Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

# Using Big Data to study subjective well-being

Maike Luhmann

Subjective well-being comprises emotional experiences and life satisfaction. This article reviews how Big Data can be used to measure, study, and change subjective well-being. Most Big Data approaches measure subjective well-being by analyzing language patterns on Twitter or Facebook. These approaches provide satisfactory accuracy for emotional experiences, but not yet for life satisfaction. Other measurement approaches include the analysis of other digital traces such as Facebook profiles and the analysis of mobile phone usage patterns. Big Data can be used to study subjective well-being on individual levels, regional levels, and across time. Potentials and limitations of using Big Data in studies on subjective well-being are discussed.

**Address**
Ruhr University Bochum, Department of Psychology, Universitaetsstrasse 150, 44780 Bochum, Germany

Corresponding author: Luhmann, Maike (maike.luhmann@rub.de)

**Current Opinion in Behavioral Sciences** 2017, **18**:28–33

This review comes from a themed issue on **Big data in the behavioural sciences**

Edited by **Michal Kosinski** and **Tara Behrend**

For a complete overview see the Issue and the Editorial

Available online 21st July 2017

http://dx.doi.org/10.1016/j.cobeha.2017.07.006



## Introduction

The science of subjective well-being (SWB) focuses on the definition, measurement, and correlates of happiness. Subjective well-being encompasses people's emotional experiences (i.e. positive and negative emotions and moods) as well as their evaluations of their lives (i.e. life satisfaction) [1]. SWB is an inherently subjective experience, meaning that each person knows best whether he or she is happy. Consequently, SWB has traditionally been measured with self-reports [2,3]. Self-report measures of SWB exhibit adequate reliability and validity [4] but can be distorted by irrelevant factors such as item order, momentary mood, or motivated self-enhancement [4–6]. Although the effects of these factors tend to be weak [4,7], researchers have nevertheless tried to improve the measurement of SWB through modifications of the research design (e.g. using experience sampling methodology instead of single-occasion surveys) as well as through non-self-report measures (e.g. peer reports or psychophysiological indicators such as cortisol levels and facial expressions). However, studies using these alternative measures remain rare compared to studies relying on self-reports. Moreover, because all of these measures are rather expensive to collect, studies on SWB are often limited by a low temporal and/or geographical resolution and by small and selective samples.

Big Data offer new opportunities to study SWB in ways that circumvent some of these limitations. This paper offers an overview of how Big Data are currently used to measure, study, and change SWB.

## Using Big Data to measure SWB

Approaches to measure SWB using Big Data can be distinguished in terms of data source, measurement level, and SWB facet (Table 1).

### Data sources

The predominant measurement approach relies on the analysis of so-called digital traces. Digital traces comprise all recorded online activities of an individual that can be accessed through publicly available databases (e.g. Twitter, Google Trends) or by obtaining individuals' permission (e.g. private Facebook profiles). Twitter is particularly popular because it allows researchers to access massive amounts of data quickly, cost-effectively, and without having to obtain informed consent from the users. Tweets are analyzed with respect to language patterns that predict SWB (e.g. [8–11]). This methodology has also been applied to other texts such as Facebook status updates (e.g. [12–14]).

Language patterns are studied using either a closed-vocabulary approach or an open-vocabulary approach [15]. The closed-vocabulary approach scans the text for keywords that are retrieved from a predefined dictionary. Sometimes, these dictionaries contain only a few dozen highly specific words [16,17]. Most dictionaries, however, comprise thousands of words. For example, the LabMT dictionary that is used in a few studies [9,18] contains 10,000 words that were rated by Amazon Mechanical Turk workers on a scale from sad to happy [19]. The most popular dictionary in SWB research, however, is Linguistic Inquiry and Word Count (LIWC) [20] which comprises >6000 words assigned to one or more dimensions. The dimensions most frequently used by SWB researchers are positive and negative emotion words and dimensions related to specific emotions such as sadness or anger. LIWC is easy to use and widely accepted as a validated method to study language patterns. However, as

**Table 1**

**Overview of measurement approaches.**

| Publication | General approach | Data source | SWB facet | Measurement level |
|---|---|---|---|---|
| Algan *et al.* [47] | Frequency of specific search terms on Google | Google Trends | Life satisfaction<br>Emotional well-being | Longitudinal trends within one nation (United States) |
| Carlquist *et al.* [16] | Closed vocabulary (self-constructed lexicon) | Newspaper articles | Emotional well-being | Longitudinal trends within one nation (Norway) |
| Collins *et al.* [12] | Online activity; closed vocabulary (LIWC) | Facebook status updates and likes | Life satisfaction | Individual |
| Curini *et al.* [45] | Open vocabulary | Twitter | Emotional well-being | Italian provinces |
| Doré *et al.* [46••] | Closed vocabulary (LIWC) | Twitter | Emotional well-being | Local with exact coordinates |
| Durahim & Coskun [44] | Closed vocabulary (SentiStrength, Turkish version) | Twitter | Emotional well-being | Turkish provinces and national |
| Hao *et al.* [24] | Open vocabulary | Weibo posts | Emotional well-being | Individual |
| Hung *et al.* [34] | Mobile phone usage | Calling states, app usage | Emotional well-being | Individual |
| Jones *et al.* [41] | Closed vocabulary (LIWC) | Twitter | Emotional well-being | Regional |
| Kosinski *et al.* [26] | Online activity | Facebook likes | Life satisfaction | Individual |
| Kramer [40] | Closed vocabulary (LIWC) | Facebook status updates | Life satisfaction | Regional |
| Lee *et al.* [27] | Online activity | Daily activity on Facebook and Twitter | Emotional well-being | Individual |
| LiKamWa *et al.* [39] | Mobile phone usage | Call and messaging logs, app usage, location | Emotional well-being | Individual |
| Liu *et al.* [13] | Closed vocabulary (LIWC) | Facebook status updates | Life satisfaction | Individual |
| MacKerron & Mourato [33] | Mobile phone app | Location data as basis for data type of location, weather | Emotional well-being | Individual |
| Mitchell *et al.* [18] | Closed vocabulary (LabMT) | Twitter | Emotional well-being | Neighborhood (U.S. cities) |
| Miura *et al.* [17] | Closed vocabulary (self-constructed lexicon) | Twitter | Emotional well-being | Japanese regions |
| Nguyen *et al.* [23] | Open vocabulary | Twitter | Emotional well-being | U.S. ZIP codes |
| Nguyen *et al.* [9] | Closed vocabulary (LabMT) | Twitter | Emotional well-being | Neighborhoood (3 U.S. cities) |
| Prata *et al.* [43] | Open vocabulary | Twitter | Emotional well-being | Precise location in Brazil |
| Rickard *et al.* [38•] | Combination of language analysis, other online activities | Facebook and Twitter activity and posts | Emotional well-being | Individual |
| Saeb *et al.* [36] | Mobile phone usage | Location, usage patterns | Emotional well-being | Individual |
| Schwartz *et al.* [8] | Combination of closed vocabulary (LIWC) and open vocabulary | Twitter | Life satisfaction | U.S. counties |
| Schwartz *et al.* [14] | Open vocabulary | Facebook status updates | Life satisfaction | Individual |
| Settanni & Marengo [51] | Closed vocabulary (LIWC) | Facebook status updates | Emotional well-being | Individual |
| Volkova & Bachrach [22] | Open vocabulary | Twitter | Emotional well-being | Individual |
| Volkova *et al.* [42] | Open vocabulary | Twitter | Emotional well-being | Neighborhood (U.S. universities) |
| Wang *et al.* [52] | Closed vocabulary (LIWC) | Facebook status updates | Life satisfaction | Individual |
| Wang *et al.* [11] | Closed vocabulary (LIWC) | Twitter | Emotional well-being | National |
| Wojcik *et al.* [56•] | Closed vocabulary (LIWC) for language analysis | Twitter, LinkedIn, etc. | Emotional well-being | Individual |
| Yang & Srinivasan [10] | Closed vocabulary (LIWC) | Twitter | Life satisfaction | Individual |
| Yu & Wang [54] | Closed vocabulary (NRC Word-Emotion Association lexicon) | Twitter | Emotional well-being | National |

all dictionaries, it comprises a limited number of words and does not work well for colloquial language, leading some to question its validity for the analysis of social media language [21••].

Open-vocabulary approaches, in contrast, do not rely on an a priori defined list of words but rather identify relevant words, topics (i.e. clusters of words used frequently together), and other linguistic features using a data-driven approach [8,14,22–24]. Open-vocabulary-based predictions of SWB tend to exhibit greater predictive validity than closed-vocabulary-based predictions. In one exemplary study, life satisfaction correlated weaker with words identified using a closed-vocabulary approach (cross-validated $r = .26$) than with topics identified using an open-vocabulary approach (cross-validated $r = .31$) [14]. However, closed-vocabulary approaches require more sophisticated statistical skills and computational power, and can yield results that are harder to interpret because they may lack face validity [15].

Digital traces are not limited to user-generated texts. Facebook profiles, for example, include information on demographics, political and sexual orientation, preferences (e.g. 'likes'), social networks, profile pictures, etc. [25]. To date, few studies have tapped into these data sources to measure SWB. In one exemplary study, self-reported life satisfaction was predicted by Facebook likes; however, the correlation between predicted and self-reported life satisfaction was low (cross-validated $r = .17$) [26]. Another small-scale study predicted daily mood fluctuations using data on daily activities on Twitter and Facebook, including frequency of posting, commenting on other posts, likes, etc., but the results were inconsistent across participants, with correlations between online activity and mood being positive for some, negative for others, and non-significant for a third group of participants [27]. Better results were achieved in a study linking smiling on profile pictures to self-reported life satisfaction ($r = .34$) [28].

Another promising data source are mobile phones and other wearable devices such as smartwatches or fitness trackers. As part of the 'quantified self' movement [29••], more and more people regularly record their momentary SWB outside of the context of specifically designed research studies. In addition to accumulating massive amounts of self-report data on SWB and other variables (e.g. [30]), some of these apps can be integrated with experimental paradigms [31] and allow the unobtrusive collection of data on, for example, location, phone usage patterns, or context information (e.g. noise, light) [32,33]. First results on using such non-self-report data to predict momentary mood are promising, but with few exceptions [33], most of these studies were based on very small samples [34–37,38•,39] and need to be replicated with bigger samples.

### Measurement level

Most data sources discussed above use data that are produced by individuals and can therefore in principle be used to measure individual SWB. For accurate individual predictions, however, a sufficient number of individual data points must be available, for instance, a specific number of tweets or likes [15] or a specific number of data points on different combined variables. But even if these requirements are not met, data collected from individuals can still be aggregated within geographic regions (if they are geo-tagged) or within time periods (if they are time-stamped). One of the first attempts to use Big Data to measure the geographic distribution of SWB (so-called 'gross national happiness') was based on Facebook status updates [40]. To date, Twitter data are particularly popular to study SWB on regional levels [8,9,23,41–45] and over time [45,46••]. Another promising data source is Google Trends which provides data on the frequency of specific search terms over time. One study using these data found strong correlations between SWB trends observed in self-report data and SWB trends predicted by search terms (cross-validated $rs \geq .85$) [47].

With these studies, it is important to acknowledge that individual information gets lost in the aggregation process. In particular, one needs to be careful to interpret effects only on the level on which they are measured and to not draw inferences from the individual level to the aggregate level or vice versa (ecological fallacy) [48].

### Emotional well-being versus life satisfaction

Emotional well-being and life satisfaction are correlated but conceptually and functionally distinct [49,50]. As Table 1 illustrates, most approaches tap into emotional aspects of well-being. These approaches appear to yield valid indicators of emotional well-being, as indicated by studies showing that these alternative measures fluctuate in similar weekly and diurnal rhythms as self-reported emotional well-being [19], and by studies finding moderate correlations ($rs$ between .20 and .40) of these measures with self-reported emotional well-being [51]. For life satisfaction, however, the correlations between self-reported scores and alternative indicators are typically weaker ($rs < .20$) [12,13,26,52], suggesting that individual-level life satisfaction cannot yet be predicted reliably from people's digital traces.

### Summary

Big Data provide multiple data sources that can potentially be used to measure SWB. To date, no single data source seems reliable and valid enough to replace traditional self-report measures of well-being. However, this may change as more data sources are developed, validated, and combined. Studies combining language pattern analysis with indicators such as likes or online activity patterns consistently find that these combined indicators predict SWB more accurately than any single indicator [12,14,24,38•].

## Using Big Data to study correlates of SWB

Major events such as natural disasters or terrorist attacks are hard to study because they are unpredictable. In the past, prospective studies that examined the impact of such events on SWB had to rely on data that happened to be collected around the time of the event [53]. Because these studies had not been designed to study the impact of major events on SWB, they were often associated with limitations such as small samples and inadequately spaced measurements. These issues can partially be resolved by using Big Data indicators of SWB, particularly tweets. Twitter data have been used to study changes in emotional well-being after earthquakes [17] and violence in schools and colleges [41,46••].

Twitter data can also be used to study real-time emotional experiences, something that is not practical with traditional measures of SWB. For example, Yu and Wang analyzed emotions in tweets during 2014 World Cup soccer matches and found that the emotions in these tweets matched what actually happened on the field [54].

Finally, Big Data can be used in multimethod studies to supplement traditional measures of SWB. The logic of multimethod studies is that evidence for an effect is stronger if that effect can be found across multiple different methods [55]. This is not always the case, as one study on the correlation between political orientation and SWB illustrates [56•]. In this study, conservatives scored higher than liberals on self-reported SWB (replicating earlier work) but lower on Big Data measures of SWB such as emotional language on Twitter or smiling in profile pictures. These findings do not necessarily speak against the validity of Big Data measures of SWB but rather suggest that Big Data and self-reports measure different aspects of SWB, something that needs to be explored in future research.

## Using Big Data for SWB interventions

Probably the greatest potential for future research lies in not simply observing, but changing SWB through Big Data. Such changes can be temporary (e.g. mood manipulations in experiments) or enduring (e.g. interventions aimed at permanently improving people's SWB). One example for the former is the 'Facebook experiment' in which the Facebook newsfeed was manipulated to reduce the amount of positive or negative emotional content, which led to fewer positive or negative status updates, respectively [57]. This study has been criticized on both ethical and methodological grounds [21••]. Examples for the latter are self-tracking apps with built-in interventions such as feedback on one's self-reported mood or specific advice on how to improve one's mood. A few studies suggest that regular usage of these kinds of apps may improve SWB [35,58].

## Conclusion

Big Data provide SWB researchers with easy and cheap access to large samples that can, depending on the data source, be tracked over time and space. Big Data may therefore be particularly useful to study the impact of rare events on SWB, including rare collective events such as natural or man-made disasters as well as rare individual events such as major life events. Big Data can also be used to study SWB on regional and national levels. Finally, because most Big Data sources are social online networks such as Facebook and Twitter, these data lend themselves to studies on SWB in social networks [59].

However, it is also important to keep in mind some of the limitations of Big Data. First, whereas the emotional content of tweets and status updates appears to predict emotional well-being with adequate accuracy, satisfactory predictive models for life satisfaction are still absent. This may change, however, as the combination of different indicators improves these predictive models. Second, Big Data studies rely on individuals who are willing to provide data, for instance, active Twitter and Facebook users. These samples may differ from the general population in terms of demographic composition and other characteristics. For example, younger people use Facebook for different motives than older people, which may affect what kind of content they share on Facebook [60,61] and how well this content can be used to measure SWB [51]. Third, most studies to date used either Twitter or Facebook data. Future studies should expand the data sources. Finally, recent studies have focused on English-speaking users (for exceptions, see [17,43–45]). Studies including non-English-speaking users are necessary to evaluate to what extent the findings generalize to other populations and to identify cultural moderators that affect how well SWB can be measured using Big Data [62].

In conclusion, Big Data will probably not replace traditional self-report measures of SWB in the foreseeable future, but they can serve as a useful additional source of data in multimethod studies of SWB.

## Conflict of interest statement

Nothing declared.

## Acknowledgements

I am grateful to Nick Augustat for his assistance in the literature search and review and to Nick Augustat, Sophia Terwiel, and Kira Wolff for feedback on an earlier draft of this paper.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Diener E: **Subjective well-being**. *Psychol Bull* 1984, **95**:542-575.
2. Diener E, Emmons RA, Larsen RJ, Griffins S: **The satisfaction with life scale**. *J Pers Assess* 1985, **49**:71-75.
3. Watson D, Clark LA, Tellegen A: **Development and validation of brief measures of positive and negative affect: the PANAS scales**. *J Pers Soc Psychol* 1988, **54**:1063-1070.
4. Diener E, Inglehart R, Tay L: **Theory and validity of Life Satisfaction Scales**. *SOCI* 2013, **112**:497-527.
5. Wojcik SP, Ditto PH: **Motivated happiness: self-enhancement inflates self-reported subjective well-being**. *Soc Psychol Pers Sci* 2014, **5**:825-834.
6. Schwarz N, Strack F: **Reports of subjective well-being: judgmental processes and their methodological implications**. In *Well-being. The Foundations of Hedonic Psychology*. Edited by

Kahneman D, Diener E, Schwarz N. Russell Sage Foundation; 1999:61-84.

7. Schimmack U, Oishi S: **The influence of chronically and temporarily accessible information on life satisfaction judgments**. *J Pers Soc Psychol* 2005, **89**:395-406.

8. Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Lucas RE, Agrawal M, Park GJ, Lakshmikanth SK, Jha S, Seligman MEP *et al.*: **Characterizing geographic variation in well-being using tweets**. *Seventh International AAAI Conference on Weblogs and Social Media*. 2013:583-591.

9. Nguyen QC, Kath S, Meng HW, Li D, Smith KR, VanDerslice JA, Wen M, Li F: **Leveraging geotagged Twitter data to examine neighborhood happiness, diet, and physical activity**. *Appl Geogr* 2016, **73**:77-88.

10. Yang C, Srinivasan P: **Life satisfaction and the pursuit of happiness on Twitter**. *PLOS ONE* 2016, **11**.

11. Wang W, Hernancez I, Newman DA, He J, Bian J: **Twitter analysis: studying US weekly trends in work stress and emotion**. *Appl Psychol* 2016, **65**:355-378.

12. Collins S, Sun Y, Kosinski M, Stilwell D, Markuzon N: **Are you satisfied with life? Predicting satisfaction with life from Facebook**. In *Social Computing, Behavioral-Cultural Modeling, and Prediction*. Edited by Agarwal N, Xu K, Osgood N. Springer; 2015:24-33.

13. Liu P, Tov W, Kosinski M, Stillwell DJ, Qui L: **Do Facebook status updates reflect subjective well-being?** *Cyberpsychol Behav Soc Netw* 2015, **18**:373-379.

14. Schwartz HA, Sap M, Kern ML, Eichstaedt JC, Kapelner A, Agrawal M, Ungar LH: **Predicting individual well-being through the language of social media**. *Pac Symp Biocomput* 2016, **21**:516-527.

15. Kern ML, Park G, Eichstaedt JC, Schwartz HA, Sap M, Smith LK, Ungar LH: **Gaining insights from social media language: methodologies and challenges**. *Psychol Methods* 2016, **21**:507-525.

16. Carlquist E, Nafstad HE, Blakar RM, Ulleberg P, Delle Fave A, Phelps JM: **Well-being vocabulary in media language: an analysis of changing word usage in Norwegian newspapers**. *J Positive Psychol* 2017, **12**:99-109.

17. Miura A, Komori M, Matsumara N, Maeda K: **Expression of negative emotional responses to the 2011 Great East Japan Earthquake: analysis of big data from social media**. *Jpn J Psychol* 2015, **86**:102-111.

18. Mitchell L, Frank MR, Harris KD, Dodds PS, Danforth CM: **The geography of happiness: connecting Twitter sentiment and expression, demographics, and objective characteristics of place**. *PLoS ONE* 2013, **8**:1-15.

19. Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM: **Temporal patterns of happiness and information in a global social network: hedonometrics and twitter**. *PLoS ONE* 2011, **6**:1-26.

20. Tausczik YR, Pennebaker JW: **The psychological meaning of words: LIWC and computerized text analysis methods**. *J Lang Soc Psychol* 2010, **29**:24-54.

21. •• Panger G: **Reassessing the Facebook experiment: critical thinking about the validity of Big Data research**. *Inf Commun Soc* 2016, **19**:1108-1126.
Critical review of methodological and ethical issues associated with the Facebook experiment, including a general critique of using Facebook updates as a measure of subjective well-being.

22. Volkova S, Bachrach Y: **On predicting sociodemographic traits and emotions from communications in social networks and their implications to online self-disclosure**. *Cyberpsychol Behav Soc Netw* 2015, **18**:726-736.

23. Nguyen QC, Li D, Meng HW, Kath S, Nsoesie E, Li F, Wen M: **Building a national neighborhood dataset from geotagged Twitter data for indicators of happiness, diet, and physical activity**. *JMIR Public Health Surveill* 2016, **2**:e158.

24. Hao B, Li L, Gao R, Li A, Zhu T: **Sensing subjective well-being from social media**. In *Active Media Technology*. Edited by Ślezak D, Schaefer G, Vuong ST, Kim YS. Springer; 2014:324-335.

25. Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D: **Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines**. *Am Psychol* 2015, **70**:543-556.

26. Kosinski M, Stillwell D, Graepel T: **Private traits and attributes are predictable from digital records of human behavior**. *Proc Natl Acad Sci* 2013, **110**:5802-5805.

27. Lee JA, Efstratiou C, Bai L: **OSN mood tracking: exploring the use of online social network activity as an indicator of mood changes**. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*; *ACM: 2016*.

28. Seder JP, Oishi S: **Intensity of smiling in Facebook photos predicts future life satisfaction**. *Soc Psychol Pers Sci* 2012, **3**:407-413.

29. •• Fawcett T: **Mining the quantified self: personal knowledge discovery as a challenge for data science**. *Big Data* 2015, **3**:249-266.
Introduction and overview of quantified-self approaches (e.g. specific apps, data-analytic approaches).

30. Killingsworth MA, Gilbert DT: **A wandering mind is an unhappy mind**. *Science* 2010, **330**:932.

31. Rutledge RB, Skandali N, Dayan P, Dolan RJ: **A computational and neural model of momentary subjective well-being**. *Proc Natl Acad Sci* 2014, **111**:12252-12257.

32. Harari GM, Lane ND, Wang R, Corsier BS, Campbell AT, Gosling SD: **Using smartphones to collect behavioral data in psychological science: opportunities, practical considerations, and challenges**. *Perspect Psychol Sci* 2016, **11**:838-854.

33. MacKerron G, Mourato S: **Happiness is greater in natural environments**. *Global Environ Change* 2013, **23**:992-1000.

34. Hung GCL, Yang PC, Chang CC, Chiang JH, Chen YY: **Predicting negative emotions based on mobile phone usage patterns: an exploratory study**. *JMIR Res Protoc* 2016, **5**:e160.

35. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, Mohr DC: **Harnessing context sensing to develop a mobile intervention for depression**. *J Med Internet Res* 2011, **13**:158-174.

36. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, Mohr DC: **Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study**. *J Med Internet Res* 2015, **17**:e175.

37. Asselbergs J, Ruwaard J, Ejdys M, Schrader N, Sijbrandij M, Riper H: **Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: an explorative study**. *J Med Internet Res* 2016, **18**:e72.

38. • Rickard N, Arjmand HA, Bakker D, Seabrook E: **Development of a mobile phone app to support self-monitoring of emotional well-being: a mental health digital innovation**. *JMIR Mental Health* 2016, **3**:e49.
Example for a study that uses mobile technology to combine traditional data sources (self-reports using ESM) with novel data sources (Facebook and Twitter data).

39. LiKamWa R, Liu Y, Lane ND, Zhong L: **MoodScope: building a mood sensor from smartphone usage patterns**. *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*. ACM; 2013.

40. Kramer ADI: **An unobtrusive behavioral model of gross national happiness**. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Atlanta, Georgia: ACM; 2010, .

41. Jones NM, Wojcik SP, Sweeting J, Silver RC: **Tweeting negative emotion: an investigation of Twitter data in the aftermath of violence on college campuses**. *Psychol Methods* 2016, **21**:526-541.

42. Volkova, Han K, Corley C: **Using social media to measure student wellbeing: a large-scale study of emotional response in academic discourse**. In *Social Informatics.* Edited by Spiro E, Ahn YY. Springer; 2016.

43. Prata DN, Soares KP, Silva MA, Trevisan DQ, Letouze P: **Social data analysis of Brazilian's mood from twitter**. *Int J Soc Sci Hum* 2016, **6**:179-183.

44. Durahim AO, Coşkun M: **#iamhappybecause: Gross National Happiness through Twitter analysis and big data**. *Tech Forecast Soc Change* 2015, **99**:92-105.

45. Curini L, Iacus S, Canova L: **Measuring idiosyncratic happiness through the analysis of Twitter: an application to the Italian case**. *Soc Ind Res* 2015, **121**:525-542.

46. •• Doré B, Ort L, Braverman O, Ochsner KN: **Sadness shifts to anxiety over time and distance from the national tragedy in Newtown, Connecticut**. *Psychol Sci* 2015, **26**:363-373.
This study illustrates the power of using Big Data to study the impact of collective negative events on subjective well-being by focusing on how this impact changes as a function of time and space.

47. Algan Y, Beasley E, Guyot F, Higa K, Murtin F, Senik C: **Big data measures of well-being: evidence from a Google well-being index in the United States**. *OECD Statistics Working Papers* 2016, **2016**.

48. Robinson WS: **Ecological correlations and the behavior of individuals**. *Am Soc Rev* 1950, **15**:351-357.

49. Lucas RE, Diener E, Suh E: **Discriminant validity of well-being measures**. *J Pers Soc Psychol* 1996, **71**:616-628.

50. Luhmann M, Hawkley LC, Eid M, Cacioppo JT: **Time frames and the distinction between affective and cognitive well-being**. *J Res Pers* 2012, **46**:431-441.

51. Settanni M, Marengo D: **Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts**. *Front Psychol* 2015, **6**:1045.

52. Wang N, Kosinski M, Stillwell DJ, Rust J: **Can well-being be measured using Facebook status updates? Validation of Facebook's gross national happiness index**. *Soc Ind Res* 2014, **115**:483-491.

53. Salguero JM, Cano-Vindel A, Irruarrizaga I, Fernández-Berrocal P, Galea S: **Trajectory and predictors of depression in a 12-month prospective study after the Madrid March 11 terrorist attacks**. *J Psychiatr Res* 2011, **45**:1395-1403.

54. Yu Y, Wang X: **World Cup 2014 in the Twitter World: a big data analysis of sentiments in U.S. Sports fans' tweets**. *Comput Hum Behav* 2015, **48**:392-400.

55. Eid M, Diener E: **Introduction: the need for multimethod measurement in psychology**. In *Handbook of Multimethod Measurement in Psychology.* Edited by Eid M, Diener E. American Psychological Association; 2006:3-8.

56. • Wojcik SP, Hovasapian A, Graham J, Motyl M, Ditto PH: **Conservatives report, but liberals display, greater happiness**. *Science* 2015, **347**:1243-1246.
Multimethod study combining self-report measures and different Big Data measures of subjective well-being to study the correlation between political orientation and subjective well-being.

57. Kramer ADI, Guillory JE, Hancock JT: **Experimental evidence of massive-scale emotional contagion through social networks**. *Proc Natl Acad Sci USA* 2014, **111**:8788-8790.

58. Carpenter J, Crutchley P, Zilca RD, Schwartz HA, Smith LK, Cobb AM, Parks AC: **Seeing the 'big' picture: Big data methods for exploring relationships between usage, language, and outcome in Internet intervention data**. *J Med Internet Res* 2016, **18**:374-391.

59. He S, Zheng X, Zeng D, Luo C, Zhang Z: **Exploring entrainment patterns of human emotion in social media**. *PLoS ONE* 2016, **11**:e0150630.

60. Christofides E, Muise A, Desmarais S: **Hey mom, what's on your Facebook? Comparing Facebook disclosure and privacy in adolescents and adults**. *Soc Psychol Pers Sci* 2012, **3**:48-54.

61. Davenport SW, Bergmann SM, Bergmann JZ, Fearrington ME: **Twitter versus Facebook: exploring the role of narcissism in the motives and usage of different social media platforms**. *CHB* 2014, **32**:212-220.

62. Liu P, Chan D, Qiu L, Kosinski M, Stillwell DJ: **Culture moderates the relationship between emotional expression and subjective well-being**. *Int J Psychol* 2016, **51**:534.