Natural Language Content Mediates the Association Between Active Interactions on Social Network Services and Subjective Well-Being

Kazuma Mori, PhD,1,2 Hugo Hadjur, PhD,1 and Masahiko Haruno, PhD1,3

Abstract

Previous studies indicated that active interactions on social networking services (SNS) are positively linked to subjective well-being (SWB). However, how semantic SNS content affects the association between the degree of SNS interaction and SWB has not been investigated. We addressed this issue by conducting a mediation analysis using natural language processing. We first analyzed Twitter data and SWB scores from 217 participants and found that the degree of active interactions on Twitter (i.e., frequency of reply) was positively correlated with SWB. Next, our multivariate mediation analysis demonstrated that positive words served as SWB-promoting mechanisms for highly interactive people, whereas worrying words led to lower SWB for less interactive people, but negative words did not. This study revealed that natural language content explains why individuals who are highly interactive on SNS have higher SWB, whereas less interactive individuals show lower SWB.

Keywords: social media, subjective well-being, active interaction, natural language, mediation analysis

Introduction

SUBJECTIVE WELL-BEING (SWB) refers to three distinct aspects of life: how people evaluate their lives, frequent positive emotional experiences, and infrequent negative emotional experiences.1 SWB can reflect personal goals individuals make throughout their lives2 and promotes health and longevity.3,4 Because social networking services (SNS) influence broad aspects of people’s lives, researchers are interested in the relationship between SNS use and SWB.5

Early empirical studies that attempted to identify the connection between SNS use and SWB reported inconsistent results. Several studies associated the frequency of SNS use with lower SWB6 by showing that too much SNS use was linked to psychological dysfunction (e.g., depression, loneliness).7,8 Conversely, other studies reported that frequent SNS use was associated with positive SWB.9,10

Recent studies indicated that the link between SNS use and SWB depends on how people use SNS rather than whether they use SNS.11–14 An important factor in this context is the link between active interaction and SWB. Social connection reflects a basic human psychological desire to feel a sense of belonging, and feel close to and connected with others15; social connection through SNS has been correlated with SWB.16,17 Several studies reported that receiving more responses to SNS updates was associated with higher SWB, rather than higher frequency of SNS posting.18–20 This suggested that active interactions on SNS have a key role in promoting SWB through establishing and maintaining social connections.9,21,22
Few studies have examined how actual natural language content on SNS affects the link between active interactions and SWB. Previous studies investigated the emotional content of SNS postings and SWB using questionnaire surveys, but data for SNS usage were noisy as these surveys were based on participants’ retrospective thoughts. To accurately quantify the effects of the content of SNS postings, it is necessary to assess the natural language content of SNS interactions.

This study examined whether natural language content was a key determinant of the link between SNS active interactions and SWB. Because the emotional valence of text may reflect SWB, we hypothesized that positive and negative language expressions had opposite mediating effects on the relationship between active interactions and SWB. We extracted SNS (i.e., Twitter) information and used this to quantify the relationship between active interactions and SWB. We then conducted a mediation analysis to test whether SNS natural language content formed bridges in the pathway from active interactions to SWB.

Methods

Participants

We recruited 217 participants with Twitter accounts (76 females; mean age 22.1 years, standard deviation [SD] = 3.3 years). All Twitter accounts met our activity requirement of more than 100 tweets before data collection. We confirmed that each participant’s last tweet was posted in the same year of data collection. The time periods for tweets collected by API (see the Twitter Behaviors section) varied across participants’ accounts (M = 878 days, SD = 539 days, range = 18–2,766 days). This study was approved by the NICT Institutional Ethics and Safety Committee, and we obtained written informed consent from all participants before the experiments.

Subjective well-being

SWB was measured using a subjective happiness scale that was suitable for a range of age groups and had good validity and reliability (Table 1). Participants also completed other psychological scales related to SWB (Fig. 1; Supplementary Table S1). An in-house experiment system based on the Lime Survey program (LimeSurvey GmbH, Germany) was used to present and answer the scales on web browsers.

**Table 1. The Four Subjective Happiness Scale Items**

| Item                                                                 | Response                                      |
|----------------------------------------------------------------------|-----------------------------------------------|
| In general, I consider myself                                         | 1: not a very happy person ~ 7: a very happy person |
| Compared with most of my peers, I consider myself                    | 1: less happy ~ 7: more happy                  |
| Some people are generally very happy. They enjoy life regardless of what is going on, getting the most out of everything. | 1: not at all ~ 7: a great deal |
| To what extent does this characterization describe you?              |                                               |
| Some people are generally not very happy. Although they are not depressed, they never seem as happy as they might be. |                                               |
| To what extent does this characterization describe you?              |                                               |

We treated the total value of the subjective happiness scale as the subjective well-being index.

Twitter behaviors

We used the official Twitter streaming API through the python-Twitter library to collect tweets. We used the API interface to retrieve Twitter account information and past tweets posted by all 217 participants (up to a maximum of the 3,200 most recent tweets at the time of data collection). We calculated the frequency of SNS use for each Twitter account as the number of all tweets divided by the account use period (i.e., days), which varied across accounts. The SNS frequency was log-transformed because its distribution was positively skewed.

We also extracted the reply behavior from the most recent 3,200 tweets (maximum number of API). As “reply” describes exchanging messages with friends and others on Twitter, we concluded that a reply was an active interaction similar to those used in previous SWB studies focused on texting and Facebook comments. The number of tweets extracted by API varied among participants (M = 2,217; SD = 1,132). We therefore defined the active interaction score as the frequency of replies divided by the total number of tweets (i.e., the proportion of tweets that were replies).

Natural language content

We applied natural language processing methods to tweet sentences. We used MeCab (morphological analyzer segmenting Japanese sentences into a sequence of morphemes) with MeCab-ipadic-NEologd (Japanese dictionary that includes new words). The Affective Norms for English Words (ANEW) was used to analyze affective words. After the translation procedure, we obtained 1,331 positive Japanese words and 1,643 negative Japanese words as defined in a previous study (scored 0–10). For each participant, we computed positive and negative emotional expression scores using the sum of the weighted affective scores for all ANEW words used in the tweets.

Next, we modeled each user’s Twitter data as a document and the clusters of words as topics based on the R package topicmodels. Latent Dirichlet allocation (LDA) is a cluster analysis method used to compute semantic dimensions in language. This open-vocabulary analysis method enables examination of the associations between language content and SWB in a data-driven manner and is not limited to a priori defined affective words. We modeled 20 topics from the 510,854 Twitter sentences drawn from the 217 participants’ Twitter accounts using Gibbs sampling.

Some people are generally not very happy. Although they are not...
FIG. 1. Twitter behaviors and SWB. (a) Scatter plots relating Twitter behaviors to SWB. (b) Results of a correlation analysis between the two types of Twitter behaviors and 11 trait scores. Grayscale represents correlation coefficients. Significant positive and negative results with the FDR correction are shown in bright and dark gray. BAS, behavioral approach system; BAS-D, BAS-drive; BAS-R, BAS-reward; BAS-S, BAS-sensation seeking; BDI, Beck Depression Inventory; BIS, behavioral inhibition system; FDR, false discovery rate; PSS, perceived subjective stress; RSS, Rosenberg Self-Esteem Scale; SDS, Self-Rating Depression Scale; STAI, State Trait Anxiety Inventory; STAI-S, STAI-state; STAI-T, STAI-trait; SWB, subjective well-being.

(Supplementary Fig. S2). As LDA allows the setting of multiple topics for a document, it is possible that a word in a document shows high probability for several topics. 17

Statistical analysis

Using z-transformed scores, we conducted a partial correlation analysis between Twitter behaviors (active interaction, SNS frequency) and SWB (and related trait scores) by setting demographic characteristics (age and gender) as control variables with false discovery rate multiple comparisons. Similarly, we conducted a partial correlation analysis between SWB scores and the natural language indices (Supplementary Fig. S1).

We performed a multiple mediation analysis 39,40 using the lavaan package in R 41 to reveal potential internal mechanisms underlying the relationship between active interactions and SWB. Multiple mediators were included in the model to estimate the indirect effect of a certain linguistic mediator. Age, gender, and SNS frequency were set as control variables. Significance tests were conducted based on nonparametric percentile bootstrap resamples repeated 5,000 times. The direct and indirect effects were statistically significant if 95% bootstrap confidence intervals did not contain zero.

Results

Twitter behaviors and SWB

Participants’ SWB score was positively correlated with the active interaction score ($r=0.32$, $p<0.001$) but not with the frequency of SNS ($r=-0.09$, $p=0.168$) (Fig. 1a). Only active interaction scores consistently showed positive correlations with positive traits and negative correlations with negative traits ($p$s $<0.05$) (Fig. 1b).

Natural language content and SWB

We found a significant correlation between positive words and the SWB score ($r=0.29$, $p<0.001$), and negative correlations between negative words and the SWB score ($r=-0.31$, $p<0.001$) (Fig. 2a, b). A similar relationship was found for emotional words and SWB-related traits ($p$s $<0.05$) (Fig. 2c).

Figure 3a shows the words that were most strongly positively correlated (topic 2, left) and negatively correlated (topic 3, right) with SWB. These two topic scores showed a negative correlation ($r=-45$, $p<0.001$). We observed that the language markers of high-SWB topics included joy-related words (e.g., many “lol”). The language markers of low-SWB topics were suggestive of worrying-related words (e.g., “may,” “feel some,” and “end up”). We found significant correlations between SWB score and the high-SWB topic score ($r=0.28$, $p<0.001$) and the low-SWB topic score ($r=-0.32$, $p<0.001$) (Fig. 3b). A similar tendency was found for these topics and other traits ($p$s $<0.05$) (Fig. 3c).

Multiple mediation analysis

We tested whether the natural language indices mediated the relationship between active interaction and SWB (Fig. 4). Standardized regression coefficients showed that active interactions were positively associated with positive words and the high-SWB topic and negatively associated with
negative words and the low-SWB topic. In turn, positive words were positively associated with SWB, and the low-SWB topic was negatively associated with SWB.

The direct effect of active interactions on SWB was not significant, although the total effect was significant (Table 2), suggesting that the relationship between active interactions and SWB may be explained by natural language. The indirect effects of active interactions on SWB through the low-SWB topic were not significant, whereas positive words had a significant positive effect on the association between active interactions and SWB. Therefore, the association between active interactions and SWB was mediated and explained by individual differences in the use of positive words and the low-SWB topic.

Discussion

This study explored internal mechanisms for the relationship between active interactions on SNS and SWB. We found a positive correlation between SNS active interaction (i.e., reply) and SWB, and revealed that positive words and low-SWB topic words mediated the association between SNS active interactions and SWB through a multiple mediation analysis based on natural language processing.

We showed that active interactions on SNS were associated with SWB, whereas the frequency of SNS use was not. A similar relationship between active interactions and SWB was previously reported in both offline and online contexts. We speculated that SNS interactions with others (mostly friends in our student sample) may promote SWB at least partly through the reinforcement of offline social connections, as online interactions may complement existing personal connections.

The significant mediation effect of positive words indicated that the higher SWB in an individual with a wide active SNS network was explained by the higher frequency of positive-word use, which was consistent with previous literature. Although positive words in interactions do not necessarily share a positive event, the sharing of positive
FIG. 3. Correlation between the topic score and SWB. (a) The two language topics most strongly positively and negatively associated with SWB. Font size reflects the relative prevalence of words within topics. Grayscale is for readability. (b) Scatter plot displaying the correlations between the two topic scores and SWB. (c) Results of the correlation analysis between the topic scores and 11 trait scores. Grayscale represents correlation coefficients. Significant positive and negative results with the FDR correction are shown in bright and dark gray.

FIG. 4. Model examining the multiple mediational effects of natural language content on the association between active interactions and SWB. Standardized regression coefficients (β) are shown above each arrow. **p < 0.01, ***p < 0.001.

| Variables                  | Direct effect | Indirect effects | Lower limit | Upper limit |
|----------------------------|---------------|------------------|-------------|-------------|
| Active interaction         | 0.103         |                  | -0.014      | 0.237       |
| Positive words             |               |                  | 0.030       | 0.174       |
| Negative words             | -0.004        | -0.042           | 0.037       | 0.081       |
| High-SWB topic             | 0.021         | -0.033           | 0.081       | 0.169       |
| Low-SWB topic              | 0.087**       | 0.030            | 0.169       |             |

Total effect = 0.296, CI [0.166–0.430], p < 0.001; **p < 0.01. CI, confidence interval; SWB, subjective well-being.
events through active communication may lead to greater positive effects, because active interaction with positive words may satisfy a basic human psychological demand to enhance the sensation of social connections.

Another key finding of our study was the mediation effect of the low-SWB topic, which appeared to be a cluster of worrying words. This suggested that increasing worrying word use reduced both SNS interactions and SWB. Less interactive SNS use may displace time spent on other activities that are beneficial to mental health (e.g., real-life active interactions) and possibly increase the user’s loneliness. Our finding that the predetermined negative words did not mediate the association between lower SNS interactions and lower SWB was surprising. However, emotional intensity decreases rapidly after explicit expressions of negative emotion in postings, SNS posts with negative words may attenuate negative emotions in the short term. Conversely, when people use worrying words, they may be unable to attenuate their emotions because they are describing concerns and may end up with the lower SWB associated with being less interactive on SNS.

We analyzed data for tweets that were posted over several years. This strategy could capture stable SWB but may not be optimal for dynamic changes in SWB. It is important that future studies consider temporal variations in the link between SNS behavior and SWB in more detail.

In summary, this study found that positive words and worry-related words serve as key factors through which SNS active interactions impact SWB, which clarifies how linguistic information contributes to the associations between SNS active interactions and SWB. We believe that the knowledge accumulated in this study will be useful for the further development of theories on the use of digital technologies and SWB.

Acknowledgments

We thank S. Tada and R. Iwai for their technical assistance and data collection. We also thank Peter Karagiannis and Edanz for editing the drafts of the article.

Authors’ Contributions

K.M., H.H., and M.H. conceived the design of this study. K.M. conducted the experiments and performed the analyses. K.M. and M.H. interpreted the results and wrote the article.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

Preparation of this manuscript was supported by Core Research for Evolutional Science and Technology (CREST), Grant Number JPMICR15E4, Japan Science and Technology Agency (JST), Moonshot R&D Grant Number JPMJMS2011, JST and KAKENHI (17H06314, 22H05155 and 20K21562) to M.H.

Supplementary Material

Supplementary Figure S1
Supplementary Figure S2
Supplementary Table S1

References

1. Diener E. Subjective well-being. Psychol Bull 1984;95(3): 542–575; doi: 10.1037/0033-2909.95.3.542.
2. Tay L, Kuykendall L, Diener E. Satisfaction and Happiness—The Bright Side of Quality of Life. In: Global Handbook of Quality of Life: Exploration of Well-Being of Nations and Continents. (Glazer W, Camfield L, Moller V, et al. eds) Netherlands: Springer; 2015; pp. 839–853.
3. Boehm JK, Peterson C, Kivimaki M, et al. A prospective study of positive psychological well-being and coronary heart disease. Heal Psychol 2011;30(3):259–267; doi: 10.1037/a0023124.
4. Diener E, Chan MY. Happy people live longer: Subjective well-being contributes to health and longevity. Appl Psychol Heal Well-Being 2011;3(1):1–43; doi: 10.1111/j.1758-0854.2010.01045.x.
5. Jaidka K, Giorgi S, Schwartz HA, et al. Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. 2020; 117(19):10165–10171.
6. Song H, Zmyslinski-Seelig A, Kim J, et al. Does Facebook make you lonely? A meta analysis. Comput Human Behav 2014;36:446–452; doi: 10.1016/j.chb.2014.04.011.
7. Liu yi Lin, Sidani JE, Shensa A, et al. Association between social media use and depression among U.S. young adults. Depress Anxiety 2016;33:323–331; doi: 10.1016/j.physbeh.2017.03.040.
8. Sampaşa-Kanyinga H, Lewis RF. Frequent use of social networking sites is associated with poor psychological functioning among children and adolescents. Cyberpsychol Behav Soc Netw 2015;18(7):380–385; doi: 10.1089/cyber.2015.0055.
9. Burke M, Marlow C, Lento T. Social network activity and social well-being. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2010; pp. 1909–1912.
10. Deters FG, Mehl MR. Does posting Facebook status updates increase or decrease loneliness? An online social networking experiment. Soc Psychol Personal Sci 2013; 4(5):579–586; doi: 10.1177/1948550612469233.
11. Verduyn P, Ybarra O, Réisbois M, et al. Do social networking sites enhance or undermine subjective well-being? A critical review. Soc Issues Policy Rev 2017;11(1):274–302; doi: 10.1111/sipr.12033.
12. Verduyn P, Lee DS, Park J, et al. Passive Facebook usage undermines affective well-being: Experimental and longitudinal evidence. J Exp Psychol Gen 2015;144(2):480–488; doi: 10.1037/0033-2909.144.2.480.
13. Bailey ER, Matz SC, Youyou W, et al. eds) Netherlands: Springer; 2015; pp. 839–853.
14. Kross E, Verduyn P, Sheppes G, et al. Social media and well-being: Pitfalls, progress, and next steps. Trends Cogn Sci 2021;25(1):55–66; doi: 10.1016/j.tics.2020.01.005.
15. Deci EL, Ryan RM. The “What” and “Why” of goal pursuits: Human needs and the self-determination of behavior. Psychol Inq 2000;11(4):227–268; doi: 10.1207/S15327965PLI1104_01.
16. Ellison NB, Steinfield C, Lampe C. The benefits of Facebook “Friends”: Social capital and college students’ use of online social network sites. J Comput Commun
2007;12(4):1143–1168; doi: 10.1111/j.1083-6101.2007 .00367.x.
17. Kim C, Shen C. Connecting activities on social network sites and life satisfaction: A comparison of older and younger users. Comput Human Behav 2020;105:106222; doi: 10.1016/j.chb.2019.106222.
18. Burke M, Kraut RE. The relationship between Facebook use and well-being depends on communication type and tie strength. J Comput Commun 2016;21(4):265–281; doi: 10.1111/jcc4.12162.
19. Zell AL, Moeller L. Are You Happy for Me … on Facebook? The potential importance of “likes” and comments. Comput Human Behav 2018;78:26–33; doi: 10.1016/j.chb.2017.08.050.
20. Wenninger H, Krasnova H, Buxmann P. Activity matters: Investigating the influence of Facebook on Life Satisfaction of Teenage Users. ECIS 2014 Proceedings of the 22nd European Conference on Information Systems 2014, pp. 1–18.
21. Burke M, Kraut R, Marlow C. (2011) Social capital on Facebook: Differentiating uses and users. In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems. New York, NY: ACM, pp. 571–580.
22. Ahn D, Shin DH. Is the social use of media for seeking connectedness or for avoiding social isolation? Mechanisms underlying media use and subjective well-being. Comput Human Behav 2013;29(6):2453–2462; doi: 10.1016/j.chb.2012.12.022.
23. Choi M, Toma CL. Social sharing through interpersonal media: Patterns and effects on emotional well-being. Comput Human Behav 2014;36:530–541; doi: 10.1016/j.chb.2014.04.026.
24. Scharkow M. The accuracy of self-reported Internet use—A validation study using client log data. Commun Methods Meas 2016;10(1):13–27; doi: 10.1080/19312458.2015 .1118446.
25. Ernala SK, Burke M, Leavitt A, et al. How well do people report time spent on Facebook? An evaluation of established survey questions with recommendations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, 2020, pp. 1–4.
26. Hills TT, Proto E, Sgroi D, et al. Historical analysis of national subjective wellbeing using millions of digitized books. Nat Hum Behav 2019;3:1271–1275; doi: 10.1038/s41562-019-0750-z.
27. Lyubomirsky S, Lepper HS. A measure of subjective happiness: Preliminary reliability and construct validation. Soc Indic Res 1999;46(2):137–155; doi: 10.1023/A:1006824100041.
28. Shimai S, Otake K, Utsuki N, et al. Development of a Japanese Version of the Subjective Happiness Scale (SHS), and Examination of its validity and reliability. Jpn J Public Health 2004;10:845–853.
29. Kudo T, Yamamoto K, Matsumoto Y. Applying conditional random fields to Japanese morphological analysis. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing—EMNLP’04 Barcelona, Spain; 2004; pp. 230–237.
30. Satou T. Mecab-Ipadic-Neologd; 2015. Available from: https://github.com/neologd/mecab-ipadic-neologd [Last accessed: August 22, 2022].
48. Pang D, Eichstaedt JC, Buffone A, et al. The language of character strengths: predicting morally valued traits on social media. J Pers 2020;88(2):287–306; doi: 10.1111/jopy.12491.

49. Twenge JM. More time on technology, less happiness? Associations between digital-media use and psychological well-being. Curr Dir Psychol Sci 2019;28(4):372–379; doi: 10.1177/0963721419838244.

50. Fan R, Varol O, Varamesh A, et al. The minute-scale dynamics of online emotions reveal the effects of affect labeling. Nat Hum Behav 2019;3:92–100; doi: 10.1038/s41562-018-0490-5.

51. Schemer C, Masur PK, Geiß S, et al. The impact of Internet and social media use on well-being: A longitudinal analysis of adolescents across nine years. J Comput Commun 2021; 26(1):1–21; doi: 10.1093/jcme/zaaa014.

52. Kross E, Verduyn P, Boyer M, et al. Does counting emotion words on online social networks provide a window into people’s subjective experience of emotion? A case study on Facebook. Emotion 2019;19(1):97–107; doi: 10.1037/emo0000416.

Address correspondence to:
Dr. Kazuma Mori
Center for Information and Neural Networks (CiNet)
National Institute of Information and Communications Technology (NICT)
1-4 Yamadaoka
Suita
Osaka 565-0871
Japan
E-mail: kazumamori@nict.go.jp

Dr. Masahiko Haruno
Center for Information and Neural Networks (CiNet)
National Institute of Information and Communications Technology (NICT)
1-4 Yamadaoka
Suita
Osaka 565-0871
Japan
E-mail: mharuno@nict.go.jp