculating the proportion of photos in which we detect smiling faces, we proffer our methodological contribution: the *Smile Index* — a quantifiable proxy for happiness that capitalizes on the ever-expanding supply of social media hosted images, which are shared by users frequently, naturally, and globally.

Examining temporal trends, our Smile Index reveals patterns of happiness that are highest during mornings, evenings, and weekends and that drop during late night, mid-day, and mid-week. In our case studies on four emotionally significant U.S. events, we verify that the Smile Index considerably increases in response to celebratory events and holidays, while it substantially sinks during tragedies and disasters. In connecting emotional health with demographic and economic indicators, we find positive correlations between income and positive sentiment in white-majority geographical areas yet the reverse relationship in black-majority areas. In testing the predictive power of the Smile Index, we demonstrate the metric's ability to forecast consumer confidence 7 days ahead using just 13 days of historical data. Altogether, these results showcase our approach's potential for monitoring population-level emotion.

Further, our methodology can benefit stakeholders from multiple domains including CSCW disciplines, government, and business. For example, a primary intended outcome of our work was unlocking a rich data source and offering a new lens through which future research can extend existing understandings of human behavior, emotion, and communities of practice. In particular, our work helps advance computational social science's aim to leverage social media as part of data-driven social science research, which continues to gain interest within CSCW. Along the same lines, our work further provides novel image-based computational techniques that scholars can apply to develop psychological and sociological theory.

Our findings, especially those regarding the relationship between events and public mood, highlight the insight our method provides about one such notable behavioral dynamic: how people's offline activities and emotional states impact their online behaviors. In this way, social media's ability to reflect the well-being of real world communities has important implications for the role such online communication platforms can play in connecting leaders and constituents in order to help improve people's quality of life and welfare, for example through better public health intervention, allocation of public services, or regional planning.

Looking for evidence of the reverse phenomenon would be desirable as well — that is, to what extent being exposed online to emotionally-charged content influences emotions and behaviors in the physical world. Prior work on mimicry and emotional contagion find that viewing pictures containing smiles can influence a person's own happiness [39, 65], and it would therefore be worthwhile to explore whether pictures containing smiles not only signal current levels of collective sentiment but also are responsible for stimulating similar emotions in other users or even fueling emotional cascades. Researchers may similarly investigate whether content containing pictures has more influence over other users' behavior than purely text-based content does, for instance in posts that encourage users to join protests, charity events, or other forms of collective action. This sort of study could also make space for new generations of personalized early warning systems that inform people about major events that will personally impact them, based on local patterns of emotional arousal trending in their social media streams.

Practical Implications
From a practical perspective, our proposed system additionally offers ways to overcome a number of limitations faced by current methods for assessing sentiment. First, in contrast to survey-based methodologies — which may struggle with low response rates, are typically administered infrequently and retrospectively, and can be difficult to scale — our approach is unobtrusive, enables live and continuous affective monitoring, and provides access to ever-wider global populations as social media adoption continues to grow. Such real-time monitoring, which our prediction and forecasting analyses demonstrate is feasible, can have particularly important implications during times of crisis, for instance when communities require prompt and accurately targeted aid. Second, current NLP techniques for text-based sentiment analysis are language-dependent, and most available corpora and dictionaries only exist for the English language. This impedes not only international-level studies but even research within the United States. For instance, an estimated 200 languages are spoken in New York City alone [78], and studies using sentiment analyses restricted to text in English may be biasing results and missing opportunities to investigate the full emotional characteristics of such highly diverse regions. Again in contrast, our text (and hence language) independent method facilitates an analysis of emotions expressed by individuals spanning multiple languages and cultures.

Cultural Differences in Emotion and Photography
To further push on this idea of using photographed smiles to assess sentiment cross-culturally, we investigate viability by comparing our Smile Index across countries. For the 15 countries with more than $100,000$ tweets in our dataset, we find it to be relatively consistent, as shown in Figure 11. Additionally, the standard deviation of the Smile Index across these countries is 0.02, and this low value further suggests that our proposed sentiment assessment framework would be applicable for use with countries beyond the US. These results provide some evidence that the universal nature of smiling would permit the use of our system cross-culturally, though researchers should be aware of cultural variations.

While the universality of emotion across cultures has been long debated by psychologists [21, 55], an extensive survey of the related literature by Elfenbein et al. [22] concludes that the core six emotions are universally recognized. Key to our research is that even researchers critical of universality

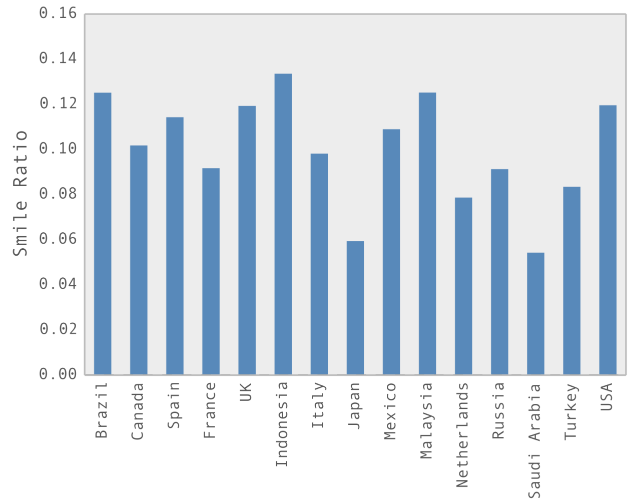

Figure 11. Smile ratio across countries with more than $100,000$ tweets in our dataset. Low standard deviation (0.02) indicates that Smile Index is consistent, highlighting the potential of using our framework in cross-cultural setups

models find that facial expression for *happiness* is consistent across cultures [34].

Regarding the influence of culture on photography, East Asians' facial expressions in particular tend to display lower intensity in photographs, and images from East Asia also focus more on contextual details. This phenomenon is observed on social media as well — for instance, East Asian users often de-emphasize faces in Facebook profile pictures, resulting in a smaller face-to-background size ratio compared to users from the United States [46]. Nonetheless, vital to our research is the determination that the *ratio of smiling and non-smiling images is the same across cultures* [32] — though as seen in Figure 11, photos from Japan do exhibit lower smile ratios as expected.

Overall, such prior work as well as our own preliminary inspection are a sign that using smile features to build an index of happiness could improve cross-cultural sentiment analysis, especially compared to word-based approaches, which are tightly coupled with specific languages and obligate complicated feature engineering such as cross-language dictionaries [3] or multilingual annotations [1] to study the sentiment of speakers with different languages.

Future Work
This paper illustrates the merits of our current study, but as a nascent step towards using pictures as a measurement tool for societal happiness, further attention is required to address some limitations.

The first is a question of population bias. While the wide adoption of social media platforms like Twitter increasingly offers the potential to access broad groups of people, at this point concerns can be raised about the issue of representativeness for these sorts of datasets. Around 15% of online adults in the U.S. use Twitter, resulting in a user base in which younger generations and minorities are overrepresented compared to the overall population [58]. Combining this with the fact that we have access to only 10% of all posts, we acknowledge that our dataset is a non-uniform subsample from a non-representative population and that our happiness metric is based on the sentiments of some people more than others. Especially since a core motivation of our work is the creation of techniques with which to perform global and cross-cultural assessments of sentiment, a key next step is therefore extending our study to ensure more representative coverage and to utilize additional datasets, including from countries around the world of diverse languages and cultures.

Second, it is also necessary to more fully investigate the quality of image-based data streams for sentiment analysis. Though extant research and our own comparison studies suggest our data is reliable in terms of the authenticity of positive affect that detected smiles express, there are some complex questions herein that require deeper investigation. One is whether the social convention of smiling while being photographed might artificially inflate our smile based happiness index. That is, smiling might not always indicate good mood but rather be the result of nuanced posing etiquette and customs, which can vary with culture, gender, and age as discussed in earlier portions of the paper. On the reverse hand, pictures suggestive of happiness may not always contain smiles or even faces or people, necessitating strategies to detect alternate signals of sentiment in photos, such as sustainable crowdsourcing techniques for labeling emotions in photos. Another issue is the extent to which the images people share on social media are pictures of themselves and friends as opposed to popular memes that may contain smiles but not be truly reflective of the contributing user's own emotional state. In our earlier description of our dataset, we also noted potential complications due to discrepancies between when and where images are captured vs. posted, though we explained why this concern should likely be minimal.

Alongside these issues regarding the reliability of image data, it is worth pointing out that the *existing* approaches to happiness measurement suffer from their own analogous challenges. For instance, text-based sentiment analysis techniques typically struggle with syntactic errors and properly handling linguistic mechanisms like sarcasm, and survey-based approaches must deal with response veracity and can suffer from reflective self-report biases. Going forward, it would be worthwhile to make further comparisons between various image, text, and survey-based methods for happiness assessment to determine the key advantages and limitations of each approach, with a focus on combining the methods in ways most appropriate for the phenomenon being studied.

Finally, in this study our focus was on developing a new technique for *measuring* happiness, so it is desirable to next look more deeply into what factors may actually *generate* that happiness. While our findings demonstrate a number of clear correlations between smiling and sentiment, we cannot yet claim causation without such further inquiry. Similarly, careful examination of potential confounding factors, such as the effect of education on income and happiness, will help to strengthen interpretations regarding the reasons underlying observed emotional characteristics.

Given these concerns, we therefore emphasize the need to further unravel our findings by undertaking additional qualitative work. Such research could involve more rigorous analyses of the images themselves to identify subtleties in the expression of emotion in photographs. Interview studies could also be conducted with individuals who contribute image content in order to better understand the motivations behind their sharing practices and the authenticity of emotion displayed in their pictures, as well as the extent of geographical and temporal differences between when they capture vs. post images, and finally their reactions to having their content used as part of automatic sentiment tracking.

CONCLUSION

In this paper, we presented a novel methodology for utilizing user-contributed pictures on social media to infer public happiness. We provided evidence of the potential to assess population-wide sentiment from such large-scale, geographically diverse, and temporally dispersed image-based content.

In the first attempt to use images for large-scale sentiment analysis, we leveraged established computer vision techniques to detect smiling faces from photos posted on Twitter in order to track societal patterns of happiness, sense underlying political and cultural events, and uncover community characteristics. We verified that this data source aligns well with established survey-based methods for assessing public mood and yet is less costly to obtain in terms of time, financing, and human resources. In addition to this comparison with ground-truth data, we also compared our technique to traditional text-based approaches for analyzing sentiment, laid out strengths and weaknesses of each, and considered how they might be combined.

Through multiple analytic experiments and case studies, we explored how publicly posted pictures can reveal temporal trends in happiness, serve as an indicator of mood during societal events, and be used to study relations between socioeconomic factors. Finally, we discussed the theoretical and practical implications of our research along with directions for future work, giving particular attention to the issues of overcoming population bias, achieving cross-cultural generalizability, and remaining cognizant of the social contexts in which individuals who share photos are embedded.

Acknowledgements

The authors would like to thank Chris Danforth and Peter Dodds from UVM's Computational Story Lab for providing valuable comments and making the research data available.

Development of these methods for non-US data was originally supported by the Intelligence Advanced Research Projects Activity (IARPA) via DoI/NBC contract number D12PC000337. The US Government is authorized to reproduce and distribute reprints of this work for Governmental purposes notwithstanding any copyright annotation thereon. Application of these methods to US data was undertaken by the authors on their own.

REFERENCES

1. Boiy, E., and Moens, M.-F. A machine learning approach to sentiment analysis in multilingual web texts. *Information retrieval 12*, 5 (2009), 526–558.
2. Bollen, J., Mao, H., and Pepe, A. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *ICWSM* (2011).
3. Boyd-Graber, J., and Resnik, P. Holistic sentiment analysis across languages: multilingual supervised latent Dirichlet allocation. In *EMNLP'10*, ACL (2010), 45–55.
4. Bradley, M. M., and Lang, P. J. Affective norms for english words (anew): Instruction manual and affective ratings (1999).
5. Campbell, A., Converse, P. E., and Rodgers, W. L. *The quality of American life: Perceptions, evaluations, and satisfactions*, vol. 3508. Russell Sage Foundation, 1976.
6. Clark, A. Happiness, habits and high rank: Comparisons in economic and social life.
7. Clark, A. E., Frijters, P., and Shields, M. A. Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature* (2008), 95–144.
8. Clark, L. A., Watson, D., and Leeka, J. Diurnal variation in the positive affects. *Motivation and Emotion 13*, 3 (1989), 205–234.
9. Cohn, M. A., Mehl, M. R., and Pennebaker, J. W. Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science 15*, 10 (2004), 687–693.
10. Coverdill, J. E., López, C. A., and Petrie, M. A. Race, ethnicity and the quality of life in america, 1972–2008. *Social Forces 89*, 3 (2011), 783–805.
11. De Choudhury, M., Counts, S., and Horvitz, E. Major life changes and behavioral markers in social media: Case of childbirth. CSCW (2013), 1431–1442.
12. De Choudhury, M., Counts, S., Horvitz, E. J., and Hoff, A. Characterizing and predicting postpartum depression from shared facebook data. CSCW (2014), 626–638.
13. De Choudhury, M., Monroy-Hernández, A., and Mark, G. "Narco" emotions: Affect and desensitization in social media during the mexican drug war. CHI, ACM (2014), 3563–3572.
14. Di Tella, R., and MacCulloch, R. Some uses of happiness data in economics. *The Journal of Economic Perspectives 20*, 1 (2006), 25–46.
15. Diener, E. Subjective well-being. *Psychological Bulletin 95* (1984), 542–575.
16. Diener, E., Emmons, R. A., Larsen, R. J., and Griffin, S. The satisfaction with life scale. *Journal of personality assessment 49*, 1 (1985), 71–75.
17. Dodds, P. S., and Danforth, C. M. Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. *Journal of Happiness Studies 11*, 4 (2010), 441–456.
18. Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., and Danforth, C. M. Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PloS one 6*, 12 (2011).
19. Ekman, P. An argument for basic emotions. *Cognition & Emotion 6*, 3-4 (1992), 169–200.
20. Ekman, P., and Friesen, W. V. Facial action coding system: A technique for the measurement of facial movement. *Consulting Psych. Press* (1978), 271–302.
21. Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., and Ricci-Bitti, P. E. Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology 53*, 4 (1987), 712.
22. Elfenbein, H. A., and Ambady, N. On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological bulletin 128*, 2 (2002).
23. Ferrer-i Carbonell, A. Income and well-being: an empirical analysis of the comparison income effect. *Journal of Public Economics 89*, 5 (2005), 997–1019.
24. Fowler, J. H., Christakis, N. A., Steptoe, and Roux, D. Dynamic spread of happiness in a large social network: longitudinal analysis of the framingham heart study social network. *British medical journal* (2009), 23–27.
25. Girardin, F., Calabrese, F., Fiore, F. D., Ratti, C., and Blat, J. Digital footprinting: Uncovering tourists with user-generated content. *Pervasive 7*, 4 (2008), 36–43.
26. Golder, S. A., and Macy, M. W. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science 333*, 6051 (2011), 1878–1881.
27. Hall, J. A., LeBeau, L. S., Reinoso, J. G., and Thayer, F. Status, gender, and nonverbal behavior in candid and posed photographs: A study of conversations between university employees. *Sex Roles 44* (2001), 677–692.
28. Harker, L., and Keltner, D. Expressions of positive emotion in women's college yearbook pictures and their relationship to personality and life outcomes across adulthood. *Personality and social psych. 80*, 1 (2001), 112.
29. Hernandez, J., Hoque, M. E., Drevo, W., and Picard, R. W. Mood meter: counting smiles in the wild. In *UbiComp*, ACM (2012), 301–310.
30. Hess, U., Beaupré, M. G., Cheung, N., et al. Who to whom and why: cultural differences and similarities in the function of smiles. 187.
31. Hromada, D. D., Tijus, C., Poitrenaud, S., and Nadel, J. Zygomatic smile detection: The semi-supervised Haar training of a fast and frugal system: A gift to opencv community. IEEE (2010), 1–5.

32. Huang, C.-M., and Park, D. Cultural influences on facebook photographs. *International Journal of Psychology 48*, 3 (2013), 334–343.
33. Hughes, M., and Thomas, M. E. The continuing significance of race revisited: A study of race, class, and quality of life in america, 1972 to 1996. *American Sociological Review* (1998), 785–795.
34. Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., and Schyns, P. G. Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences 109*, 19 (2012), 7241–7244.
35. Joshi, D., Datta, R., Fedorovskaya, E., Luong, Q.-T., Wang, J. Z., Li, J., and Luo, J. Aesthetics and emotions in images. *Signal Processing Mag. 28*, 5 (2011), 94–115.
36. Keltner, D. Ekman, emotional expression, and the art of empirical epiphany. *Journal of Research in Personality 38*, 1 (2004), 37–44.
37. Kim, E., Gilbert, S., Edwards, M. J., and Graeff, E. Detecting sadness in 140 characters: Sentiment analysis and mourning Michael Jackson on twitter. *Web Ecology 3* (2009).
38. Kivran-Swaine, F., and Naaman, M. Network properties and social sharing of emotions in social awareness streams. CSCW, ACM (2011), 379–382.
39. Kleinke, C. L., Peterson, T. R., and Rutledge, T. R. Effects of self-generated facial expressions on mood. *Jour. Personality & Social Psych. 74*, 1 (1998), 272.
40. Kotchemidova, C. Why we say “cheese”: Producing the smile in snapshot photography. *Critical Studies in Media Communication 22*, 1 (2005), 2–25.
41. Kramer, A. D. An unobtrusive behavioral model of gross national happiness. In *CHI*, ACM (2010), 287–290.
42. Levenson, R. W., Ekman, P., and Friesen, W. V. Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology 27*, 4 (1990), 363–384.
43. Lienhart, R., Kuranov, A., and Pisarevsky, V. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. *Pattern Rec.* (2003), 297–304.
44. Logunov, A., and Panchenko, V. Characteristics and predictability of twitter sentiment series. In *Internat. Congress on Modelling and Simulation* (2011).
45. Maruyama, M. T., Robertson, S. P., Douglas, S. K., Semaan, B. C., and Faucett, H. A. Hybrid media consumption: How tweeting during a televised political debate influences the vote decision. CSCW, ACM (2014), 1422–1432.
46. Masuda, T., Gonzalez, R., Kwan, L., and Nisbett, R. E. Culture and aesthetic preference: Comparing the attention to context of east asians and americans. *Personality & Social Psych. Bull. 34*, 9 (2008), 1260–75.
47. McNair, D., Lorr, M., and Droppleman, L. Profile of mood states. *San Diego, California 53* (1992), 6.
48. Mihalcea, R., and Liu, H. A corpus-based approach to finding happiness. In *AAAI, Computational Approaches to Analyzing Weblogs* (2006), 139–144.
49. Mishne, G., and de Rijke, M. Capturing global mood levels using blog posts. In *AAAI, Computational Approaches to Analyzing Weblogs* (2006), 145–152.
50. Monk, T., Buysse, D., Reynolds, C., Berga, S., Jarrett, D., Begley, A., and Kupfer, D. Circadian rhythms in human performance and mood under constant conditions. *Journal of sleep research 6*, 1 (1997), 9–18.
51. O’Connor, B., Balasubramanyan, R., Routledge, B. R., and Smith, N. A. From tweets to polls: Linking text sentiment to public opinion time series. *ICWSM’10*, 122–9.
52. Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., and Booth, R. J. The development and psychometric properties of liwc2007. *liwc.net* (2007).
53. Quercia, D. Don’t worry, be happy: The geography of happiness on facebook. In *Proceedings of the 5th Annual ACM Web Science Conference*, ACM (2013), 316–325.
54. Quercia, D., Ellis, J., Capra, L., and Crowcroft, J. Tracking gross community happiness from tweets. In *Proc. CSCW’12*, ACM (2012), 965–968.
55. Russell, J. A. Is there universal recognition of emotion from facial expressions? a review of the cross-cultural studies. *Psychological bulletin 115*, 1 (1994), 102.
56. Sakaki, T., Okazaki, M., and Matsuo, Y. Earthquake shakes twitter users: real-time event detection by social sensors. In *WWW*, ACM (2010), 851–860.
57. Seder, J. P., and Oishi, S. Intensity of smiling in facebook photos predicts future life satisfaction. *Social Psychological & Personality Sci. 3*, 4 (2012), 407–413.
58. Smith, A., and Brenner, J. Twitter use 2012. pew.
59. Stone, A. A., Hedges, S. M., Neale, J. M., and Satin, M. S. Prospective and cross-sectional mood reports offer no evidence of a “blue monday” phenomenon. *Journal of Personality and Social Psychology 49*, 1 (1985), 129.
60. Tomkins, S. S. Affect, imagery, consciousness.
61. Tsujita, H., and Rekimoto, J. Smiling makes us happier: enhancing positive mood and communication with smile-encouraging digital appliances. In *Proceedings of the 13th international conference on Ubiquitous computing*, ACM (2011), 1–10.
62. Viola, P., and Jones, M. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition*, vol. 1, IEEE (2001), I–511.
63. Wang, W., Chen, L., Thirunarayan, K., and Sheth, A. P. Cursing in english on twitter. CSCW (2014), 415–25.
64. Whitehill, J., Littlewort, G., Fasel, I., Bartlett, M., and Movellan, J. Toward practical smile detection. *Pattern Analysis and Machine Intelligence* (2009), 2106–2111.

65. Wild, B., Erb, M., and Bartels, M. Are emotions contagious? evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences. *Psychiatry research 102*, 2 (2001), 109–124.

66. Xia, C., Schwartz, R., Xie, K., Krebs, A., Langdon, A., Ting, J., and Naaman, M. Citybeat: real-time social media visualization of hyper-local city data. In *WWW* (2014), 167–170.

67. Xie, K., Xia, C., Grinberg, N., Schwartz, R., and Naaman, M. Robust detection of hyper-local events from geotagged social media data. In *13th Internat. Workshop on Multimedia Data Mining*, ACM (2013), 2.

68. Bhutanese gross national happiness commission. `http://www.gnhc.gov.bt/`.

69. Facebook has 220 billion of your photos to put on ice. `http://gigaom.com/2012/10/17/facebook-has-220-billion-of-your-photos-to-put-on-ice`.

70. Gallup Healthways Well-Being Index. `http://www.gallup.com/poll/106915/gallup-daily-us-mood.aspx`.

71. Global mobile statistics 2014 part A. `http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats/a`.

72. Gross national happiness. `http://www.grossnationalhappiness.com/`.

73. Gross national happiness USA. `http://www.gnhusa.org/`.

74. Guangdong to use index to gauge happiness. `http://china.org.cn/china/2011-10/12/content_23601655.htm`.

75. Hurricane/post-tropical cyclone sandy service assessment. `http://www.nws.noaa.gov/os/assessments/pdfs/Sandy13.pdf`.

76. Ict facts and figures. `http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf`.

77. Locals and tourists. `https://www.flickr.com/photos/walkingsf/sets/72157624209158632/`.

78. NYC planning - population facts. `http://www.nyc.gov/html/dcp/html/census/pop_facts.shtml`.

79. Oecd guidelines on measuring subjective well-being. `http://www.oecd.org/statistics/guidelines-on-measuring-subjective-well-being.htm`.

80. Office for national statistics - measuring national well-being. `http://www.ons.gov.uk/ons/guide-method/user-guidance/well-being/index.html`.

81. Photo and video sharing grow online. *PEW 2013*.

82. Planet selfie. `http://www.businessinsider.com/were-now-posting-a-staggering-18-billion-photos-to-social-media-every-day-2014-5`.

83. State of American Well-Being. `http://info.healthways.com/wbi2013`.

84. The lottery of life. `http://www.economist.com/news/21566430-where-be-born-2013-lottery-life`.

85. The role for well-being in creating the big society. `http://cles.org.uk/wp-content/uploads/2011/01/LW-105-Well-Being-and-Big-Society1.pdf`.

86. World happiness report 2013. `http://unsdsn.org/resources/publications/world-happiness-report-2013/`.