



BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Detecting Concurrent Personal Mood in Contact Networks: An Improved Diary Approach

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-020600
Article Type:	Research
Date Submitted by the Author:	13-Nov-2017
Complete List of Authors:	Chan, Ta-Chien; Academia Sinica, Research Center for Humanities and Social Sciences Yen, Tso-Jung; Academia Sinica, Institute of Statistical Science Hu, Tsuey-Hwa; Academia Sinica, Institute of Statistical Science Fu, Yang-chih; Academia Sinica, Institute of Sociology Hwang, Jing-Shiang; Academia Sinica, Institute of Statistical Science
Keywords:	contact diary, emotional contagion, social networks, transmission of mood

SCHOLARONE™
Manuscripts

Detecting Concurrent Personal Mood in Contact Networks: An Improved Diary Approach

Ta-Chien Chan^a, Tso-Jung Yen^b, Tsuey-Hwa Hu^b, Yang-chih Fu^{c,1*}, Jing-Shiang Hwang^{b,2*}

^aResearch Center for Humanities and Social Sciences, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei, 115 Taiwan

^bInstitute of Statistical Science, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei, 115 Taiwan

^cInstitute of Sociology, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei, 115 Taiwan

Major classification: Social Sciences; Minor classification: Social Sciences

Short title: mood spread through networks

Keywords: contact diary, emotional contagion, social networks, transmission of mood

*Co-corresponding Authors:

Yang-chih Fu¹:

Tel: 886-2-2652-5149; E-mail: fuyc@sinica.edu.tw

Jing-Shiang Hwang²:

Tel: 886-2-2783-5611; E-mail: hwang@sinica.edu.tw

Abstract

Objectives: This paper examines how people express personal mood concurrently with those connected with them by one or two degrees of separation.

Design: Participatory cohort study.

Setting: Online contact diary.

Participants: 133 participants kept online contact diaries for seven months in 2014, which included 127,455 contacts with 12,070 persons.

Main outcome measures: Diary keepers rated a contacted person's mood during each specific contact, as well as the strength of ties between any pairs of such contacted persons. Such rich information about ties and contacts enable us to construct a complete contact network for each diary keeper, along with the network members' mood and tie strength. We calculate one's overall mood by that person's average mood score during the study period and take the shortest path between any given pair of contacted persons as the degree of separation. We further assume that two connecting persons in a complete contact network have made contact with each other during the study period, which allows us to examine whether and how personal moods occur concurrently within these contact networks.

Results: Using mixed-effects models while controlling for covariates at both tie and contact levels, we show that personal mood score positively and significantly correlates with the average mood among those directly tied to the person. The same effect remains positive and significant, though the effect size is reduced by about one half, for those connected to the person by two degrees. The mood of anyone separated by more than two degrees is statistically irrelevant.

Conclusions: Applying core social network perspectives and rich data at both tie and contact levels to inquiries about subjective well-being, the current study sheds new light on how an improved diary approach can help explain the ways in which

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

individuals express their personal moods concurrently during social interactions in
everyday life.

For peer review only

Strengths and limitations of this study

1. The observed online contact diaries allow us to construct 133 complete contact networks with more than 12,000 members, which help uncover how personal moods vary among network members.
2. Some intertwining contact diaries further enable us to cross-examine personal mood and tie strength as rated by both parties of social interactions, solidifying our findings about how a bottom-up social network approach helps reveal concurrent personal mood in everyday life.
3. Without information about the exact timing of each contact among network members, we cannot infer the occurrence of contagion or diffusion of personal mood.
4. With limited information about the contacted persons' personal background, our model cannot fully adjust for the effects of homophily.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

Experiments have demonstrated that emotions tend to spread to others during social interactions,¹ as in the case when diseases, behaviors, and ideas transmit through social networking.^{2 3} Analyses from other large-scale data have further revealed similar spreading patterns of both positive and negative emotions, such as happiness and depression.^{4 5} Like the diffusion of behaviors and attitudes amid face-to-face social networking, emotive sharing and contagion may also cover large social circles through online social contacts by text or voice on Facebook, Twitter, Skype, Google talk, and other social media.^{6 7}

Most existing studies have demonstrated that certain emotions tend to spread from direct contacts over a short period of time, but relatively few have examined whether personal mood may also spread from, or at least occur concurrently among indirect contacts within social networks over a longer period of time. Two recent studies, for example, showed that our mood can be influenced not only by those familiar to us, but also by friends’ friends whom we do not personally know.^{2 3} More specifically, the spread of personal happiness reaches up to three degrees of separation along social networks, according to one such rare empirical study, the Framingham Heart Study, which analyzed historical data over 20 years.⁴ In that particular study, the network data consisted of 5,124 egocentric networks with an average of 10.4 “alters,” or network members. The researchers measured the emotions of these individuals by retrieving data from three waves of physical examinations in 1986, 1996, and 2000. Based on the rich longitudinal data, they modeled personal happiness as a function of one’s happiness in the previous wave, and of friends’ and relatives’ happiness in both current and previous waves. All of the network members were classified by the degrees of separation from the focal person, in four categories. The longitudinal analyses indicated that an individual was 15.3% more likely to be happy if a directly

connected network member (with one degree of separation) was happy; the effect decreased to 9.8% for those separated by two degrees and 5.6% for those by three degrees.⁴

As such findings may have inspired studies of mood diffusion, most non-experimental studies are insufficient to examine the actual mechanism of “diffusion,” because the observational data on which they are based offer no such advantages. Thus, it would be particularly difficult to determine and infer from observational data how personal moods spread or diffuse.⁸ Even without claiming the causal effects that are essential for explaining spread or diffusion, however, it would be revealing to examine whether and how similar patterns of concurrent mood may exist under other circumstances. While previous longitudinal studies have focused on how individuals’ happiness or depression changes across waves of surveys, an alternative “bottom-up” approach tends to uncover subtle patterns of concurrent mood by examining how one’s mood fluctuates at the micro level, contact by contact.⁹⁻¹¹ Due to the methodological limitation, it may be inappropriate to pursue inquiries about “mood diffusion or contagion” solely by using observational data, even if one could replicate the decades-long study by creating another huge number of personal networks with records of participants’ relationships and emotions. As an alternative approach, our study, based on a different format, should provide a new perspective that would further enrich studies on how personal moods can be linked to one another in contact networks.

By extending the bottom-up approach to social network studies, we aim to examine whether and how personal moods occur concurrently within a contact network, using data collected via an online platform, ClickDiary, over a seven-month period between May 1 and November 30, 2014. The ClickDiary program uses a web-based platform to collect data on participants’ health behaviors and all

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

one-on-one interpersonal contacts in everyday life.¹² The data retrieved for this study have a nested hierarchical structure, including detailed information about 133 participants (or “diary keepers,” who recorded details about their social interactions in daily life), 12,070 contacted persons (or “network members,” including 74 persons who also participated in the same diary keeping platform), as well as 127,455 contacts.

Two features in the ClickDiary platform are central to our research design. First, diary keepers reported their own personal mood for each one-on-one contact by selecting one of the following score categories that best matched their estimates: (1) poor, (2) good, (3) very good, or (4) excellent; they then evaluated the mood of each network member during the contact. Second, diary keepers, to their best knowledge, rated how well a given pair of network members knew each other. Once a diary keeper confirmed all of their interpersonal ties, we used these two critical features as the backbone to construct a “complete contact network” surrounding that diary keeper, the focal person. Within such a complete contact network, we linked the nodes that represent the network members to each other by interpersonal ties. The overall mood of each network member can be represented by taking the average of the person’s mood scores recorded in the contact diary during the whole study period. Furthermore, we assume that two connecting persons in a complete contact network also have made contact with each other during the study period. This assumption allows us to evaluate whether and how the network members’ personal moods correlate with one another. We applied a mixed-effects model to analyze overall mood scores of the members in the 133 complete contact networks to examine whether and how an individual’s mood is associated with those within one or two degrees of separation, as well as other members in the complete contact networks.

Materials and Methods

The ClickDiary Program

The ClickDiary program (<http://cdiary.tw>) uses a web-based platform, written in Chinese, to collect data on participants' daily health behaviors and interpersonal one-on-one contacts.¹² One unique feature of ClickDiary is the friendly interface designed for clicking options on structured diary items via a website or mobile apps, making it easier to record responses whenever it is convenient for participants. Upon signing up for the program, participants provided socio-demographic information, including age, gender, place of residence, marital status, and current job. The program also collects participants' Big Five personality traits (openness, conscientiousness, extraversion, agreeableness, neuroticism),¹³ height and weight, perceived health status and happiness, the number (and characteristics) of people contacted during the day, along with a baseline health module that borrowed items from the International Social Survey Programme.¹⁴ This current study has been approved by the Institutional Review Board on Humanities and Social Science Research (IRB-HS), Academia Sinica (AS-IRB-HS 02-13022).

Contact diary

When adding a person to the contact list for the first time, the diary keepers provided the person's background information, such as age and gender, and evaluated several aspects of their relationship with the person, including the duration of acquaintanceship, degree of familiarity, the most frequent mode used for contact (face-to-face, voice only, or text only), contact frequency, and the likelihood of discussing important matters. In this particular diary program, a "contact" refers to one-on-one exchange that involves at least three verbal or written sentences, a definition somewhat narrower than most previous studies using the contact diary

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

approach.¹⁵

The program also asked participants to evaluate the degree to which any pair of persons on the contact list was familiar with each other. Before starting to enter the contact details with any new person, a diary keeper had the chance to estimate, on a scale from 1 to 3, how well this particular person knew each of randomly selected five other persons already recorded in the diary. After this first step, another random sample of five different persons' names popped up, so that the diary keeper could continue judging all the ties between any pair of persons within the contact network. The process continued, randomly adding five new names to the list at a time, until the diary keeper finished rating the strength of all alter pairs. The design allowed the diary keeper to evaluate all alter-alter ties when it was convenient to do so, thus achieving maximum flexibility and encouraging a higher completion rate. In addition, the diary entries also focused on 11 contact attributes, including when the contact took place, who initiated the contact, the major mode of the contact, the duration and content of the contact, where the diary keeper and the contacted person were during the contact, the extent to which the contact felt beneficial to the diary keeper, the mood of each party during the contact, and whether the contacted person showed any cold symptoms.

Data retrieved for the study

From May 1 to November 30, 2014, 133 residents in Taiwan each completed at least 30 days of contact diaries and recorded one-on-one interpersonal contacts with at least 30 persons, along with other diary keepers who kept records on contacts for fewer than 30 days or with fewer than 30 persons. We retrieved the seven-month contact diaries from these 133 participants, with all personal identities removed, which consisted of 141,909 contacts with 16,139 contacted persons. The length of the

contact lists varied substantially among the 133 participants, ranging from 30 to 1,399, with a median of 76. The participants recorded an average of 12 contacts a day, with a minimum of 1 and a maximum of 56. The contacted persons varied greatly in terms of how often they appeared in the diaries, ranging from only once to daily (214 times during the seven-month period), even though the average frequency was 8 times.

To measure each person's mood, diary keepers selected one from the following score categories that best matched their estimates: (1) poor, (2) good, (3) very good, or (4) excellent. As expected, diary keepers sometimes were unable to estimate a contacted person's mood and consequently answered "Don't Know" for the item. The records during the study period showed such "Don't Know" answers for the mood item appearing in 9,042 contacts. We treated these contacts (with 845 contacted persons) as missing, which reduced the valid number of contacted persons to 15,294.

We further excluded the cases where the diary keepers did not know a contacted person (i.e., the category that accounts for about 21.1% of the relationships) from the subsequent analysis for three reasons. First, interactions with strangers carry very different implications in studies of network diffusion. Even though they had actual contacts with the diary keepers, keeping total strangers in a personal network would have left too many unknown or uncertain links, because the diary keepers were unlikely to judge how well these strangers knew one another. Second, it would have been more difficult for the diary keepers to evaluate strangers' personal mood, which also tended to yield missing or less reliable mood rating. Third, probably due to such uncertainties, our different modeling efforts while retaining the ties with these strangers resulted in unreasonable noise to the analysis. The final data for modeling mood spread included 12,070 contacted persons of 133 diary keepers who had made 127,455 contacts during the study period.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Network construction

As a whole, diary keepers were able to confirm nearly all (99.97%) interpersonal ties in terms of the familiarity between any pair of contacted persons. Of these ties, averaged across the 133 diary keepers, 78.4% were considered absent (the pairs did not know each other), 10.6% were strong (knew each other well), and 11.1% were weak (knew each other, though not well). For each diary keeper, we constructed a complete contact network, in which each alter, or contacted person, represents a node, and two nodes (other than the focal person) are deemed to be connected when the reported alter-alter tie is either strong or weak, rather than absent. In contrast, nodes are not linked when the pair of network members are strangers to each other. Using the shortest path between any given pair of nodes in the network, we calculate “distance” (or “the degree of separation”) between the network members pair by pair. We focus on the links among these network members only, while excluding any links leading to and from the unique node that represented the diary keeper (or “ego-alter ties”). Excluding such ties linked by the diary keepers helps simplify the calculation of “distance.” Otherwise, any two given nodes in the network would have at most two degrees of separation, because all the network members would be directly linked to the diary keepers. Finally, we defined a network member’s overall mood by the average of the mood scores reported by diary keepers during the study period, ranging from 1 to 4.

To illustrate how alter-alter ties cluster into a subset of ego’s complete contact network, we drew a figure, using R package igraph, to display the network patterns in each of four diary keepers’ complete contact networks during the study period (Figure 1). The figure shows that the clustering patterns may differ significantly among diary keepers. Within each of the four complete contact networks, those members who are closer to each other in terms of distance also tend to average similar scores in personal

mood during all interpersonal contacts.

Statistical Analysis

To examine whether and how a network member's mood may be associated with those separated by one or two degrees in each of the 133 complete contact networks, we first calculated the average mood of those members who were directly connected to a particular member (that is, those connected by one degree of separation), and then we obtained the average mood of those at two degrees of separation and the average mood of all other members. The average mood score of the j^{th} member in the i^{th} complete contact network can be simply obtained by calculating $Y_{ij} = \sum_{k=1}^{K_{ij}} O_{ijk} / K_{ij}$, where O_{ijk} is the mood score of their k^{th} contact given by the i^{th} diary keeper and K_{ij} is the number of contacts during the study period. Let D_{1ij} and D_{2ij} indicate the mean mood scores of those separated by one and two degrees from the j^{th} member in the i^{th} network, respectively. Finally, D_{3ij} measures the average mood scores for those beyond two degrees of separation from the j^{th} network member.

For this study, we applied a mixed-effects model to analyze the relationship between a person's mood score and the average mood scores of those surrounding the person in the network (or "network neighbors"), while controlling for the effects of potential covariates on the person's mood. The model is given as

$$Y_{ij} = (\alpha_0 + a_{0i}) + (\alpha_1 + a_{1i})D_{1ij} + (\alpha_2 + a_{2i})D_{2ij} + (\alpha_3 + a_{3i})D_{3ij} + \sum_{h=1}^p \beta_h X_{hij} + \varepsilon_{ij} \quad \text{for } i = 1, 2, \dots, n \text{ and } j = 1, 2, \dots, n_i,$$

where n is the number of networks; n_i is the size of the i^{th} network; random components a_{li} are assumed to be normally distributed with a mean of 0 and

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

variance σ_l^2 for $l = 0, 1, 2, 3$; and the error term is normally distributed with a mean of 0 and variance τ^2 . Our main interest is to estimate the fixed effects of α_1 , α_2 , and α_3 , which measure whether and how one's mood correlates with those separated by one degree, two degrees, and the others, respectively, within the same complete contact network.

We applied exploratory data analysis tools to identify $p = 15$ potential covariates. In addition to controlling for the possible gender effect, the other covariates included in the mixed-effects model measure a variety of features at both tie and contact levels. Two groups of covariates tap the relationships between the diary keepers and their network members. The first is a binary variable indicating that diary keepers knew a network member either well or “not well” (the latter serves as the base category in the model). The second group of covariates distinguishes four types of the network members' relationships to the diary keepers: family members or relatives (13.2%); good friends (13.5%); coworkers or trade partners (12.7%); and schoolmates, teachers, or students (18.4%). The base category for the group is “others.”

The next group of covariates covers three major contact features: mode, purpose, and duration. Because the main goal of the analysis is to examine how personal moods vary among network members, we sum up these contact features for each member. The modes of contact, for example, may play a key role in determining how well the diary keepers judged others' mood. To verify such an effect, we include two contact modes as covariates: face-to-face and “voice only,” leaving “text only” out of the model. From all of the contacts between the diary keepers and each particular network member during the seven-month study period, we calculate the proportion (percentage) of each of the three modes. Because the percentages of all three modes

for each network member add up to 1, we keep only the first two modes (face-to-face and voice) in the analysis. The averaged percentages of face-to-face and voice only contacts among the network members were about 79.8% and 8.2%, respectively.

The model also includes the percentage for each of the two kinds of contact purposes, “work- or school-related” and “daily routine,” which were about 30.4% and 12.4% on average, respectively, while excluding “other purposes.” We further add contact duration into the model, using the percentages for contacts that last 5-59 minutes, 1-4 hours, or more than 4 hours, excluding those lasting under 5 minutes. On average, 47% of the contact durations was 5-59 minutes, 20% was 1-4 hours, and 6% was more than 4 hours.

Because diary keepers reported their own mood as well as the network members' mood in the same diaries, the two scores are expected to be highly correlated. To take such an effect into account, we further controlled for the most influential covariate that measures the average mood score of the i^{th} diary keeper when contacted with the j^{th} network member during the study period. We used the lme function from the R package “nlme” to estimate the model parameters.¹⁶

Results

Among the original 133 diary keepers, about 80% were female (106/133). The group tended to be younger and better educated. Percentages of those under age 23 (college students), 23-39, 40-59 and 60 or more were 33.8%, 36.1%, 21.8% and 8.3%, respectively. At least 82.7% had ever gone to college. Like previous contact diary studies, the study participants were overrepresented by females and better educated subpopulations.¹¹

The high percentage of female diary keepers probably yielded more females on the contact lists (65.5%), although the latter was actually more balanced than the

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

gender distribution among the diary keepers. The age distribution of the 12,070 contacted persons was not very skewed, with age groups of 1-19, 20-29, 30-49, 50-59, and 60 or more accounting for about 8.4%, 21.8%, 31.1%, 18.5%, and 16.9%, respectively. About 53.3% of the contacted persons the diary keepers knew were known very well.

In fitting the mixed-effects model, we excluded about 12% of network members with any covariates missing. As a result, the following results were obtained from the remaining 10,581 network members with complete covariates. Table 1 shows model estimates of the parameters associated with individual mood variation. On average, a person's mood score increased about 0.13 ($p < .0001$) for every additional point scored in the average mood of fellow members who were directly connected to the person (with one degree of separation). The average mood score of those members linked by two steps also contributed significantly ($p = .002$), with an effect size of about 0.06, to the person. Such contribution diminishes, however, for those members at degree three and beyond (linked by at least three steps).

These key findings are noteworthy, because they have been adjusted by several highly relevant covariates in the mixed-effects model. Since the diary keepers rated the mood of both parties after a contact, the two scores were often highly correlated. As shown in Table 1, a member's mood score was strongly determined by 0.74 of the average mood score of the diary keeper. When the diary keeper knew a member well, that member had a better chance of receiving a higher score on personal mood. A network member also tended to receive a higher mood score for a face-to-face contact and a contact that lasted longer. When a contact was about work or school, or was part of the daily routine, the network member's mood was not as good as that of other contacts, at least based on the diary keeper's judgment.

Robustness Checks

To verify the effects of different degrees of separation, we tried a separate analysis of those members at three degrees of separation, using only 8,505 network members with all valid variates. The alternative analytic design showed little change in both the coefficient estimates of the covariates and the coefficient estimates of one, two, and three degrees of separation (0.13, 0.08, and 0.03, respectively). Thus, it would be unnecessary and unfruitful to further determine whether concurrent mood could occur up to three, four, or more degrees of separation. Unlike previous studies that have indicated how happiness is linked to more indirectly linked network members in the long run,⁴ personal mood in everyday life appears to coexist among those separated by only one or two degrees in one's contact network. To further check how robust it is to represent each network member's overall mood with the average of the person's mood scores during the whole study period, we also fitted the mixed-effects model by replacing the average with the median. The fitted models showed similar results regardless of whether we used the average or the median.

Validation

To further verify the findings from the study, we applied the same mixed-effects model to a similar data set collected in a later study period of seven months from April to October 2015. The results from the supplementary Table S1 indicate that the model estimates and significance levels are similar between the two study periods. In particular, while the average mood score of network members linked by only one step significantly contributed for about 0.22 to a given member's mood score, the coefficient estimate dropped to 0.12 for those separated by two degrees. Like the first study period, one's mood in this second period had little to do with those separated beyond two degrees.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Sensitivity analysis

Our online diary platform allowed us to obtain contact information reported by diary keepers, not by the network members themselves, which raises a big concern about how accurately the diary keepers judged a network member’s mood during a specific contact. To address this potential issue of measurement error on a network member’s mood, we checked the extent to which such judged mood scores were accurate and reliable by matching part of them with the mood scores rated by the network members themselves. This critical step of cross-checking with both actors of a particular contact was possible, because some of our participants joined the ClickDiary study as a group. Among 133 diary keepers, 74 also appeared on the lists of others’ “network members.” As a result, we were able to compare how other diary keepers estimated the mood of these 74 network members during 2,368 contacts with how these 74 network members rated their own mood for each of the identical contacts in their own contact diaries.

Counting the original answering categories, concordant pairs accounted for only 49.1% of all mood rating pairs (Table 2). Of the 50.9% pairs that were discordant, however, 43.9% showed only a one-category difference (e.g., while a diary keeper rated a network member’s mood as “excellent” during a specific contact, that member rated her own mood during that contact as “very good,” which accounted for 14.0% of all 2,368 pairs). Therefore, about 93% of these score differences between the moods rated by diary keeper and network member on the identical contact fell between -1 and 1 .

The initial findings of such cross-checking indicate that we can model the measurement error of a network member’s mood score given by the diary keeper with a normal distribution with a mean of 0 and a standard deviation of 0.5 approximately.

To evaluate how such a measurement error affects the parameter estimates of the mixed-effects model, we simulated 100 datasets of mood scores $O_{ijk}^{(b)}$ by adding a random number from $N(0, 0.5)$ to the score of the j^{th} member given by the i^{th} diary keeper at their k^{th} contact, for $b = 1, 2, \dots, 100$. Fitting the same mixed-effects model with mood scores from the b^{th} simulated dataset, we obtained estimates and standard errors of the l^{th} model parameters, denoted by $\hat{\theta}_l^{(b)}$ and $\hat{s}_l^{(b)}$, respectively.

Taking the influence of measurement error into account, we then estimated the parameter θ_l by the average of these 100 $\hat{\theta}_l^{(b)}$ with the standard error equal to the square root of the sample mean of these 100 $\hat{s}_l^{(b)}$ squared, plus the sample variance of these 100 $\hat{\theta}_l^{(b)}$. It is clear that the augmented measurement errors increased the standard errors of the parameter estimates, as shown in Table 3. Consequently, most of the covariate effects were no more significant. A person's mood score, however, still significantly increased about 0.11 ($p = .0005$) for every additional point scored in the average mood of fellow network members with one degree of separation. The effect size was about 0.07 for those members separated by two degrees, though the strength reduced to marginal significance ($p = .044$). Such effects eventually diminished for those network members at degree three and beyond. The average mood score of diary keepers remained highly influential, with an estimated size of 0.69 ($p < .0001$), while other covariates were irrelevant to how network members' moods varied.

Discussion

With higher quality data collected through a new method (online diary), our results about concurrent personal mood over interpersonal contacts in daily life are consistent with those of the Framingham Heart Study that analyzed 20 years of

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

historical data.⁴ Unlike other studies of egocentric networks, however, we analyzed how personal moods of more than 10,000 network members were linked to one another in 133 complete contact networks, based on the information our diary keepers provided.

To take potential “raters’ effects” into account, we took the essential step to include the diary keepers’ mood scores in the model, which were most influential, as shown in Table 1. Furthermore, we used random components to minimize the effects of the diary keepers’ own scores on how personal mood may occur concurrently among network members. Specifically, our mixed-effects model showed that the average mood of network members linked by only one step contributed significantly to a given member’s mood, as had the average mood of those members separated by two degrees. Other members, or those separated at three degrees and beyond, did not show a clear association.

We reached the findings by a special longitudinal design that followed up more than 100 study participants with online contact diaries for seven months. With the advantage of a web application, the ClickDiary program offers a friendly interface to collect detailed information about an egocentric network, the estimated relationships among all network members, and the mood status of both parties during each contact. To minimize recall bias, the program allowed diary keepers to record only the main contacts that occurred with the same person only within the past 24 hours. That is, ClickDiary encouraged participants to enter the information about their daily contacts as soon as possible.

While there is no gold standard to evaluate the extent to which these diary entries are valid, it would help to cross check the basic entries against similar studies of contact diaries. Participants in this study, for example, recorded an average of 12 contacts per day, which was very close to, although slightly fewer than, the number of

1
2
3 contacts in compatible social surveys and more conventional paper-pencil diary
4 studies in China, Hong Kong, and Taiwan.^{17 18} As discussed earlier, one major
5 difference of the ClickDiary lies in the stricter criteria about what counts as a contact.
6
7 Most previous diary studies either included all fleeting contacts or used “two to three
8 words in exchange” as the minimum requirement for enlisting contacts, whereas the
9 ClickDiary asked for only the contacts that involved at least “three sentences.” This
10 last unique feature may also justify the validity of the basic profile of the findings
11 from the ClickDiary.
12
13

14
15 By cross-checking the mutual ratings of a network member’s personal mood, we
16 were able to evaluate the measurement errors from the way diary keepers rated how
17 others felt during a specific contact. The same rare data also enabled us to evaluate
18 another major concern: how accurately the diary keepers judged the strength of ties
19 among their network members. As in large probability sampling surveys on
20 self-reported egocentric networks, which always involved a risk of informant
21 inaccuracy, we asked the focal persons to judge the strength of ties among those
22 surrounding them. Unlike those surveys where the respondents rated the degree of
23 acquaintanceship among a small number of confidants¹⁹⁻²¹, however, our diary
24 keepers tended to have a tough task, because their contact networks usually stretched
25 far and beyond such core networks. As a result, they often needed to estimate how
26 well any two individuals on their contact lists knew each other even though they did
27 not know either individual well enough in the first place. Although some diary
28 keepers were likely to report “don’t know” under such a circumstance, it remains
29 critical to seek a validity criterion to cross-check their responses.
30
31

32
33 Such a validity criterion again came from a network member’s own answer to
34 how well she actually knew each particular member in the diary keeper’s network.
35
36 Among 133 diary keepers, 74 also appeared on the lists of others’ “network members.”
37
38

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

As in the case with cross-checking personal mood, we further cross-checked the strength of ties among network members, or the alter-to-alter ties. In total, 7,310 individuals appeared on both contact lists of any two diary keepers who also appeared on each other’s contact diary. In 6,956 cases, when a first diary keeper believed that the second diary keeper (who happened to be the first diary keeper’s network member, thus an alter here, or Alter 1) knew one of these overlapped individuals (another alter, or Alter 2), the second diary keeper (Alter 1) also said she did indeed know this particular person (Alter 2). Likewise, in 22 cases when a first diary keeper said that the second diary keeper did not know a network member they shared, the second diary keeper also confirmed that such a tie was absent. As a result, the diary keepers had judged the alter-to-alter ties among the members in their contact networks with an accuracy rate at about 95.5%, which helps justify our strategy of using diary keeper’s contact records to reconstruct part of their contact networks in everyday life.

Even though we were able to validate the alter-to-alter ties of constructed contact networks from these 74 diary keepers, nonetheless, it is noteworthy that these data may not represent all the ties among the 12,070 members in 133 contact networks. Thus, constructing “complete contact networks” out of the diary keepers’ evaluations on how their network members were tied to each other remains a clear limitation. Were all of these network members also involved in keeping a ClickDiary, the estimated ties among them could have been verified by the extent to which they actually contacted one another during the same study period. The task of asking 12,070 network members to record every contact they made for seven months, however, would have been too costly, enormous, and hard to imagine.

As a more realistic, alternative strategy, our design of constructing “proxy” complete contact networks by relying on 133 diary keepers has facilitated a rare analysis of mood correlations among members in different positions in egocentric

networks. Such an approach could be further justified if other conditions also satisfied the assumption that two linked members (that is, any two alters who knew each other according to diary keeper's judgement) indeed made contacts with each other during the study period. To check this assumption, we first identified 5,249 individuals who knew those "network members" (i.e., those on the contact lists) who happened to be diary keepers as well. Then we went back to these diary keepers' own contact diaries and counted how many days each diary keeper actually made contact with each of these 5,249 individuals during the seven-month period.

The days of actual contacts between the pairs were well fitted to a negative binomial distribution, with a mean of 25.3 and a size of 0.63. That is, on average these 5,249 pairs of network members contacted with each other on 25.3 days during the study period, but the range of contacts varies widely and is quite skewed (s.d.=32 days). The results indicate that any two members in a contact network who knew each other had a 90.5% of chance to have at least one contact with each other during the seven months. The finding further supports our underlying assumption that the network members who knew each other indeed had contact with each other during the seven-month period. Such interpersonal contacts, in turn, facilitated structural circumstances under which personal moods could disperse or emerge in parallel among network members.

Even though we are able to demonstrate that two members connecting with each other had a high probability of making actual contacts during the study period, we do not know when the contacts occurred. Without information about the exact timing of each contact among the network members, we are unable to infer the occurrence of contagion or diffusion of personal mood within the network. In other words, our

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

observational data do not support claims of mood contagion. Furthermore, with limited information about the contacted persons' personal background, our model cannot fully adjust for the effects of homophily.²² After conducting a detailed analysis of the contact networks and cross-checking both parties' reports on personal moods involved in the identical contacts from a portion of the network data, however, we do observe concurrent personal mood in the first-order and second-order social ties.

Moreover, the subjects who participated in the ClickDiary study volunteered without a strict sampling procedure. The resulting sample of diary keepers is thus skewed towards female, younger, and better-educated subjects. As common in other diary studies that rely on a small sample of subjects, however, the main goal of our study was not using a representative sample to make an inference to the general population.¹¹ Rather, we used the detailed information about all contacts and ties to build 133 sophisticated complete contact networks, some of which intertwined with one another, which allowed us to examine how personal mood may occur concurrently in everyday life.

Conclusions

As with most other social network studies, it is relatively easy to collect empirical data about the ties between a focal person and those surrounding him/her (or "ego-alter ties"), which are key indicators to understand the structure of an egocentric network.²³ It becomes highly challenging, however, to collect helpful information about, or reconstruct from any sources, the relationships among network members, which allows researchers to analyze the structures of a complete network. One convenient and flexible design in this study relied on some incentives and the sampling strategy to help diary keepers evaluate and confirm how well any pairs of

their network members knew each other. In particular, our system assigned an “absent tie” as the default value of the alter-alter tie (meaning the pair did not know each other), which was the case in about 78.4% of all alter-alter ties. When rating these ties, the diary keeper only needed to either confirm such an absent tie or change the option to either “knew each other well” or “knew each other, but not well.” With a median of 76 alters per ego, an average diary keeper managed to evaluate the strength of $76 \times 75/2 = 2,850$ alter pairs within the study period. Being better motivated to report and confirm such alter-alter ties, as a result, diary keepers in this study completed and verified about 99.97% of all ties, which allowed us to analyze concurrent mood among nearly all network members in egocentric networks.

Using special study designs in ClickDiary, we have been able to cross-check both the network members’ moods and the tie strength among members by matching the diary keepers’ estimates and some of the network members’ own ratings. Future studies could make the best use of all network members’ own reports to reconfirm the strength of ties with one another in complete networks. Such ultimate validity criteria would further verify, in a more comprehensive manner, how accurately diary keepers had judged the ties among the members in their personal networks, even though the distributions of such estimates were similar to those of previous paper-pencil diary studies. Most notably, our findings imply that similar personal mood can occur simultaneously, to varying extents, among the friends, relatives, and other acquaintances clustered around different locations within personal networks. Applying the core concepts of network diffusion and richly designed contact-by-contact data to the inquiries about personal well-being, the current study sheds new light on how social network perspectives can help explain the ways individuals express their personal moods concurrently during social interactions in everyday life.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgments

This research was supported by a grant from the Academia Sinica, Taiwan (AS-103-TP-C03). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We would like to thank Nan Lin for valuable comments and Jie-Yu Sung for administrative support.

Data sharing statement

The data (<https://figshare.com/s/f2223e95335f5dac44df>; DOI: 10.6084/m9.figshare.5208703) and the codes (<https://figshare.com/s/f7247dbcf3fade138859>; DOI:10.6084/m9.figshare.5208706) used in this study are all public available from figshare.com.

Conflicts of Interest

The authors declare that we do not have any competing interests related to this study.

Contributors

JSH and YCF proposed the research project to and obtained funding from Academia Sinica, Taiwan. TCC, TJY, and JSH designed the study. TCC, TJY, THH, and JSH were involved in data collection. TJY, THH, and JSH provided statistical expertise. THH analyzed the data. TCC wrote the first draft of the manuscript. YCF and JSH revised and reorganized the manuscript extensively. All authors contributed to the interpreting the results and critically revising the manuscript for important intellectual content and approved the final version of the manuscript.

References

1. Hatfield E, Cacioppo JT, Rapson RL. Emotional Contagion. *Current Directions in Psychological Science* 1993;2(3):96-100.
2. Hill EM, Griffiths FE, House T. Spreading of healthy mood in adolescent social networks. *P Roy Soc B-Biol Sci* 2015;282(1813).
3. Hill AL, Rand DG, Nowak MA, et al. Emotions as infectious diseases in a large social network: the SISa model. *P Roy Soc B-Biol Sci* 2010;277(1701):3827-35.
4. Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *Bmj* 2008;337:a2338.
5. Rosenquist JN, Fowler JH, Christakis NA. Social network determinants of depression. *Mol Psychiatr* 2011;16(3):273-81.
6. Centola D. The Spread of Behavior in an Online Social Network Experiment. *Science* 2010;329(5996):1194-97.
7. Coviello L, Sohn Y, Kramer AD, et al. Detecting emotional contagion in massive social networks. *PLoS One* 2014;9(3):e90315.
8. Shalizi CR, Thomas AC. Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. *Sociological Methods & Research* 2011;40(2):211-39.
9. Fu YC, Ho HC, Chen HM. Weak ties and contact initiation in everyday life: Exploring contextual variations from contact diaries. *Social Networks* 2013;35(3):279-87.
10. Kahneman D, Krueger AB, Schkade DA, et al. A survey method for characterizing daily life experience: The day reconstruction method. *Science* 2004;306(5702):1776-80.
11. Fu YC. Contact diaries: Building archives of actual and comprehensive personal networks. *Field Methods* 2007;19(2):194-217.
12. Chan TC, Yen TJ, Fu YC, et al. ClickDiary: Online Tracking of Health Behaviors and Mood. *Journal of Medical Internet Research* 2015;17(6):e147.
13. Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. *J Res Pers* 2003;37(6):504-28.
14. Chang YH, Fu YC. The evolution of the Taiwan Social Change Survey. In: Japanese General Social Survey, ed. JGSS Symposium 2003: Birth of JGSS and Its Fruit: Social Surveys in Different Countries and Areas and JGSS. East Osaka City: Institute of Regional Studies, Osaka University of Commerce, 2004:149-60.
15. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of

southern China. *Proceedings Biological sciences / The Royal Society* 2014;281(1785):20140268.

16. José Pinheiro DB, Saikat DebRoy, Deepayan Sarkar, EISPACK authors, Siem Heisterkamp, Bert Van Willigen, R-corenlme: Linear and Nonlinear Mixed Effects Models), R package version 3.1-128, 2016.

17. Fu YC. Measuring personal networks with daily contacts: a single-item survey question and the contact diary. *Social Networks* 2005;27(3):169-86.

18. Fu YC, Wang DW, Chuang JH. Representative Contact Diaries for Modeling the Spread of Infectious Diseases in Taiwan. *PLoS One* 2012;7(10).

19. Burt RS. Network items and the general social survey. *Social Networks* 1984;6(4):293-339.

20. Marsden PV. Network Data and Measurement. *Annu Rev Sociol* 1990;16:435-63.

21. Marsden PV. Survey Methods for Network Data. In: Scott J, Carrington PJ, eds. *The Sage Handbook of social Network Analysis*. London: Sage Publications 2011:370-88.

22. Bollen J, Goncalves B, van de Leemput I, et al. The happiness paradox: your friends are happier than you. *Epj Data Sci* 2017;6.

23. Kaptein M, Nass C, Parvinen P, & Markopoulos P. Nice to Know You: Familiarity and Influence in Social Networks. *System Sciences (HICSS)*, 2013 46th Hawaii International Conference on, pp 2745-2752. 2013.

Table 1. Estimates of the effects associated with a network member's mood score in the mixed-effects models using diary data during May-November 2014.

Variable	Value	Std. Error	t-value	p-value
Male [female]	0.0032	0.0052	0.6236	0.5329
<i>Tie strength with diary keeper</i>				
Knew well	0.0163	0.0062	2.6177	0.0089
[knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0003	0.0097	0.0330	0.9737
Good friend	0.0057	0.0101	0.5634	0.5731
Coworker/trade partner	0.0036	0.0094	0.3887	0.6975
Schoolmate/teacher/student	0.0301	0.0084	3.5603	0.0004
[Other]				
<i>Contact mode (%)</i>				
Face-to-face	0.0277	0.0097	2.8486	0.0044
Voice only	0.0202	0.0136	1.4877	0.1369
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0142	0.0073	-1.9522	0.0509
Daily routine	-0.0205	0.0105	-1.9516	0.0510
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0208	0.0079	2.6152	0.0089
1-4 hours	0.0385	0.0101	3.8141	0.0001
4 hours or more	0.0474	0.0149	3.1866	0.0014
Average mood of diary keeper	0.7427	0.0074	100.9489	0.0000
<i>Average mood of network neighbors</i>				
one degree of separation	0.1326	0.0149	8.8671	0.0000
two degrees of separation	0.0590	0.0191	3.0962	0.0020
All others	-0.0024	0.0167	-0.1460	0.8839

Note: Those listed in brackets are the categories excluded from the model.

Table 2. Pairs of mood ratings from the diary keepers and 74 of their network members who also rated their own moods during the same contacts

		Member's moods rated by the diary keeper			
		(1) Poor	(2) Good	(3) Very good	(4) Excellent
Network member's self-rated moods	(1)	0 (0.0%)	3 (0.1%)	1 (0.0%)	5 (0.2%)
	(2)	5 (0.2%)	42 (1.8%)	203 (8.6%)	76 (3.2%)
	(3)	5 (0.2%)	171 (7.2%)	877 (37.0%)	331 (14.0%)
	(4)	3 (0.1%)	75 (3.2%)	326 (13.8%)	245 (10.3%)

Table 3. Combined effect estimates of the same mixed-effects models fitted with 100 different simulated datasets of mood scores augmented with measurement errors.

Variable	Value	Std. Error	t-value	p-value
Male [female]	0.0041	0.0110	0.3727	0.7094
<i>Tie strength with diary keeper</i>				
Knew well	0.0156	0.0128	1.2209	0.2222
[knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0011	0.0196	0.0577	0.9540
Good friend	0.0079	0.0191	0.4146	0.6784
Coworker/trade partner	0.0031	0.0178	0.1764	0.8600
Schoolmate/teacher/student	0.0291	0.0163	1.7893	0.0736
[Other]				
<i>Contact mode (%)</i>				
Face-to-face	0.0234	0.0196	1.1951	0.2321
Voice only	0.0210	0.0288	0.7289	0.4661
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0170	0.0147	1.1505	0.2500
Daily routine	-0.0170	0.0221	0.7667	0.4433
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0193	0.0183	1.0508	0.2934
1-4 hours	0.0349	0.0211	1.6545	0.0981
4 hours or more	0.0436	0.0288	1.5164	0.1294
Average mood of diary keeper	0.6859	0.0167	41.0983	0.0000
<i>Average mood of network neighbors</i>				
one degree of separation	0.1117	0.0319	3.5004	0.0005
two degrees of separation	0.0714	0.0354	2.0157	0.0439
All others	0.0118	0.0289	0.4091	0.6825

Note: Measurement errors were simulated from normal distributions with a zero mean and 0.5 as the standard deviation for each dataset.

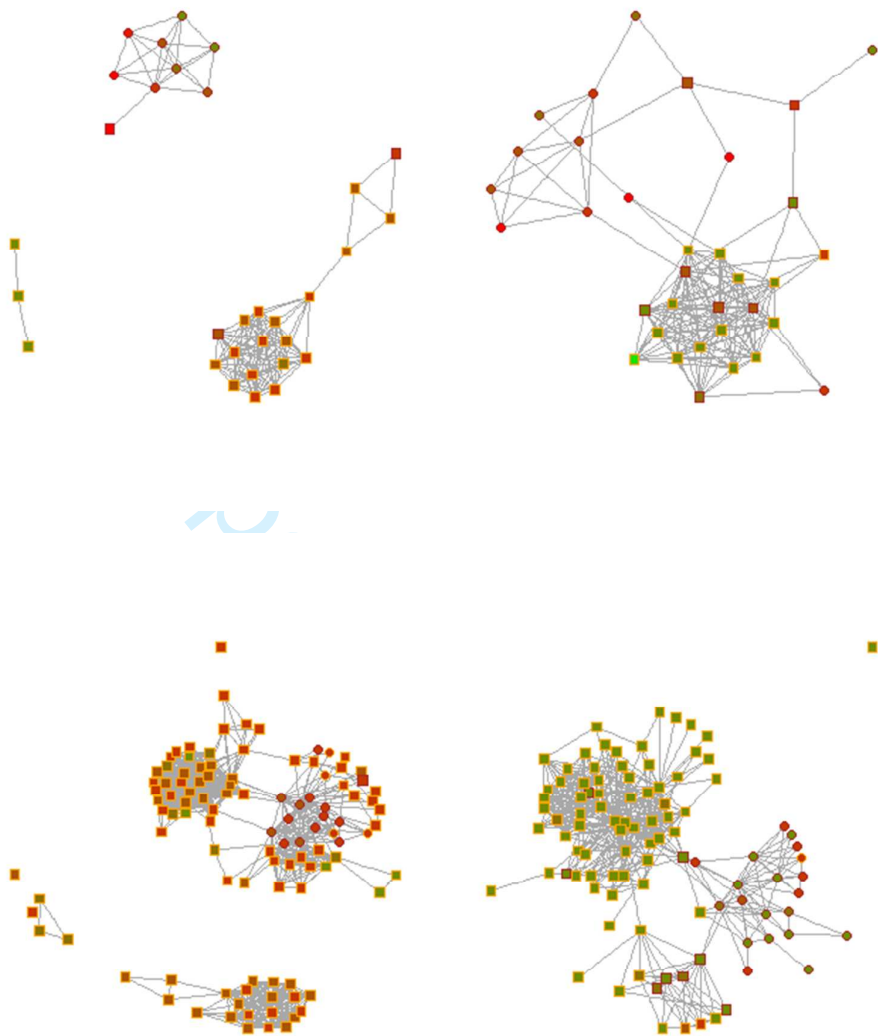


Figure 1. The clustering patterns of average personal mood in the complete contact networks of four diary keepers. Each node represents a person, whose relationship with the diary keeper is displayed with a circle for family members, relatives, and good friends and a square for the others. The frame color of the node reveals the strength of tie to the diary keeper (brown for “know each other very well”; orange for “know each other, but not well”). Node color denotes average mood scores of the persons during the study period, with a color gradient ranging from green, which indicates the worst mood, to red, which indicates the best mood. The figure does not include the diary keeper, who is linked to everyone in respective contact network.

Supplementary

We launched a second wave of data collection with the same incentives to promote the participation in April 2015. To verify the findings from the data collected in the seven months of 2014, we retrieved seven-month contact diaries from April 1 to October 31, 2015. During this second study period, 130 participants completed at least 30 days of contact diaries with at least 30 persons. Note that 54 of them had participated in the May-November 2014 study as well. Very similar to the data collected in the previous study period, these diary keepers recorded 156,892 contacts with 13,539 persons in 2015. The contact lists among the 130 participants averaged 104 persons, with a minimum, median, and maximum of 30, 57 and 1,372, respectively. Diary keepers recorded about 9 contacts per day in 2015, with a minimum of 3, a median of 7, and a maximum of 64. On average, the persons on the contact lists appeared 11 times during the 7 months, ranging from 1 to 211, with a median of 3 times. During this second study period, 8,680 contacts were missing mood scores. As a result, 365 persons were excluded for having no average mood scores. We also excluded network members whom the diary keepers did not know (12%), and the number of contacted persons was reduced to 11,604 with 145,813 contacts for modeling mood variation. Diary keepers confirmed nearly all (99.69%) of the ties between any two persons on the contact lists. The aggregated data showed that about 80.9% of the pairs did not know each other, 9.6% of them knew each other well, and 9.5% just knew each other. For a comparison with the diary data collected during 2014, we listed the model estimates of the 2015 data in Table S1.

Table S1. Estimates of the effects associated with a network member’s mood score in the mixed-effects models using diary data during April-October 2015.

Variables	Value	Std. Error	t-value	p-value
Male [female]	-0.0090	0.0045	-1.9745	0.0484
<i>Acquaintanceship with diary keeper</i>				
Knew well	0.0115	0.0057	2.0085	0.0446
[Knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0017	0.0080	0.2094	0.8342
Good friend	0.0048	0.0091	0.5229	0.6011
Coworker/trade partner	-0.0005	0.0078	-0.0580	0.9538
Schoolmate/teacher/student	0.0057	0.0078	0.7335	0.4633
[Other]				
<i>Contact mode (%)</i>				
Face-to-face	0.0353	0.0083	4.2326	0.0000
Voice only	0.0212	0.0132	1.6103	0.1074
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0265	0.0074	-3.5783	0.0003
Daily routine	-0.0499	0.0107	-4.6710	0.0000
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0093	0.0077	1.2072	0.2274
1-4 hours	0.0309	0.0094	3.2984	0.0010
4 or more hours	0.0359	0.0132	2.7261	0.0064
Average mood of diary keeper	0.6667	0.0074	89.9680	0.0000
<i>Average mood of network neighbors</i>				
One degree of separation	0.2180	0.0155	14.0250	0.0000
Two degrees of separation	0.1156	0.0195	5.9321	0.0000
All others	-0.0295	0.0164	-1.7921	0.0732

Note: Those listed in brackets are the categories excluded from the model.

STROBE statement: checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Checklist	Page #
Title and abstract				
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	Y	p2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	Y	p2
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Y	p5-p6
Objectives	3	State specific objectives, including any prespecified hypotheses	Y	p7
Methods				
Study design	4	Present key elements of study design early in the paper	Y	p8-p9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Y	p9-10
Participants	6	(a) Cohort study ? Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study ?Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross sectional study ?Give the eligibility criteria and the (b) Cohort study ? For matched studies, give matching criteria and number of exposed and unexposed Case-control study ?For matched studies, give matching criteria and the number of controls per case	Y <	

Participants	13*	(b) Give reasons for non-participation at each stage	NA	NA
		(c) Consider use of a flow diagram	NA	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Y	p14-p15
		(b) Indicate number of participants with missing data for each variable of interest	Y	p14-p15
		(c) Cohort study ? Summarise follow-up time (eg average and total amount)	Y	p14-p15
Outcome data	15*	Cohort study ? Report numbers of outcome events or summary measures over time	Y	p14-p15
		Case-control study? Report numbers in each exposure category, or summary measures of exposure	NA	NA
		Cross sectional study? Report numbers of outcome events or summary measures	NA	NA
Main results	16	(a) Report the numbers of individuals at each stage of the study?eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Y	p14-p15
		(b) Give reasons for non-participation at each stage	NA	NA
		(c) Consider use of a flow diagram	NA	NA
Other analyses	17	Report other analyses done?eg analyses of subgroups and interactions, and sensitivity analyses	Y	p16-p18
Discussion				
Key results	18	Summarise key results with reference to study objectives	Y	p19-21
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Y	p20-p23
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other	Y	p20-p23
Generalisability	21	Discuss the generalisability (external validity) of the study results	Y	p23-p25
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Y	p25

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross sectional studies.

BMJ Open

Detecting concurrent mood in daily contact networks: an online participatory cohort study with a diary approach

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-020600.R1
Article Type:	Research
Date Submitted by the Author:	08-Mar-2018
Complete List of Authors:	Chan, Ta-Chien; Academia Sinica, Research Center for Humanities and Social Sciences Yen, Tso-Jung; Academia Sinica, Institute of Statistical Science Hu, Tsuey-Hwa; Academia Sinica, Institute of Statistical Science Fu, Yang-chih; Academia Sinica, Institute of Sociology Hwang, Jing-Shiang; Academia Sinica, Institute of Statistical Science
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Sociology
Keywords:	contact diary, emotional contagion, social networks, transmission of mood

SCHOLARONE™
Manuscripts

1
2
3
4 Detecting concurrent mood in daily contact networks: an online
5
6
7 participatory cohort study with a diary approach
8

9 Ta-Chien Chan^a, Tso-Jung Yen^b, Tsuey-Hwa Hu^b, Yang-chih Fu^{c,1*}, Jing-Shiang
10
11 Hwang^{b,2*}
12
13
14

15 ^aResearch Center for Humanities and Social Sciences, Academia Sinica, 128
16
17 Academia Road, Section 2, Nankang, Taipei, 115 Taiwan
18

19 ^bInstitute of Statistical Science, Academia Sinica, 128 Academia Road, Section 2,
20
21 Nankang, Taipei, 115 Taiwan
22

23 ^cInstitute of Sociology, Academia Sinica, 128 Academia Road, Section 2, Nankang,
24
25 Taipei, 115 Taiwan
26
27
28
29

30 Major classification: Social Sciences; Minor classification: Social Sciences
31

32 Short title: mood spread through networks
33
34
35

36 Keywords: contact diary, emotional contagion, social networks, transmission of mood
37
38
39

40 *Co-corresponding Authors:
41

42 Yang-chih Fu¹:
43

44 Tel: 886-2-2652-5149; E-mail: fuyc@sinica.edu.tw
45

46 Jing-Shiang Hwang²:
47

48 Tel: 886-2-2783-5611; E-mail: hwang@sinica.edu.tw
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objectives: This paper examines how people express personal mood concurrently with those connected with them by one or two degrees of separation.

Design: Participatory cohort study.

Setting: Online contact diary.

Participants: 133 participants kept online contact diaries for seven months in 2014, which included 127,455 contacts with 12,070 persons.

Main outcome measures: Diary keepers rated a contacted person's mood during each specific contact, as well as the strength of ties between any pairs of such contacted persons. Such rich information about ties and contacts enable us to construct a complete contact network for each diary keeper, along with the network members' mood and tie strength. We calculate one's overall mood by that person's average mood score during the study period and take the shortest path between any given pair of contacted persons as the degree of separation. We further assume that two connecting persons in a complete contact network have made contact with each other during the study period, which allows us to examine whether and how personal moods occur concurrently within these contact networks.

Results: Using mixed-effects models while controlling for covariates at individual, tie and contact levels, we show that personal mood score positively and significantly correlates with the average mood among those directly tied to the person. The same effect remains positive and significant, though the effect size is reduced by about one half, for those connected to the person by two degrees. The mood of anyone separated by more than two degrees is statistically irrelevant.

Conclusions: Applying core social network perspectives and rich data at both tie and contact levels to inquiries about subjective well-being, the current study sheds new light on how an improved diary approach can help explain the ways in which

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

individuals express their personal moods concurrently during social interactions in
everyday life.

For peer review only

Strengths and limitations of this study

1. The observed online contact diaries allow us to construct 133 complete contact networks with more than 12,000 members, which help uncover how personal moods vary among network members.
2. Some intertwining contact diaries further enable us to cross-examine personal mood and tie strength as rated by both parties of social interactions, solidifying our findings about how a bottom-up social network approach helps reveal concurrent mood in everyday life.
3. Without information about the exact timing of each contact among network members, we cannot infer the occurrence of contagion or diffusion of personal mood.
4. With limited information about the contacted persons' personal background, our model cannot fully adjust for the effects of homophily.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

Experiments have demonstrated that emotions tend to spread to others during social interactions,¹ as in the case when diseases, behaviors, and ideas transmit through social networking.^{2 3} Analyses from other large-scale data have further revealed similar spreading patterns of both positive and negative emotions, such as happiness and depression.^{4 5} Like the diffusion of behaviors and attitudes amid face-to-face social networking, emotive sharing and contagion may also cover large social circles through online social contacts by text or voice on Facebook, Twitter, Skype, Google talk, and other social media.^{6 7}

The literature has identified such a common phenomenon, generally known as “emotional contagion,” through various study designs in different research settings. One seminal work argued that emotion or mood can easily be transmitted within social networks, because people are inclined to synchronize others’ facial expressions, voices, postures, movements, and “emotional behaviors” and get feedback from such “mimicry.”⁸ Not only does such a diffusion of emotions and moods occur among those who close to one another, but it also happens from moment to moment in everyday encounters with less known others. More recent experimental studies have shown similar contagion of happiness and fear through facial mimicry by measuring facial electromyography.^{9 10} In addition to face-to-face contacts, furthermore, some posts in social media also can trigger users’ emotional contagion on a massive scale.¹¹

Most existing studies have demonstrated that certain emotions tend to spread from direct contacts over a short period of time, but relatively few have examined whether personal mood may also spread from, or at least occur concurrently among indirect contacts within social networks over a longer period of time. Two recent studies, for example, showed that our mood can be influenced not only by those familiar to us, but also by friends’ friends whom we do not personally know.^{2 3} More

specifically, the spread of personal happiness reaches up to three degrees of separation along social networks, according to one such rare empirical study, the Framingham Heart Study, which analyzed historical data over 20 years.⁴ The longitudinal analyses indicated that an individual was 15.3% more likely to be happy if a directly connected network member (with one degree of separation) was happy; the effect decreased to 9.8% for those separated by two degrees.⁴

As such findings may have inspired studies of mood diffusion, most non-experimental studies are insufficient to examine the actual mechanism of “diffusion,” because the observational data on which they are based offer no such advantages. Thus, it would be particularly difficult to determine and infer from observational data how personal moods spread or diffuse.¹² Even without claiming the causal effects that are essential for explaining spread or diffusion, however, it would be revealing to examine whether and how similar patterns of concurrent mood may exist under other circumstances. While previous longitudinal studies have focused on how individuals’ happiness or depression changes across waves of surveys, an alternative “bottom-up” approach tends to uncover subtle patterns of concurrent mood by examining how one’s mood fluctuates at the micro level, contact by contact.¹³⁻¹⁵ Due to the methodological limitation, it may be inappropriate to pursue inquiries about “mood diffusion or contagion” solely by using observational data, even if one could replicate the decades-long study by creating another huge number of personal networks with records of participants’ relationships and emotions. As an alternative approach, our study, based on a different format, should provide a new perspective that would further enrich studies on how personal moods can be linked to one another in contact networks.

By extending the bottom-up approach to social network studies, we aim to examine whether and how personal moods occur concurrently within a contact

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

network, using data collected via an online platform, ClickDiary, over a seven-month period between May 1 and November 30, 2014. The ClickDiary program uses a web-based platform to collect data on participants’ health behaviors and all one-on-one interpersonal contacts in everyday life.¹⁶ The data retrieved for this study have a nested hierarchical structure, including detailed information about 133 participants (or “diary keepers,” who recorded details about their social interactions in daily life), 12,070 contacted persons (or “network members,” including 74 persons who also participated in the same diary keeping platform), as well as 127,455 contacts.

Two features in the ClickDiary platform are central to our research design. First, diary keepers reported their own mood for each one-on-one contact by selecting one of the following score categories that best matched their estimates: (1) poor, (2) good, (3) very good, or (4) excellent; they then evaluated the mood of each network member during the contact. Second, diary keepers, to their best knowledge, rated how well a given pair of network members knew each other. Once a diary keeper confirmed all of their interpersonal ties, we used these two critical features as the backbone to construct a “complete contact network” surrounding that diary keeper, the focal person. Within such a complete contact network, we linked the nodes that represent the network members to each other by interpersonal ties. The overall mood of each network member can be represented by taking the average of the person’s mood scores recorded in the contact diary during the whole study period. Furthermore, we assumed that two connecting persons in a complete contact network have made contact with each other during the study period, which allows us to examine whether and how personal moods occur concurrently within these contact networks. That is, with the assumption, concurrent mood of the linked members could be explained partially as caused by personal contact. We applied a mixed-effects model to analyze

overall mood scores of the members in the 133 complete contact networks to examine whether and how an individual's mood is associated with those within one or two degrees of separation, as well as other members in the complete contact networks.

Materials and Methods

The ClickDiary Program

The ClickDiary program (<http://cdiary.tw>) uses a web-based platform, written in Chinese, to collect data on participants' daily health behaviors and interpersonal one-on-one contacts.¹⁶ One unique feature of ClickDiary is the friendly interface designed for clicking options on structured diary items via a website or mobile apps, making it easier to record responses whenever it is convenient for participants.

Public Involvement

Our research team recruited participants from various channels including university students, school teachers and administrative employees, volunteers at health-promotion centers, hospital patients, and community college students, as well as other adults in the general population. When registering in the ClickDiary platform, all diary keepers are required to give online informed consent before starting to keep the diaries. Upon signing up for the program, moreover, participants provided socio-demographic information, including age, gender, place of residence, marital status, and current job. The program also collects participants' Big Five personality traits (openness, conscientiousness, extraversion, agreeableness, neuroticism),¹⁷ height and weight, perceived health status and happiness, the number (and characteristics) of people contacted during the day, along with a baseline health module that borrowed items from the International Social Survey Programme.¹⁸ We provided an interactive Web chart summarizing the records in each participant's contact and health diaries. In addition, participants could gain insight from their

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

overall contact patterns by checking their contact network tree we developed. This current study has been approved by the Institutional Review Board on Humanities and Social Science Research (IRB-HS), Academia Sinica (AS-IRB-HS 02-13022).

Contact diary

When adding a person to the contact list for the first time, the diary keepers provided the person’s background information, such as age and gender, and evaluated several aspects of their relationship with the person, including the duration of acquaintanceship, degree of familiarity, the most frequent mode used for contact (face-to-face, voice only, or text only), contact frequency, and the likelihood of discussing important matters. In this particular diary program, a “contact” refers to one-on-one exchange that involves at least three verbal or written sentences, a definition somewhat narrower than most previous studies using the contact diary approach.¹⁹

The program also asked participants to evaluate the degree to which any pair of persons on the contact list was familiar with each other. Before starting to enter the contact details with any new person, a diary keeper had the chance to estimate, on a scale from 1 to 3, how well this particular person knew each of randomly selected five other persons already recorded in the diary. After this first step, another random sample of five different persons’ names popped up, so that the diary keeper could continue judging all the ties between any pair of persons within the contact network. The process continued, randomly adding five new names to the list at a time, until the diary keeper finished rating the strength of all alter pairs. The design allowed the diary keeper to evaluate all alter-alter ties when it was convenient to do so, thus achieving maximum flexibility and encouraging a higher completion rate. In addition, the diary entries also focused on 11 contact attributes, including when the contact took

place, who initiated the contact, the major mode of the contact, the duration and content of the contact, where the diary keeper and the contacted person were during the contact, the extent to which the contact felt beneficial to the diary keeper, the mood of each party during the contact, and whether the contacted person showed any cold symptoms.

Data retrieved for the study

From May 1 to November 30, 2014, 133 residents in Taiwan each completed at least 30 days of contact diaries and recorded one-on-one interpersonal contacts with at least 30 persons. We retrieved the seven-month contact diaries from these 133 participants, with all personal identities removed, which consisted of 141,909 contacts with 16,139 contacted persons. The length of the contact lists varied substantially among the 133 participants, ranging from 30 to 1,399, with a median of 76. The participants recorded an average of 12 contacts a day, with a minimum of 1 and a maximum of 56. The contacted persons varied greatly in terms of how often they appeared in the diaries, ranging from only once to daily (214 times during the seven-month period), even though the average frequency was 8 times.

To measure each person's mood, diary keepers selected one from the following score categories that best matched their estimates: (1) poor, (2) good, (3) very good, or (4) excellent. As expected, diary keepers sometimes were unable to estimate a contacted person's mood and consequently answered "Don't Know" for the item. The records during the study period showed such "Don't Know" answers for the mood item appearing in 9,042 contacts. We treated these contacts (with 845 contacted persons) as missing, which reduced the valid number of contacted persons to 15,294.

We further excluded the cases where the diary keepers did not know a contacted person (i.e., the category that accounts for about 21.1% of the relationships) from the

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

subsequent analysis for three reasons. First, interactions with strangers carry very different implications in studies of network diffusion. Even though they had actual contacts with the diary keepers, keeping total strangers in a personal network would have left too many unknown or uncertain links, because the diary keepers were unlikely to judge how well these strangers knew one another. Second, it would have been more difficult for the diary keepers to evaluate strangers' mood, which also tended to yield missing or less reliable mood rating. Third, probably due to such uncertainties, our different modeling efforts while retaining the ties with these strangers resulted in unreasonable noise to the analysis. The final data for modeling mood spread included 12,070 contacted persons of 133 diary keepers who had made 127,455 contacts during the study period.

Network construction

As a whole, diary keepers were able to confirm nearly all (99.97%) interpersonal ties in terms of the familiarity between any pair of contacted persons. Of these ties, averaged across the 133 diary keepers, 78.4% were considered absent (the pairs did not know each other), 10.6% were strong (knew each other well), and 11.1% were weak (knew each other, though not well). For each diary keeper, we constructed a complete contact network, in which each alter, or contacted person, represents a node, and two nodes (other than the focal person) are deemed to be connected when the reported alter-alter tie is either strong or weak, rather than absent. In contrast, nodes are not linked when the pair of network members are strangers to each other. Using the shortest path between any given pair of nodes in the network, we calculate "distance" (or "the degree of separation") between the network members pair by pair. We focus on the links among these network members only, while excluding any links leading to and from the unique node that represented the diary keeper (or "ego-alter

ties”). Excluding such ties linked by the diary keepers helps simplify the calculation of “distance.” Otherwise, any two given nodes in the network would have at most two degrees of separation, because all the network members would be directly linked to the diary keepers. Finally, we defined a network member’s overall mood by the average of the mood scores reported by diary keepers during the study period, ranging from 1 to 4.

To illustrate how alter-alter ties cluster into a subset of ego’s complete contact network, we drew a figure, using R package igraph, to display the network patterns in each of four diary keepers’ complete contact networks during the study period (Figure 1). The figure shows that the clustering patterns may differ significantly among diary keepers. Within each of the four complete contact networks, those members who are closer to each other in terms of distance also tend to average similar scores in personal mood during all interpersonal contacts.

Statistical Analysis

To examine whether and how a network member’s mood may be associated with those separated by one or two degrees in each of the 133 complete contact networks, we first calculated the average mood of those members who were directly connected to a particular member (that is, those connected by one degree of separation), and then we obtained the average mood of those at two degrees of separation and the average mood of all other members. The average mood score of the j^{th} member in the i^{th} complete contact network can be simply obtained by calculating $Y_{ij} = \sum_{k=1}^{K_{ij}} O_{ijk} / K_{ij}$, where O_{ijk} is the mood score of their k^{th} contact given by the i^{th} diary keeper and K_{ij} is the number of contacts during the study period. Let D_{1ij} and D_{2ij} indicate the mean mood scores of those separated by one and two degrees from the j^{th} member in the i^{th}

network, respectively. Finally, D_{3ij} measures the average mood scores for those beyond two degrees of separation from the j^{th} network member.

For this study, we applied a mixed-effects model to analyze the relationship between a person's mood score and the average mood scores of those surrounding the person in the network (or "network neighbors"), while controlling for the effects of potential covariates on the person's mood. The model is given as

$$Y_{ij} = (\alpha_0 + a_{0i}) + (\alpha_1 + a_{1i})D_{1ij} + (\alpha_2 + a_{2i})D_{2ij} + (\alpha_3 + a_{3i})D_{3ij} + \sum_{h=1}^p \beta_h X_{hij} + \varepsilon_{ij} \quad \text{for } i = 1, 2, \dots, n \text{ and } j = 1, 2, \dots, n_i,$$

where n is the number of networks; n_i is the size of the i^{th} network; random components a_{li} are assumed to be normally distributed with a mean of 0 and variance σ_l^2 for $l = 0, 1, 2, 3$; and the error term is normally distributed with a mean of 0 and variance τ^2 . Our main interest is to estimate the fixed effects of α_1 , α_2 , and α_3 , which measure whether and how one's mood correlates with those separated by one degree, two degrees, and the others, respectively, within the same complete contact network.

We applied exploratory data analysis tools to identify potential covariates X_{hij} , $h = 1, 2, \dots, p$, for the j^{th} member in the i^{th} network. The final $p = 14$ covariates selected in the model are described in the following. In addition to controlling for the possible gender effect of the individual member, the other covariates included in the mixed-effects model measure a variety of features at both tie and contact levels. Two groups of covariates tap the relationships between the diary keepers and their network members. The first is a binary variable indicating that diary keepers knew a network member either well or "not well" (the latter serves as the base category in the model). The second group of covariates distinguishes four types of the network members'

relationships to the diary keepers: family members or relatives (13.2%); good friends (13.5%); coworkers or trade partners (12.7%); and schoolmates, teachers, or students (18.4%). The base category for the group is “others.”

The next group of covariates covers three major contact features: mode, purpose, and duration. Because the main goal of the analysis is to examine how personal moods vary among network members, we sum up these contact features for each member. The modes of contact, for example, may play a key role in determining how well the diary keepers judged others' mood. To verify such an effect, we include two contact modes as covariates: face-to-face and “voice only,” leaving “text only” out of the model. From all of the contacts between the diary keepers and each particular network member during the seven-month study period, we calculate the proportion (percentage) of each of the three modes for the network member. Because the percentages of all three modes for each network member add up to 1, we keep only the first two modes (face-to-face and voice) in the analysis. Suppose that a member-keeper pair had K times of contacts during the seven-month study period, and among which, the contact modes of “face-to-face” and “voice only” were k_1 and k_2 times, respectively. We defined the two covariates face-to-face as the proportion k_1/K and voice only as k_2/K for the analysis, and treated the other mode of “text only” as the reference category and excluded it from the model. The averaged percentages of face-to-face and voice only contacts among the network members were about 79.8% and 8.2%, respectively.

The model also includes the percentage for each of the two kinds of contact purposes, “work- or school-related” and “daily routine,” which were about 30.4% and 12.4% on average, respectively, while excluding “other purposes.” We further add contact duration into the model, using the percentages for contacts that last 5-59 minutes, 1-4 hours, or more than 4 hours, excluding those lasting under 5 minutes. On

average, 47% of the contact durations was 5-59 minutes, 20% was 1-4 hours, and 6% was more than 4 hours.

Because diary keepers reported their own mood as well as the network members' mood in the same diaries, the two scores are expected to be highly correlated. To take such an effect into account, we further controlled for the most influential covariate that measures the average mood score of the i^{th} diary keeper when contacted with the j^{th} network member during the study period. In other words, the member's mood score was influenced by the diary keeper's mood score on the contact. Hence, the covariate of the average mood score of the i^{th} diary keeper when making contact with the j^{th} network member should be influential and must be included in the model for adjustment. We used the lme function from the R package "nlme" to estimate the model parameters.²⁰

Results

Among the original 133 diary keepers, about 80% were female (106/133). Compared to a representative sample of general population, the group tended to be younger and better educated. Percentages of those under age 23 (college students), 23-39, 40-59 and 60 or more were 33.8%, 36.1%, 21.8% and 8.3%, respectively. At least 82.7% had ever gone to college. Like previous contact diary studies, the study participants were overrepresented by females and better educated subpopulations.¹⁵

The high percentage of female diary keepers probably yielded more females on the contact lists (65.5%), although the latter was actually more balanced than the gender distribution among the diary keepers. The age distribution of the 12,070 contacted persons was not very skewed, with age groups of 1-19, 20-29, 30-49, 50-59, and 60 or more accounting for about 8.4%, 21.8%, 31.1%, 18.5%, and 16.9%, respectively. About 53.3% of the contacted persons the diary keepers knew were

known very well.

In fitting the mixed-effects model, we excluded about 12% of network members with any covariates missing, mainly those who had either no degree 1 or 2 neighbors. As a result, the following results were obtained from the remaining 10,581 network members with complete covariates. Table 1 shows model estimates of the parameters associated with individual mood variation. On average, a person's mood score increased about 0.13 ($p < .0001$) for every additional point scored in the average mood of fellow members who were directly connected to the person (with one degree of separation). The average mood score of those members linked by two steps also contributed significantly ($p = .002$), with an effect size of about 0.06, to the person. Such contribution diminishes, however, for those members at degree three and beyond (linked by at least three steps).

These key findings are noteworthy, because they have been adjusted by several highly relevant covariates in the mixed-effects model. Since the diary keepers rated the mood of both parties after a contact, the two scores were often highly correlated. As shown in Table 1, a member's mood score was strongly determined by 0.74 of the average of all the diary keeper's self-rated mood scores when making contact with that specific member. When the diary keeper knew a member well, that member had a better chance of receiving a higher score on personal mood. A network member also tended to receive a higher mood score for a face-to-face contact and a contact that lasted longer. When a contact was about work or school, or was part of the daily routine, the network member's mood was not as good as that of other contacts, at least based on the diary keeper's judgment.

The estimated standard deviations of the three random components for one degree of separation, two degrees of separation and beyond two degrees were 0.0015, 0.0372, and 0.0037, respectively. It is clear that the strong mood associations between

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

directly linked members were consistent across the 133 networks. The less strong mood associations between members and neighbors separated by two degrees varied a little large, compared to the estimated fixed effect, among the 133 networks. One possible reason for the large variation among the 133 networks may be the network sizes which ranged from 30 to 1,399, with a median of 76. Members in a small network tended to have relatively few neighbors separated by two degrees. The average mood of smaller number of neighbors may incur large uncertainty on the estimated association effect.

Robustness Checks

To verify the effects of different degrees of separation, we tried a separate analysis of those members at three degrees of separation, using only 8,505 network members with all valid variates. The alternative analytic design showed little change in both the coefficient estimates of the covariates and the coefficient estimates of one, two, and three degrees of separation (0.13, 0.08, and 0.03, respectively). Thus, it would be unnecessary and unfruitful to further determine whether concurrent mood could occur up to three, four, or more degrees of separation. Unlike previous studies that have indicated how happiness is linked to more indirectly linked network members in the long run,⁴ personal mood in everyday life appears to coexist among those separated by only one or two degrees in one’s contact network. To further check how robust it is to represent each network member’s overall mood with the average of the person’s mood scores during the whole study period, we also fitted the mixed-effects model by replacing the average with the median. The fitted models showed similar results regardless of whether we used the average or the median.

Validation

To further verify the findings from the study, we applied the same mixed-effects model to a similar data set collected in a later study period of seven months from April to October 2015. The results from the supplementary Table S1 indicate that the model estimates and significance levels are similar between the two study periods. In particular, while the average mood score of network members linked by only one step significantly contributed for about 0.22 to a given member's mood score, the coefficient estimate dropped to 0.12 for those separated by two degrees. Like the first study period, one's mood in this second period had little to do with those separated beyond two degrees.

We conducted a subset validation to cross-check the accuracy from a network member's own answer with how well she actually knew each particular member in the diary keeper's network. This critical step of cross-checking alter-to-alter ties was possible because some of our participants joined the ClickDiary study as a group. Among 133 diary keepers, 74 also appeared on the lists of others' "network members." In total, 7,310 individuals appeared on both contact lists of any two diary keepers who also appeared on each other's contact diary. In 6,956 cases, when a first diary keeper believed that the second diary keeper (who happened to be the first diary keeper's network member, thus an alter here, or Alter 1) knew one of these overlapped individuals (another alter, or Alter 2), the second diary keeper (Alter 1) also said she did indeed know this particular person (Alter 2). Likewise, in 22 cases when a first diary keeper said that the second diary keeper did not know a network member they shared, the second diary keeper also confirmed that such a tie was absent. As a result, the diary keepers had judged the alter-to-alter ties among the members in their contact networks with an accuracy rate at about 95.5%, which helps justify our strategy of using diary keeper's contact records to reconstruct part of their contact networks in everyday life.

To check the assumption that two connecting persons in a complete contact network have made contact with each other during the study period, we first identified 5,249 individuals who knew those "network members" (i.e., those on the contact lists) who happened to be diary keepers as well. Then we went back to these diary keepers' own contact diaries and counted how many days each diary keeper actually made contact with each of these 5,249 individuals during the seven-month period. The days of actual contacts between the pairs were well fitted to a negative binomial distribution, with a mean of 25.3 and a size of 0.63. That is, on average these 5,249 pairs of network members contacted with each other on 25.3 days during the study period, but the range of contacts varies widely and is quite skewed (s.d.=32 days).

The results indicate that any two members in a contact network who knew each other had a 90.5% of chance to have at least one contact with each other during the seven months. The finding further supports our underlying assumption that the network members who knew each other indeed had contact with each other during the seven-month period. Such interpersonal contacts, in turn, facilitated structural circumstances under which personal moods could disperse or emerge in parallel among network members.

Sensitivity analysis

Our online diary platform allowed us to obtain contact information reported by diary keepers, not by the network members themselves, which raises a big concern about how accurately the diary keepers judged a network member's mood during a specific contact. To address this potential issue of the diary keeper's error in judging a network member's mood, we checked the extent to which such judged mood scores were accurate and reliable by matching part of them with the mood scores rated by the network members themselves, as with the case with cross-checking alter-to-alter ties.

As a result, we were able to compare how other diary keepers estimated the mood of these 74 network members during 2,368 contacts with how these 74 network members rated their own mood for each of the identical contacts in their own contact diaries.

Counting the original answering categories, concordant pairs accounted for only 49.1% of all mood rating pairs (Table 2). Of the 50.9% pairs that were discordant, however, 43.9% showed only a one-category difference (e.g., while a diary keeper rated a network member's mood as "excellent" during a specific contact, that member rated her own mood during that contact as "very good," which accounted for 14.0% of all 2,368 pairs). Therefore, about 93% of these score differences between the moods rated by diary keeper and network member on the identical contact fell between -1 and 1 .

We performed a sensitivity analysis to check for the potential effects of the diary keeper's judgement error on the parameter estimates of the mixed-effects model, using a perturbation approach with the initial findings of such cross-checking. Specifically, we simulated 200 datasets of mood scores that deviated from the observed scores by 0, 1, and 2 with probabilities 0.5, 0.43, and 0.07, respectively. Fitting the same mixed-effects model with mood scores from the b^{th} simulated dataset, we obtained estimates and standard errors of the l^{th} model parameters, denoted by $\hat{\theta}_l^{(b)}$ and $\hat{s}_l^{(b)}$, respectively.

Taking the influence of judgement error into account, we then estimated the parameter θ_l by the average of these 200 $\hat{\theta}_l^{(b)}$ with the standard error equal to the square root of the sample mean of these 200 $\hat{s}_l^{(b)}$ squared, plus the sample variance of these 200 $\hat{\theta}_l^{(b)}$. It is clear that the augmented judgement errors increased the standard errors of the parameter estimates, as shown in Table 3. Consequently, most of the covariate effects were no more significant. A person's mood score, however,

still significantly increased about 0.088 ($p = .018$) for every additional point scored in the average mood of fellow network members with one degree of separation. The effect size was about 0.079 for those members separated by two degrees, though the strength reduced to marginal significance ($p = .049$). Such effects eventually diminished for those network members at degree three and beyond. The average mood score of diary keepers remained highly influential, with an estimated size of 0.507 ($p < .0001$), while other covariates were irrelevant to how network members' moods varied.

Discussion

With higher quality data collected through a new method (online diary) than those collected from one-shot survey data, our results about concurrent mood over interpersonal contacts in daily life are consistent with those of the Framingham Heart Study, which analyzed 20 years of historical data.⁴ Unlike other studies of egocentric networks, however, we analyzed how personal moods of more than 10,000 network members were linked to one another in 133 complete contact networks, based on the information our diary keepers provided.

To be more consistent with how we collected and analyzed the data, we focus on “concurrent mood” in our subsequent discussions. We actually analyzed a hybrid construct derived from the average mood for each ego-alter pair, even though conceptually, we have relied on ego’s perceptions of both ego’s and alter’s moods during social interactions. Because each diary keeper judged her own mood and the mood of each contacted person at the same time, it is reasonable to assume that the two mood scores are strongly associated. To take potential “raters’ effects” into account, more precisely, we took the essential step to include the diary keepers’ mood scores when in contact with specific members in the model, which were most

influential, as shown in Table 1.

We tried various analytic strategies and model selections before achieving our final models. For instance, some network members were dropped out of the analysis mainly because they lacked any neighbors separated by one or two degrees. We included the current covariates in the final mixed-effects model by standard variable selection procedures, which also took into account how factors on the individual, tie, and contact levels might be linked to how diary keepers rated the mood scores. The conclusions from our analyses varied slightly when we used somewhat different criteria to choose network members and covariates for the modeling.

Among members' characteristics, age and gender turned out not to be significant in the exploratory data analysis. Since more than 65% of the members were female, however, we kept gender in the final model for adjustment. Because tie strength and the relationship between the member and the diary keeper were believed to be influential factors, furthermore, we retained the relationship in the model even though it lacked strong significance. Compared to other factors, contact features were supposed to be more influential. We have included contact mode, purpose, duration, and diary keeper's own mood during each contact, after excluding variables that were not statistically significant, such as when and where the contact took place. In addition to these covariates, it is possible that some other observed or unobserved factors also might be relevant to diary keepers' features but not included in the model. To reduce such potential impacts on the estimated coefficients, we added random components representing the variation among diary keepers. Specifically, our mixed-effects model showed that the average mood of network members linked by only one step contributed significantly to a given member's mood, as did the average mood of those members separated by two degrees. Those members who were separated by three degrees or more did not show a clear association.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

We reached the findings by a special longitudinal design that followed up 133 participants with online contact diaries for seven months. With the advantage of a web application, the ClickDiary program offers a friendly interface to collect detailed information about an egocentric network, the estimated relationships among all network members, and the mood status of both parties during each contact. To minimize recall bias, the program allowed diary keepers to record only the main contacts that occurred with the same person only within the past 24 hours. That is, ClickDiary encouraged participants to enter the information about their daily contacts as soon as possible.

While there is no gold standard to evaluate the extent to which these diary entries are valid, it would help to cross check the basic entries against similar studies of contact diaries. Participants in this study, for example, recorded an average of 12 contacts per day, which was very close to, although slightly fewer than, the number of contacts in compatible social surveys and more conventional paper-pencil diary studies in China, Hong Kong, and Taiwan.^{21 22} As discussed earlier, one major difference of the ClickDiary lies in the stricter criteria about what counts as a contact. Most previous diary studies either included all fleeting contacts or used “two to three words in exchange” as the minimum requirement for enlisting contacts, whereas the ClickDiary asked for only the contacts that involved at least “three sentences.” This last unique feature may also justify the validity of the basic profile of the findings from the ClickDiary.

By cross-checking the mutual ratings of a network member’s mood, we were able to evaluate the judgement errors from the way diary keepers rated how others felt during a specific contact. The same rare data also enabled us to evaluate another major concern: how accurately the diary keepers judged the strength of ties among their network members. As in large probability sampling surveys on self-reported

egocentric networks, which always involved a risk of informant inaccuracy, we asked the focal persons to judge the strength of ties among those surrounding them. Unlike those surveys where the respondents rated the degree of acquaintanceship among a small number of confidants²³⁻²⁵, however, our diary keepers tended to have a tough task, because their contact networks usually stretched far and beyond such core networks. As a result, they often needed to estimate how well any two individuals on their contact lists knew each other even though they did not know either individual well enough in the first place. Although some diary keepers were likely to report “don’t know” under such a circumstance, it remains critical to seek a validity criterion to cross-check their responses.

Another option for conducting a concordant pairs analysis would be to use an intraclass correlation coefficient or Kappa coefficient. Since concordant pairs accounted for only 49.1% of all mood rating pairs, however, and only 74 of all diary keepers’ contact persons also rated their own moods for the same contacts, the agreement is poor, with a small weighted Kappa value of about 0.096. Instead of relying solely on a summary index of agreement, we further investigated the disagreement structure. Given that among the 50.9% cases that showed disagreement, 43.9% revealed only a one-category difference, we assume that about 93% of these score differences (between the moods rated by diary keeper and network member on the identical contact) would fall between -1 and 1, and 7% of the differences would be 2. The cross-checking results provided us an opportunity for performing a sensitivity analysis to assess robustness of our main findings.

Although we were able to validate the alter-to-alter ties of constructed contact networks from these 74 diary keepers, nonetheless, it is noteworthy that these data may not represent all the ties among the 12,070 members in 133 contact networks. Thus, constructing “complete contact networks” out of the diary keepers’ evaluations

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

on how their network members were tied to each other remains a clear limitation. Were all of these network members also involved in keeping a ClickDiary, the estimated ties among them could have been verified by the extent to which they actually contacted one another during the same study period. The task of asking 2,070 network members to record every contact they made for seven months, however, would have been too costly, enormous, and hard to imagine.

As a more realistic, alternative strategy, our design of constructing “proxy” complete contact networks by relying on 133 diary keepers has facilitated a rare analysis of mood correlations among members in different positions in egocentric networks. Such an approach could be further justified if other conditions also satisfied the assumption that two linked members (that is, any two alters who knew each other according to diary keeper’s judgement) indeed made contacts with each other during the study period. To check this assumption, we have also analyzed our subset data, which showed that any two members in a contact network who knew each other had a 90.5% chance to contact each other at least once during the seven months. If the chance to contact each other turned out to be lower in some contact networks, however, we would have overestimated network members’ effect on concurrent mood. In addition, other unmeasured external factors also might affect the mood scores. We treated such unmeasured effects as random variations among different persons.

From the model estimates shown in Table 1, we found that diary keepers tended to rate the contact persons with higher mood scores when the contacts were face-to-face or lasted for a longer period of time. According to the seminal study on “emotional contagion,” which we outlined earlier, the underlying mechanism could be facial mimicry during everyday encounters.⁸ When people had a chance to contact each other for a longer time or in person, their concurrent mood would have been more obvious.

In addition, the tendency of mood spread also could have been linked to personality traits. Extraversion, for example, could induce positive mood, and neuroticism could help predict negative mood.²⁶ Because our study only measured diary keepers' Big-5 personality traits, however, we could not examine how such personality traits influence concurrent mood among the contact persons. In other words, we asked diary keepers to judge their own Big-5 personality traits, but not those of the contact persons, who were the main actors of the study. We assume that diary keepers' Big-5 traits would be correlated with their own mood and that these traits would have directly or indirectly affected how they perceived the mood of their contact persons (network members). To adjust for the effects due to various characteristics of these diary keepers, we included relevant covariates and random components in the mixed-effects models. The models indeed showed that the covariate of diary keeper's mood with a contact person had a very large effect (.743) on the contact person's mood score.

To address the issue of potential response bias from diary keepers, we have further polished our models. For instance, diary keepers' dispositional mood would have strong effects on their own mood scores and those of contacted persons (or network members). The network members' mood scores would be affected by the tie strength and relationships between the members and diary keepers. Diary keepers' ratings also may vary by contact attributes, such as contact mode and duration of each one-on-one contact. To take these potential effects into account, we included covariates of these tie and contact factors in the mixed-effects model. To account for the various effects among the diary keepers on network members' mood scores that are not fully adjusted by the covariates in the model, we further added random components to increase the accuracy of the estimates.

Even though we were able to demonstrate that two members connecting with

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

each other had a high probability of making actual contact during the study period, we do not know exactly when the contact occurred. Based on the diary records, we do know the date when a diary keeper recorded a contact with a network member, which gives us a proxy for the timing of the contact. There was no information, however, indicating when a pair of network members actually contacted each other. Without information about the exact timing of each contact among the network members, we are unable to infer the occurrence of contagion or diffusion of personal mood within the network. In other words, our observational data do not support claims of mood contagion.

It would have been ideal to model how network members' moods associate with one another, if the temporal observations of the contacts among members had been more complete. Diary keepers, however, rated and recorded a member's mood only when they actually contacted the member, which occurred only eight times, on average, for each member during the study period. The limited number of observations of the network members' mood thus prevents us from directly analyzing mood spread within a network. Under the circumstances, we were restricted to examine concurrent mood among neighboring members, with the assumption that a pair of acquaintances had contacted each other at least once during the study period of seven months.

We were able to identify about 2.12% of the network members who had contacts with multiple diary keepers. We believe that the real percentage would be small, even though that percentage might have been underestimated, because we cannot be sure exactly how many network members actually came into contact with multiple diary keepers. For one thing, two diary keepers may have assigned different names for identical persons, or know them by different names. That circumstance would post another limit, but the effect on our major results would be minimal.

Furthermore, with limited information about the contacted persons' personal background, our model cannot fully adjust for the effects of homophily,²⁷ most notably the similarities in personality traits, as well as other relevant risk factors. In particular, some pairs of daily contacts tend to be those who resemble each other in that they systematically rate their own and other people's moods in a similar manner. After conducting a detailed analysis of the contact networks and cross-checking both parties' reports on personal moods involved in the identical contacts from a portion of the network data, however, we do observe concurrent mood in the first-order and second-order social ties. We also cannot claim any causal effects or any direction of association between ego and alters. Our study only can show network autocorrelation of personal mood among the network members.

In our study we retrieved contact records from 133 eligible diary keepers. During the study period, 259 other volunteers also registered but turned out to be ineligible because they failed to comply with the requirements of the diary keeping. About 60% of these ineligible volunteers were college students recruited from several classes, who quitted after a few tries, while many others only visited the platform once. To align with the routine practices of empirical studies, we have treated these volunteers as ineligible or "not applicable" cases, whose diary entries were largely incomplete or too scarcely completed to qualify for any network analysis.

Collecting diary data in this prospective study was not an easy task, because diary keeping has proved to be highly demanding for many participants, even with financial incentives. Such a heavy burden prevented some participants from recording online diaries as required, such as keeping diaries at least three times a week and at least 10 times a month. To have a better grasp of interpersonal contacts in everyday life, we have actually lowered our requirements for qualifications by including participants who had recorded online diaries for at least 30 days during the study

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

period of 7 months. To achieve the minimum requirements for statistical analysis within an egocentric contact network, we also included only those who had contacted at least 30 unique individuals.

In our mixed-effects model, we removed network members who had either no one-degree separation neighboring members or no two-degree neighbors. There was no extra information available to impute from neighboring members. We also excluded strangers for three reasons, as described earlier. Mainly, interactions with strangers carry very different implications in studies of network diffusion. Conceptually, strangers are by default not part of one’s personal networks. Even though we asked participants to record contacts with all individuals, it was actually unusual or unnatural for participants to judge how a stranger was connected with their network members, which would require a somewhat different research framework and analytic strategies. It would be intriguing to explore, in some extended studies, whether and how interactions with strangers would bring about somewhat unique patterns of emotional contagion.

Moreover, the subjects who participated in the ClickDiary study volunteered without a strict sampling procedure. The resulting sample of diary keepers is thus skewed towards female, younger, and better-educated subjects. As common in other diary studies that rely on a small sample of subjects, however, the main goal of our study was not using a representative sample to make an inference to the general population.¹¹ Rather, we used the detailed information about all contacts and ties to build 133 sophisticated complete contact networks, some of which intertwined with one another, which allowed us to examine how personal mood may occur concurrently in everyday life.

Conclusions

In line with earlier studies about how emotions and moods emerge concurrently among network members, we aim to make a substantive contribution to the literature by extending the investigation to the mood averaged from a series of contacts between two individuals in everyday life. While the literature has focused on how emotions and moods transmit at the tie level, our study relies on a bottom-up approach that first scrutinizes how such moods may vary at the contact level before aggregating the mood scores for each pair of a diary keeper and a contact person. We achieved this approach by collecting data with an improved version of contact diaries.

In addition to recording key contact features, as well as how each contact person was linked to the diary keeper, our participants also judged each contact person's mood during each specific contact and estimated how well each pair of contact persons knew each other. Not only does such a version of contact diaries yield all ego-alter ties in an egocentric network, but it also generates nearly all alter-alter ties, which essentially enable us to construct comprehensive network structures surrounding each diary keeper. This methodological innovation, in turn, enhances our efforts to make the substantive contribution to the literature about social networking and emotional contagion.

As with most other social network studies, it is relatively easy to collect empirical data about the ties between a focal person and those surrounding him/her (or "ego-alter ties"), which are key indicators to understand the structure of an egocentric network.²⁸ It becomes highly challenging, however, to collect helpful information about, or reconstruct from any sources, the relationships among network members, which allows researchers to analyze the structures of a complete network. One convenient and flexible design in this study relied on some incentives and the sampling strategy to help diary keepers evaluate and confirm how well any pairs of

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

their network members knew each other. In particular, our system assigned an “absent tie” as the default value of the alter-alter tie (meaning the pair did not know each other), which was the case in about 78.4% of all alter-alter ties. When rating these ties, the diary keeper only needed to either confirm such an absent tie or change the option to either “knew each other well” or “knew each other, but not well.” With a median of 76 alters per ego, an average diary keeper managed to evaluate the strength of $76 \times 75/2 = 2,850$ alter pairs within the study period. Being better motivated to report and confirm such alter-alter ties, as a result, diary keepers in this study completed and verified about 99.97% of all ties, which allowed us to analyze concurrent mood among nearly all network members in egocentric networks.

Using special study designs in ClickDiary, we have been able to cross-check both the network members’ moods and the tie strength among members by matching the diary keepers’ estimates and some of the network members’ own ratings. Future studies could make the best use of all network members’ own reports to reconfirm the strength of ties with one another in complete networks. Such ultimate validity criteria would further verify, in a more comprehensive manner, how accurately diary keepers had judged the ties among the members in their personal networks, even though the distributions of such estimates were similar to those of previous paper-pencil diary studies. Most notably, our findings imply that similar personal mood can occur simultaneously, to varying extents, among the friends, relatives, and other acquaintances clustered around different locations within personal networks. Applying the core concepts of network diffusion and richly designed contact-by-contact data to the inquiries about personal well-being, the current study sheds new light on how social network perspectives can help explain the ways individuals express their personal moods concurrently during social interactions in everyday life.

Acknowledgments

This research was supported by a grant from the Academia Sinica, Taiwan (AS-103-TP-C03). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We would like to thank Nan Lin for valuable comments and Jie-Yu Sung for administrative support.

Data sharing statement

The data (<https://figshare.com/s/f2223e95335f5dac44df>; DOI: 10.6084/m9.figshare.5208703) and the codes (<https://figshare.com/s/f7247dbcf3fade138859>; DOI:10.6084/m9.figshare.5208706) used in this study are all public available from figshare.com.

Conflicts of Interest

The authors declare that we do not have any competing interests related to this study.

Contributors

JSH and YCF proposed the research project to and obtained funding from Academia Sinica, Taiwan. TCC, TJY, and JSH designed the study. TCC, TJY, THH, and JSH were involved in data collection. TJY, THH, and JSH provided statistical expertise. THH analyzed the data. TCC wrote the first draft of the manuscript. YCF and JSH revised and reorganized the manuscript extensively. All authors contributed to the interpreting the results and critically revising the manuscript for important intellectual content and approved the final version of the manuscript.

References

1. Hatfield E, Cacioppo JT, Rapson RL. Emotional Contagion. *Current Directions in Psychological Science* 1993;2(3):96-100. doi: 10.1111/1467-8721.ep10770953

2. Hill EM, Griffiths FE, House T. Spreading of healthy mood in adolescent social networks. *P Roy Soc B-Biol Sci* 2015;282(1813) doi: ARTN 20151180 10.1098/rspb.2015.1180

3. Hill AL, Rand DG, Nowak MA, et al. Emotions as infectious diseases in a large social network: the SISa model. *P Roy Soc B-Biol Sci* 2010;277(1701):3827-35. doi: 10.1098/rspb.2010.1217

4. Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *Bmj* 2008;337:a2338. doi: 10.1136/bmj.a2338

5. Rosenquist JN, Fowler JH, Christakis NA. Social network determinants of depression. *Mol Psychiatr* 2011;16(3):273-81.

6. Centola D. The Spread of Behavior in an Online Social Network Experiment. *Science* 2010;329(5996):1194-97.

7. Coviello L, Sohn Y, Kramer AD, et al. Detecting emotional contagion in massive social networks. *PloS one* 2014;9(3):e90315. doi: 10.1371/journal.pone.0090315

8. Hatfield E, Cacioppo JT, Rapson RL. Emotional Contagion. Cambridge: Cambridge University Press 1994.

9. Deng H, Hu P. Matching Your Face or Appraising the Situation: Two Paths to Emotional Contagion. *Front Psychol* 2018;8 doi: ArtN 2278 10.3389/Fpsyg.2017.02278

10. Dezechache G, Conty L, Chadwick M, et al. Evidence for unintentional emotional contagion beyond dyads. *PloS one* 2013;8(6):e67371. doi: 10.1371/journal.pone.0067371

11. Kramer AD, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America* 2014;111(24):8788-90. doi: 10.1073/pnas.1320040111

12. Shalizi CR, Thomas AC. Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. *Sociological methods & research* 2011;40(2):211-39. doi: 10.1177/0049124111404820

13. Fu YC, Ho HC, Chen HM. Weak ties and contact initiation in everyday life: Exploring contextual variations from contact diaries. *Social Networks*

- 2013;35(3):279-87.
14. Kahneman D, Krueger AB, Schkade DA, et al. A survey method for characterizing daily life experience: The day reconstruction method. *Science* 2004;306(5702):1776-80. doi: DOI 10.1126/science.1103572
 15. Fu YC. Contact diaries: Building archives of actual and comprehensive personal networks. *Field Methods* 2007;19(2):194-217. doi: 10.1177/1525822x06298590
 16. Chan TC, Yen TJ, Fu YC, et al. ClickDiary: Online Tracking of Health Behaviors and Mood. *Journal of medical Internet research* 2015;17(6):e147. doi: 10.2196/jmir.4315
 17. Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. *J Res Pers* 2003;37(6):504-28. doi: Doi 10.1016/S0092-6566(03)00046-1
 18. Chang YH, YC F. The evolution of the Taiwan Social Change Survey. In: Japanese General Social Survey, ed. JGSS Symposium 2003: Birth of JGSS and Its Fruit: Social Surveys in Different Countries and Areas and JGSS. East Osaka City: Institute of Regional Studies, Osaka University of Commerce, 2004:149-60.
 19. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. *Proceedings Biological sciences / The Royal Society* 2014;281(1785):20140268. doi: 10.1098/rspb.2014.0268
 20. nlme: Linear and Nonlinear Mixed Effects Models [program]. R package version 3.1-128 version, 2016.
 21. Fu Y-c. Measuring personal networks with daily contacts: a single-item survey question and the contact diary. *SOCIAL NETWORKS* 2005;27(3):169-86.
 22. Fu YC, Wang DW, Chuang JH. Representative Contact Diaries for Modeling the Spread of Infectious Diseases in Taiwan. *PloS one* 2012;7(10) doi: ARTN e45113
10.1371/journal.pone.0045113
 23. Burt RS. Network items and the general social survey. *Social Networks* 1984;6(4):293-339. doi: [http://dx.doi.org/10.1016/0378-8733\(84\)90007-8](http://dx.doi.org/10.1016/0378-8733(84)90007-8)
 24. Marsden PV. Network Data and Measurement. *Annu Rev Sociol* 1990;16:435-63.
 25. Marsden PV. Survey Methods for Network Data. In: Scott J, Carrington PJ, eds. *The Sage Handbook of social Network Analysis*. London: Sage Publications 2011:370-88.
 26. Gomez R, Cooper A, Gomez A. Susceptibility to positive and negative mood states: test of Eysenck's, Gray's and Newman's theories. *Pers Indiv Differ* 2000;29(2):351-65. doi: Doi 10.1016/S0191-8869(99)00198-1
 27. Bollen J, Goncalves B, van de Leemput I, et al. The happiness paradox: your

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

friends are happier than you. *Epj Data Sci* 2017;6 doi: ARTN 4
10.1140/epjds/s13688-017-0100-1

28. Nice to Know You: Familiarity and Influence in Social Networks. System Sciences
(HICSS), 2013 46th Hawaii International Conference on; 2013 7-10 Jan. 2013.

For peer review only

Figure Legend

Figure 1. The clustering patterns of average personal mood in the complete contact networks of four diary keepers.

Each node represents a person, whose relationship with the diary keeper is displayed with a circle for family members, relatives, and good friends and a square for the others. The frame color of the node reveals the strength of tie to the diary keeper (brown for “know each other very well”; orange for “know each other, but not well”). Node color denotes average mood scores of the persons during the study period, with a color gradient ranging from green, which indicates the worst mood, to red, which indicates the best mood. The figure does not include the diary keeper, who is linked to everyone in respective contact network.

Table 1. Estimates of the effects associated with a network member’s mood score in the mixed-effects models using diary data during May-November 2014.

Variable	Value	Std. Error	t-value	p-value
Male [female]	0.0032	0.0052	0.6236	0.5329
<i>Tie strength with diary keeper</i>				
Knew well	0.0163	0.0062	2.6177	0.0089
[knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0003	0.0097	0.0330	0.9737
Good friend	0.0057	0.0101	0.5634	0.5731
Coworker/trade partner	0.0036	0.0094	0.3887	0.6975
Schoolmate/teacher/student	0.0301	0.0084	3.5603	0.0004
[Other]				
<i>Contact mode (%)</i>				
Face-to-face	0.0277	0.0097	2.8486	0.0044
Voice only	0.0202	0.0136	1.4877	0.1369
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0142	0.0073	-1.9522	0.0509
Daily routine	-0.0205	0.0105	-1.9516	0.0510
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0208	0.0079	2.6152	0.0089
1-4 hours	0.0385	0.0101	3.8141	0.0001
4 hours or more	0.0474	0.0149	3.1866	0.0014
Average mood of diary keeper	0.7427	0.0074	100.9489	0.0000
<i>Average mood of network neighbors</i>				
one degree of separation	0.1326	0.0149	8.8671	0.0000
two degrees of separation	0.0590	0.0191	3.0962	0.0020
All others	-0.0024	0.0167	-0.1460	0.8839

Note: Those listed in brackets are the categories excluded from the model.

Table 2. Pairs of mood ratings from the diary keepers and 74 of their network members who also rated their own moods during the same contacts

		Member's moods rated by the diary keeper			
		(1) Poor	(2) Good	(3) Very good	(4) Excellent
Network member's self-rated moods	(1)	0 (0.0%)	3 (0.1%)	1 (0.0%)	5 (0.2%)
	(2)	5 (0.2%)	42 (1.8%)	203 (8.6%)	76 (3.2%)
	(3)	5 (0.2%)	171 (7.2%)	877 (37.0%)	331 (14.0%)
	(4)	3 (0.1%)	75 (3.2%)	326 (13.8%)	245 (10.3%)

Table 3. Combined effect estimates of the same mixed-effects models fitted with 200 different simulated datasets of mood using a perturbation approach.

Variable	Value	Std. Error	t-value	p-value
Male [female]	0.0027	0.0159	0.1720	0.8635
<i>Tie strength with diary keeper</i>				
Knew well	0.0111	0.0199	0.5551	0.5788
[knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0008	0.0286	0.0266	0.9788
Good friend	0.0056	0.0299	0.1882	0.8507
Coworker/trade partner	0.0013	0.0264	0.0480	0.9617
Schoolmate/teacher/student	0.0233	0.0243	0.9580	0.3381
[Other]				
<i>Contact mode (%)</i>				
Face-to-face	0.0114	0.0306	0.3719	0.7099
Voice only	0.0101	0.0452	0.2236	0.8231
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0163	0.0218	0.7465	0.4554
Daily routine	-0.0123	0.0318	0.3879	0.6981
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0118	0.0257	0.4613	0.6446
1-4 hours	0.0207	0.0314	0.6607	0.5088
4 hours or more	0.0225	0.0455	0.4952	0.6205
Average mood of diary keeper	0.5074	0.0300	16.9135	0.0000
<i>Average mood of network neighbors</i>				
one degree of separation	0.0878	0.0371	2.3643	0.0181
two degrees of separation	0.0789	0.0401	1.9684	0.0490
All others	0.0217	0.0361	0.6026	0.5468

For peer review only

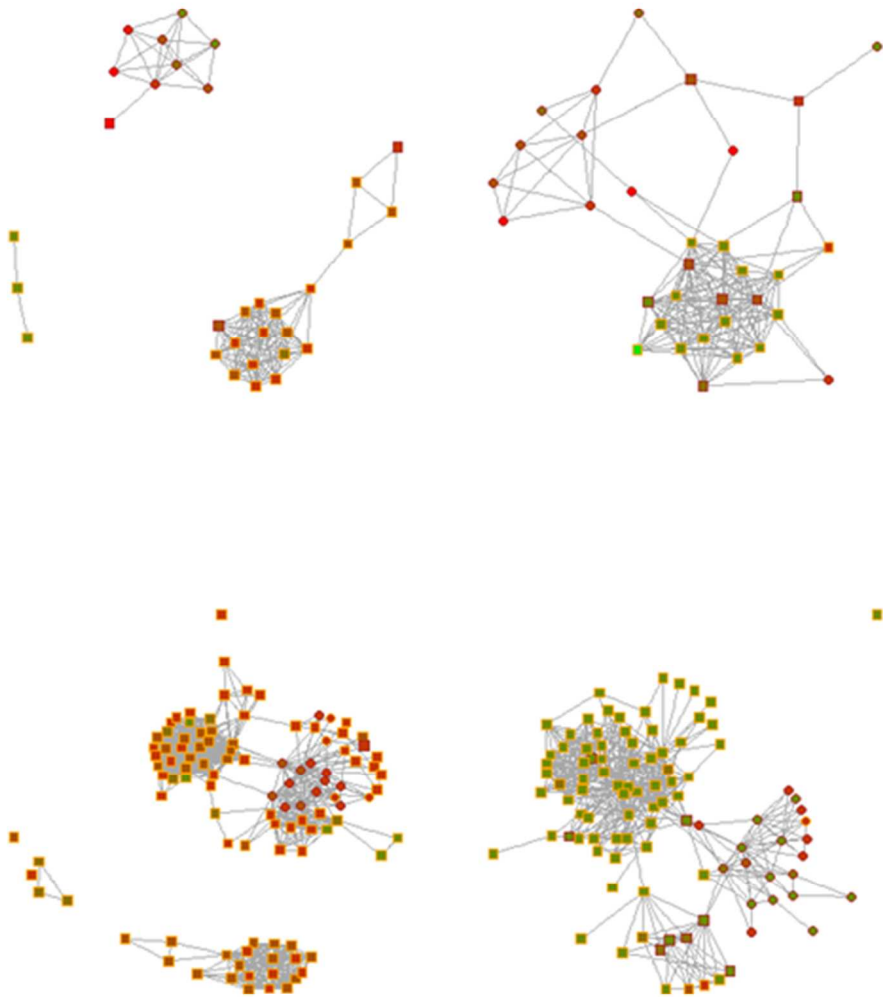


Figure 1. The clustering patterns of average personal mood in the complete contact networks of four diary keepers.

Each node represents a person, whose relationship with the diary keeper is displayed with a circle for family members, relatives, and good friends and a square for the others. The frame color of the node reveals the strength of tie to the diary keeper (brown for “know each other very well”; orange for “know each other, but not well”). Node color denotes average mood scores of the persons during the study period, with a color gradient ranging from green, which indicates the worst mood, to red, which indicates the best mood. The figure does not include the diary keeper, who is linked to everyone in respective contact network.

171x215mm (300 x 300 DPI)

Supplementary

We launched a second wave of data collection with the same incentives to promote the participation in April 2015. To verify the findings from the data collected in the seven months of 2014, we retrieved seven-month contact diaries from April 1 to October 31, 2015. During this second study period, 130 participants completed at least 30 days of contact diaries with at least 30 persons. Note that 54 of them had participated in the May-November 2014 study as well. Very similar to the data collected in the previous study period, these diary keepers recorded 156,892 contacts with 13,539 persons in 2015. The contact lists among the 130 participants averaged 104 persons, with a minimum, median, and maximum of 30, 57 and 1,372, respectively. Diary keepers recorded about 9 contacts per day in 2015, with a minimum of 3, a median of 7, and a maximum of 64. On average, the persons on the contact lists appeared 11 times during the 7 months, ranging from 1 to 211, with a median of 3 times. During this second study period, 8,680 contacts were missing mood scores. As a result, 365 persons were excluded for having no average mood scores. We also excluded network members whom the diary keepers did not know (12%), and the number of contacted persons was reduced to 11,604 with 145,813 contacts for modeling mood variation. Diary keepers confirmed nearly all (99.69%) of the ties between any two persons on the contact lists. The aggregated data showed that about 80.9% of the pairs did not know each other, 9.6% of them knew each other well, and 9.5% just knew each other. For a comparison with the diary data collected during 2014, we listed the model estimates of the 2015 data in Table S1.

Table S1. Estimates of the effects associated with a network member’s mood score in the mixed-effects models using diary data during April-October 2015.

Variables	Value	Std. Error	t-value	p-value
Male [female]	-0.0090	0.0045	-1.9745	0.0484
<i>Acquaintanceship with diary keeper</i>				
Knew well	0.0115	0.0057	2.0085	0.0446
[Knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0017	0.0080	0.2094	0.8342
Good friend	0.0048	0.0091	0.5229	0.6011
Coworker/trade partner	-0.0005	0.0078	-0.0580	0.9538
Schoolmate/teacher/student				
[Other]	0.0057	0.0078	0.7335	0.4633
<i>Contact mode (%)</i>				
Face-to-face	0.0353	0.0083	4.2326	0.0000
Voice only	0.0212	0.0132	1.6103	0.1074
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0265	0.0074	-3.5783	0.0003
Daily routine	-0.0499	0.0107	-4.6710	0.0000
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0093	0.0077	1.2072	0.2274
1-4 hours	0.0309	0.0094	3.2984	0.0010
4 or more hours	0.0359	0.0132	2.7261	0.0064
Average mood of diary keeper	0.6667	0.0074	89.9680	0.0000
<i>Average mood of network neighbors</i>				
One degree of separation	0.2180	0.0155	14.0250	0.0000
Two degrees of separation	0.1156	0.0195	5.9321	0.0000
All others	-0.0295	0.0164	-1.7921	0.0732

Note: Those listed in brackets are the categories excluded from the model.

STROBE statement: checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Checklist	Page #
Title and abstract				
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	Y	p2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	Y	p2
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Y	p5-p6
Objectives	3	State specific objectives, including any prespecified hypotheses	Y	p7
Methods				
Study design	4	Present key elements of study design early in the paper	Y	p8-p9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Y	p9-10
Participants	6	(a) Cohort study ? Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study ?Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross sectional study ?Give the eligibility criteria and the (b) Cohort study ? For matched studies, give matching criteria and number of exposed and unexposed Case-control study ?For matched studies, give matching criteria and the number of controls per case	Y 	

Participants	13*	(b) Give reasons for non-participation at each stage	NA	NA
		(c) Consider use of a flow diagram	NA	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Y	p14-p15
		(b) Indicate number of participants with missing data for each variable of interest	Y	p14-p15
		(c) Cohort study ? Summarise follow-up time (eg average and total amount)	Y	p14-p15
Outcome data	15*	Cohort study ? Report numbers of outcome events or summary measures over time	Y	p14-p15
		Case-control study? Report numbers in each exposure category, or summary measures of exposure	NA	NA
		Cross sectional study? Report numbers of outcome events or summary measures	NA	NA
Main results	16	(a) Report the numbers of individuals at each stage of the study?eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Y	p14-p15
		(b) Give reasons for non-participation at each stage	NA	NA
		(c) Consider use of a flow diagram	NA	NA
Other analyses	17	Report other analyses done?eg analyses of subgroups and interactions, and sensitivity analyses	Y	p16-p18
Discussion				
Key results	18	Summarise key results with reference to study objectives	Y	p19-21
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Y	p20-p23
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other	Y	p20-p23
Generalisability	21	Discuss the generalisability (external validity) of the study results	Y	p23-p25
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Y	p25

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross sectional studies.

BMJ Open

Detecting concurrent mood in daily contact networks: an online participatory cohort study with a diary approach

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-020600.R2
Article Type:	Research
Date Submitted by the Author:	02-May-2018
Complete List of Authors:	Chan, Ta-Chien; Academia Sinica, Research Center for Humanities and Social Sciences Yen, Tso-Jung; Academia Sinica, Institute of Statistical Science Hu, Tsuey-Hwa; Academia Sinica, Institute of Statistical Science Fu, Yang-chih; Academia Sinica, Institute of Sociology Hwang, Jing-Shiang; Academia Sinica, Institute of Statistical Science
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Sociology
Keywords:	contact diary, emotional contagion, social networks, transmission of mood

SCHOLARONE™
Manuscripts

Detecting concurrent mood in daily contact networks: an online participatory cohort study with a diary approach

Ta-Chien Chan^a, Tso-Jung Yen^b, Tsuey-Hwa Hu^b, Yang-chih Fu^{c,1}, Jing-Shiang Hwang^{b,*,1}

^aResearch Center for Humanities and Social Sciences, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei, 115 Taiwan

^bInstitute of Statistical Science, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei, 115 Taiwan

^cInstitute of Sociology, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei, 115 Taiwan

Major classification: Social Sciences; Minor classification: Social Sciences

Short title: mood spread through networks

Keywords: contact diary, emotional contagion, social networks, transmission of mood

*Co-corresponding Author

Jing-Shiang Hwang, E-mail: hwang@sinica.edu.tw.

¹Yang-chih Fu and Jing-Shiang Hwang contributed equally to this paper.

Abstract

Objectives: This paper examines how people express personal mood concurrently with those connected with them by one or two degrees of separation.

Design: Participatory cohort study.

Setting: Online contact diary.

Participants: 133 participants kept online diaries for seven months in 2014, which included 127,455 contacts with 12,070 persons.

Main outcome measures: Diary keepers rated a contacted person's mood during each specific contact, as well as the strength of ties between any pairs of such contacted persons. Such rich information about ties and contacts enable us to construct a complete contact network for each diary keeper, along with the network members' mood and tie strength. We calculate one's overall mood by that person's average mood score during the study period and take the shortest path between any given pair of contacted persons as the degree of separation. We further assume that two connecting persons in a contact network have made contact with each other during the study period, which allows us to examine whether and how personal moods occur concurrently within these contact networks.

Results: Using mixed-effects models while controlling for covariates at individual, tie and contact levels, we show that personal mood score positively and significantly correlates with the average mood among those directly tied to the person. The same effect remains positive and significant for those connected to the person by two degrees, although the effect size is reduced by about one half. The mood of anyone separated by more than two degrees is statistically irrelevant.

Conclusions: Applying network perspectives and rich data at both tie and contact levels to inquiries about subjective well-being, the current study sheds new light on how an improved diary approach can help explain the sophisticated ways in which

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

individuals express their personal moods concurrently during social interactions in
everyday life, contact by contact.

For peer review only

Strengths and limitations of this study

1. The observed online contact diaries allow us to construct 133 complete contact networks with more than 12,000 members, which help uncover how personal moods vary among network members.
2. Some intertwining contact diaries further enable us to cross-examine personal mood and tie strength as rated by both parties of social interactions, solidifying our findings about how a bottom-up social network approach helps reveal concurrent mood in everyday life.
3. Our unique approach of contact diaries yields complex network data and allows us to identify clear patterns of concurrent moods, which are particularly useful, because observational studies typically cannot provide sufficient empirical evidence as to how personal moods may spread to friends' friends through contacts.
4. Without information about the exact timing of each contact among network members, we cannot infer the occurrence of contagion or diffusion of personal mood.
5. With limited information about the contacted persons' personal background, our model cannot fully adjust for the effects of homophily.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

Experiments have demonstrated that emotions tend to spread to others during social interactions,¹ as in the case when diseases, behaviors, and ideas transmit through social networking.^{2 3} Analyses from other large-scale data have further revealed similar spreading patterns of both positive and negative emotions, such as happiness and depression.^{4 5} Like the diffusion of behaviors and attitudes amid face-to-face social networking, emotive sharing and contagion may also cover large social circles through online social contacts by text or voice on Facebook, Twitter, Skype, Google talk, and other social media.^{6 7}

The literature has identified such a common phenomenon, generally known as “emotional contagion,” through various study designs in different research settings. One seminal work argued that emotion or mood can easily be transmitted within social networks, because people are inclined to synchronize others’ facial expressions, voices, postures, movements, and “emotional behaviors” and get feedback from such “mimicry.”⁸ Not only does such a diffusion of emotions and moods occur among those who close to one another, but it also happens from moment to moment in everyday encounters with less known others. More recent experimental studies have shown similar contagion of happiness and fear through facial mimicry by measuring facial electromyography.^{9 10} In addition to face-to-face contacts, furthermore, some posts in social media also can trigger users’ emotional contagion on a massive scale.¹¹

Most existing studies have demonstrated that certain emotions tend to spread from direct contacts over a short period of time, but relatively few have examined whether personal mood may also spread from, or at least occur concurrently among indirect contacts within social networks over a longer period of time. Two recent studies, for example, showed that our mood can be influenced not only by those familiar to us, but also by friends’ friends whom we do not personally know.^{2 3} More

specifically, the spread of personal happiness reaches up to three degrees of separation along social networks, according to one such rare empirical study, the Framingham Heart Study, which analyzed historical data over 20 years.⁴ The longitudinal analyses indicated that an individual was 15.3% more likely to be happy if a directly connected network member (with one degree of separation) was happy; the effect decreased to 9.8% for those separated by two degrees.⁴

As such findings may have inspired studies of mood diffusion, most non-experimental studies are insufficient to examine the actual mechanism of “diffusion,” because the observational data on which they are based offer no such advantages. Thus, it would be particularly difficult to determine and infer from observational data how personal moods spread or diffuse.¹² Even without claiming the causal effects that are essential for explaining spread or diffusion, however, it would be revealing to examine whether and how similar patterns of concurrent mood may exist under other circumstances. While previous longitudinal studies have focused on how individuals’ happiness or depression changes across waves of surveys, an alternative “bottom-up” approach tends to uncover subtle patterns of concurrent mood by examining how one’s mood fluctuates at the micro level, contact by contact.¹³⁻¹⁵ Due to the methodological limitation, it may be inappropriate to pursue inquiries about “mood diffusion or contagion” solely by using observational data, even if one could replicate the decades-long study by creating another huge number of personal networks with records of participants’ relationships and emotions. As an alternative approach, our study, based on a different format, provides a new perspective that would further enrich studies on how personal moods can be linked to one another in contact networks.

By extending the bottom-up approach to social network studies, we aim to examine whether and how personal moods occur concurrently within a contact

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

network, using data collected via an online platform, ClickDiary, over a seven-month period between May 1 and November 30, 2014. The ClickDiary program uses a web-based platform to collect data on participants’ health behaviors and all one-on-one interpersonal contacts in everyday life.¹⁶ The data retrieved for this study have a nested hierarchical structure, including detailed information about 133 participants (or “diary keepers,” who recorded details about their social interactions in daily life), 12,070 contacted persons (or “network members,” including 74 persons who also participated in the same diary keeping platform), as well as 127,455 contacts.

Two features in the ClickDiary platform are central to our research design. First, diary keepers reported their own mood for each one-on-one contact by selecting one of the following score categories that best matched their estimates: (1) poor, (2) good, (3) very good, or (4) excellent; they then evaluated the mood of each network member during the contact. Second, diary keepers, to their best knowledge, rated how well a given pair of network members knew each other. Once a diary keeper confirmed all of their interpersonal ties, we used these two critical features as the backbone to construct a “complete contact network” surrounding that diary keeper, the focal person. Within such a complete contact network, we linked the nodes that represent the network members to each other by interpersonal ties. The overall mood of each network member can be represented by taking the average of the person’s mood scores recorded in the contact diary during the whole study period. Furthermore, we assumed that two connecting persons in a complete contact network have made contact with each other during the study period, which allows us to examine whether and how personal moods occur concurrently within these contact networks. That is, with the assumption, concurrent mood of the linked members could be explained partially as caused by personal contact. We applied a mixed-effects model to analyze

overall mood scores of the members in the 133 complete contact networks to examine whether and how an individual's mood is associated with those within one or two degrees of separation, as well as other members in the complete contact networks.

In sum, our unique approach of online contact diaries is expected to facilitate sophisticated analyses of such phenomena as concurrent patterns of personal mood among network members. Unlike earlier studies that relied on surveys separated by years, our diary approach and analyses can better capture dynamic concurrent moods in everyday life, with complicated network data that help reveal how behaviors and emotions vary contact by contact along with personal ties embedded in different network structures.

Materials and Methods

The ClickDiary Program

The ClickDiary program (<http://cdiary.tw>) uses a web-based platform, written in Chinese, to collect data on participants' daily health behaviors and interpersonal one-on-one contacts.¹⁶ One unique feature of ClickDiary is the friendly interface designed for clicking options on structured diary items via a website or mobile apps, making it easier to record responses whenever it is convenient for participants (supplementary Figures S1, S2).

Public Involvement

Our research team recruited participants from various channels including university students, school teachers and administrative employees, volunteers at health-promotion centers, hospital patients, and community college students, as well as other adults in the general population. When registering in the ClickDiary platform, all diary keepers are required to give online informed consent before starting to keep the diaries. Upon signing up for the program, moreover, participants provided

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

socio-demographic information, including age, gender, place of residence, marital status, and current job. The program also collects participants’ Big Five personality traits (openness, conscientiousness, extraversion, agreeableness, neuroticism),¹⁷ height and weight, perceived health status and happiness, the number (and characteristics) of people contacted during the day, along with a baseline health module that borrowed items from the International Social Survey Programme.¹⁸ We provided an interactive Web chart summarizing the records in each participant’s contact and health diaries. In addition, participants could gain insight from their overall contact patterns by checking their contact network tree we developed (supplementary Figure S3). This current study has been approved by the Institutional Review Board on Humanities and Social Science Research (IRB-HS), Academia Sinica (AS-IRB-HS 02-13022).

Contact diary

When adding a person to the contact list for the first time, the diary keepers provided the person’s background information, such as age and gender, and evaluated several aspects of their relationship with the person, including the duration of acquaintanceship, degree of familiarity, the most frequent mode used for contact (face-to-face, voice only, or text only), contact frequency, and the likelihood of discussing important matters. In this particular diary program, a “contact” refers to one-on-one exchange that involves at least three verbal or written sentences, a definition somewhat narrower than most previous studies using the contact diary approach.¹⁹

The program also asked participants to evaluate the degree to which any pair of persons on the contact list was familiar with each other. Before starting to enter the contact details with any new person, a diary keeper had the chance to estimate, on a

scale from 1 to 3, how well this particular person knew each of randomly selected five other persons already recorded in the diary. After this first step, another random sample of five different persons' names popped up, so that the diary keeper could continue judging all the ties between any pair of persons within the contact network. The process continued, randomly adding five new names to the list at a time, until the diary keeper finished rating the strength of all alter pairs. The design allowed the diary keeper to evaluate all alter-alter ties when it was convenient to do so, thus achieving maximum flexibility and encouraging a higher completion rate. In addition, the diary entries also focused on 11 contact attributes, including when the contact took place, who initiated the contact, the major mode of the contact, the duration and content of the contact, where the diary keeper and the contacted person were during the contact, the extent to which the contact felt beneficial to the diary keeper, the mood of each party during the contact, and whether the contacted person showed any cold symptoms.

Data retrieved for the study

From May 1 to November 30, 2014, 133 residents in Taiwan each completed at least 30 days of contact diaries and recorded one-on-one interpersonal contacts with at least 30 persons. We retrieved the seven-month contact diaries from these 133 participants, with all personal identities removed, which consisted of 141,909 contacts with 16,139 contacted persons. The length of the contact lists varied substantially among the 133 participants, ranging from 30 to 1,399, with a median of 76. The participants recorded an average of 12 contacts a day, with a minimum of 1 and a maximum of 56. The contacted persons varied greatly in terms of how often they appeared in the diaries, ranging from only once to daily (214 times during the seven-month period), even though the average frequency was 8 times.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

To measure each person’s mood, diary keepers selected one from the following score categories that best matched their estimates: (1) poor, (2) good, (3) very good, or (4) excellent. As expected, diary keepers sometimes were unable to estimate a contacted person’s mood and consequently answered “Don’t Know” for the item. The records during the study period showed such “Don’t Know” answers for the mood item appearing in 9,042 contacts. We treated these contacts (with 845 contacted persons) as missing, which reduced the valid number of contacted persons to 15,294.

We further excluded the cases where the diary keepers did not know a contacted person (i.e., the category that accounts for about 21.1% of the relationships) from the subsequent analysis for three reasons. First, interactions with strangers carry very different implications in studies of network diffusion. Even though they had actual contacts with the diary keepers, keeping total strangers in a personal network would have left too many unknown or uncertain links, because the diary keepers were unlikely to judge how well these strangers knew one another. Second, it would have been more difficult for the diary keepers to evaluate strangers’ mood, which also tended to yield missing or less reliable mood rating. Third, probably due to such uncertainties, our different modeling efforts while retaining the ties with these strangers resulted in unreasonable noise to the analysis. The final data for modeling mood spread included 12,070 contacted persons of 133 diary keepers who had made 127,455 contacts during the study period.

Network construction

As a whole, diary keepers were able to confirm nearly all (99.97%) interpersonal ties in terms of the familiarity between any pair of contacted persons. Of these ties, averaged across the 133 diary keepers, 78.4% were considered absent (the pairs did not know each other), 10.6% were strong (knew each other well), and 11.1% were

weak (knew each other, though not well). For each diary keeper, we constructed a complete contact network, in which each alter, or contacted person, represents a node, and two nodes (other than the focal person) are deemed to be connected when the reported alter-alter tie is either strong or weak, rather than absent. In contrast, nodes are not linked when the pair of network members are strangers to each other. Using the shortest path between any given pair of nodes in the network, we calculate “distance” (or “the degree of separation”) between the network members pair by pair. We focus on the links among these network members only, while excluding any links leading to and from the unique node that represented the diary keeper (or “ego-alter ties”). Excluding such ties linked by the diary keepers helps simplify the calculation of “distance.” Otherwise, any two given nodes in the network would have at most two degrees of separation, because all the network members would be directly linked to the diary keepers. Finally, we defined a network member’s overall mood by the average of the mood scores reported by diary keepers during the study period, ranging from 1 to 4.

To illustrate how alter-alter ties cluster into a subset of ego’s complete contact network, we drew a figure, using R package igraph, to display the network patterns in each of four diary keepers’ complete contact networks during the study period (Figure 1). The figure shows that the clustering patterns may differ significantly among diary keepers. Within each of the four complete contact networks, those members who are closer to each other in terms of distance also tend to average similar scores in personal mood during all interpersonal contacts.

Statistical Analysis

To examine whether and how a network member’s mood may be associated with those separated by one or two degrees in each of the 133 complete contact networks,

we first calculated the average mood of those members who were directly connected to a particular member (that is, those connected by one degree of separation), and then we obtained the average mood of those at two degrees of separation and the average mood of all other members. The average mood score of the j^{th} member in the i^{th} complete contact network can be simply obtained by calculating $Y_{ij} = \sum_{k=1}^{K_{ij}} O_{ijk} / K_{ij}$, where O_{ijk} is the mood score of their k^{th} contact given by the i^{th} diary keeper and K_{ij} is the number of contacts during the study period. Let D_{1ij} and D_{2ij} indicate the mean mood scores of those separated by one and two degrees from the j^{th} member in the i^{th} network, respectively. Finally, D_{3ij} measures the average mood scores for those beyond two degrees of separation from the j^{th} network member.

For this study, we applied a mixed-effects model to analyze the relationship between a person's mood score and the average mood scores of those surrounding the person in the network (or "network neighbors"), while controlling for the effects of potential covariates on the person's mood. The model is given as

$$Y_{ij} = (\alpha_0 + a_{0i}) + (\alpha_1 + a_{1i})D_{1ij} + (\alpha_2 + a_{2i})D_{2ij} + (\alpha_3 + a_{3i})D_{3ij} + \sum_{h=1}^p \beta_h X_{hij} + \varepsilon_{ij} \quad \text{for } i = 1, 2, \dots, n \text{ and } j = 1, 2, \dots, n_i,$$

where n is the number of networks; n_i is the size of the i^{th} network; random components a_{li} are assumed to be normally distributed with a mean of 0 and variance σ_l^2 for $l = 0, 1, 2, 3$; and the error term is normally distributed with a mean of 0 and variance τ^2 . Our main interest is to estimate the fixed effects of α_1 , α_2 , and α_3 , which measure whether and how one's mood correlates with those separated by one degree, two degrees, and the others, respectively, within the same complete contact network.

We applied exploratory data analysis tools to identify potential covariates X_{hij} ,

$h = 1, 2, \dots, p$, for the j^{th} member in the i^{th} network. The final $p = 14$ covariates selected in the model are described in the following. In addition to controlling for the possible gender effect of the individual member, the other covariates included in the mixed-effects model measure a variety of features at both tie and contact levels. Two groups of covariates tap the relationships between the diary keepers and their network members. The first is a binary variable indicating that diary keepers knew a network member either well or “not well” (the latter serves as the base category in the model). The second group of covariates distinguishes four types of the network members’ relationships to the diary keepers: family members or relatives (13.2%); good friends (13.5%); coworkers or trade partners (12.7%); and schoolmates, teachers, or students (18.4%). The base category for the group is “others.”

The next group of covariates covers three major contact features: mode, purpose, and duration. Because the main goal of the analysis is to examine how personal moods vary among network members, we sum up these contact features for each member. The modes of contact, for example, may play a key role in determining how well the diary keepers judged others’ mood. To verify such an effect, we include two contact modes as covariates: face-to-face and “voice only,” leaving “text only” out of the model. From all of the contacts between the diary keepers and each particular network member during the seven-month study period, we calculate the proportion (percentage) of each of the three modes for the network member. Because the percentages of all three modes for each network member add up to 1, we keep only the first two modes (face-to-face and voice) in the analysis. Suppose that a member-keeper pair had K times of contacts during the seven-month study period, and among which, the contact modes of “face-to-face” and “voice only” were k_1 and k_2 times, respectively. We defined the two covariates face-to-face as the proportion k_1/K and voice only as k_2/K for the analysis, and treated the other mode of “text only”

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

as the reference category and excluded it from the model. The averaged percentages of face-to-face and voice only contacts among the network members were about 79.8% and 8.2%, respectively.

The model also includes the percentage for each of the two kinds of contact purposes, “work- or school-related” and “daily routine,” which were about 30.4% and 12.4% on average, respectively, while excluding “other purposes.” We further add contact duration into the model, using the percentages for contacts that last 5-59 minutes, 1-4 hours, or more than 4 hours, excluding those lasting under 5 minutes. On average, 47% of the contact durations was 5-59 minutes, 20% was 1-4 hours, and 6% was more than 4 hours.

Because diary keepers reported their own mood as well as the network members’ mood in the same diaries, the two scores are expected to be highly correlated. To take such an effect into account, we further controlled for the most influential covariate that measures the average mood score of the i^{th} diary keeper when contacted with the j^{th} network member during the study period. In other words, the member’s mood score was influenced by the diary keeper’s mood score on the contact. Hence, the covariate of the average mood score of the i^{th} diary keeper when making contact with the j^{th} network member should be influential and must be included in the model for adjustment. We used the lme function from the R package “nlme” to estimate the model parameters.²⁰

Participant involvement

The participants were involved in neither the development of the research questions nor the design of the study. None of the participants were involved in conducting the study nor were they asked to provide input in the writing of this manuscript. There are no plans to disseminate the results of the research to the participants.

Results

Among the original 133 diary keepers, about 80% were female (106/133). Compared to a representative sample of general population, the group tended to be younger and better educated. Percentages of those under age 23 (college students), 23-39, 40-59 and 60 or more were 33.8%, 36.1%, 21.8% and 8.3%, respectively. At least 82.7% had ever gone to college. Like previous contact diary studies, the study participants were overrepresented by females and better educated subpopulations.¹⁵

The high percentage of female diary keepers probably yielded more females on the contact lists (65.5%), although the latter was actually more balanced than the gender distribution among the diary keepers. The age distribution of the 12,070 contacted persons was not very skewed, with age groups of 1-19, 20-29, 30-49, 50-59, and 60 or more accounting for about 8.4%, 21.8%, 31.1%, 18.5%, and 16.9%, respectively. About 53.3% of the contacted persons the diary keepers knew were known very well.

In fitting the mixed-effects model, we excluded about 12% of network members with any covariates missing, mainly those who had either no degree 1 or 2 neighbors. As a result, the following results were obtained from the remaining 10,581 network members with complete covariates. Table 1 shows model estimates of the parameters associated with individual mood variation. On average, a person's mood score increased about 0.13 ($p < .0001$) for every additional point scored in the average mood of fellow members who were directly connected to the person (with one degree of separation). The average mood score of those members linked by two steps also contributed significantly ($p = .002$), with an effect size of about 0.06, to the person. Such contribution diminishes, however, for those members at degree three and beyond (linked by at least three steps).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

These key findings are noteworthy, because they have been adjusted by several highly relevant covariates in the mixed-effects model. Since the diary keepers rated the mood of both parties after a contact, the two scores were often highly correlated. As shown in Table 1, a member’s mood score was strongly associated, with a large effect size of 0.74, with the average of all the diary keepers’ self-rated mood scores when making contact with that specific member. When the diary keeper knew a member well, that member had a better chance of receiving a higher score on personal mood. A network member also tended to receive a higher mood score for a face-to-face contact and a contact that lasted longer. When a contact was about work or school, or was part of the daily routine, the network member’s mood was not as good as that of other contacts, at least based on the diary keeper’s judgment.

The estimated standard deviations of the three random components for one degree of separation, two degrees of separation and beyond two degrees were 0.0015, 0.0372, and 0.0037, respectively. It is clear that the strong mood associations between directly linked members were consistent across the 133 networks. The less strong mood associations between members and neighbors separated by two degrees varied a little large, compared to the estimated fixed effect, among the 133 networks. One possible reason for the large variation among the 133 networks may be the network sizes which ranged from 30 to 1,399, with a median of 76. Members in a small network tended to have relatively few neighbors separated by two degrees. The average mood of smaller number of neighbors may incur large uncertainty on the estimated association effect.

Robustness Checks

To verify the effects of different degrees of separation, we tried a separate analysis of those members at three degrees of separation, using only 8,505 network

members with all valid variates. The alternative analytic design showed little change in both the coefficient estimates of the covariates and the coefficient estimates of one, two, and three degrees of separation (0.13, 0.08, and -0.03, respectively). Thus, it would be unnecessary and unfruitful to further determine whether concurrent mood could occur up to three, four, or more degrees of separation. Unlike previous studies that have indicated how happiness is linked to more indirectly linked network members in the long run,⁴ personal mood in everyday life appears to coexist among those separated by only one or two degrees in one's contact network. To further check how robust it is to represent each network member's overall mood with the average of the person's mood scores during the whole study period, we also fitted the mixed-effects model by replacing the average with the median. The fitted models showed similar results regardless of whether we used the average or the median.

Validation

To further verify the findings from the study, we applied the same mixed-effects model to a similar data set collected in a later study period of seven months from April to October 2015 (supplementary file 1). The results from the supplementary Table S1 indicate that the model estimates and significance levels are similar between the two study periods. In particular, while the average mood score of network members linked by only one step significantly contributed for about 0.22 to a given member's mood score, the coefficient estimate dropped to 0.12 for those separated by two degrees. Like the first study period, one's mood in this second period had little to do with those separated beyond two degrees.

We conducted a subset validation to cross-check the accuracy from a network member's own answer with how well she actually knew each particular member in the diary keeper's network. This critical step of cross-checking alter-to-alter ties was

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

possible because some of our participants joined the ClickDiary study as a group. Among 133 diary keepers, 74 also appeared on the lists of others' "network members." In total, 7,310 individuals appeared on both contact lists of any two diary keepers who also appeared on each other's contact diary. In 6,956 cases, when a first diary keeper believed that the second diary keeper (who happened to be the first diary keeper's network member, thus an alter here, or Alter 1) knew one of these overlapped individuals (another alter, or Alter 2), the second diary keeper (Alter 1) also said she did indeed know this particular person (Alter 2). Likewise, in 22 cases when a first diary keeper said that the second diary keeper did not know a network member they shared, the second diary keeper also confirmed that such a tie was absent. As a result, the diary keepers had judged the alter-to-alter ties among the members in their contact networks with an accuracy rate at about 95.5%, which helps justify our strategy of using diary keeper's contact records to reconstruct part of their contact networks in everyday life.

To check the assumption that two connecting persons in a complete contact network have made contact with each other during the study period, we first identified 5,249 individuals who knew those "network members" (i.e., those on the contact lists) who happened to be diary keepers as well. Then we went back to these diary keepers' own contact diaries and counted how many days each diary keeper actually made contact with each of these 5,249 individuals during the seven-month period. The days of actual contacts between the pairs were well fitted to a negative binomial distribution, with a mean of 25.3 and a size of 0.63. That is, on average these 5,249 pairs of network members contacted with each other on 25.3 days during the study period, but the range of contacts varies widely and is quite skewed (s.d.=32 days).

The results indicate that any two members in a contact network who knew each other had a 90.5% of chance to have at least one contact with each other during the

seven months. The finding further supports our underlying assumption that the network members who knew each other indeed had contact with each other during the seven-month period. Such interpersonal contacts, in turn, facilitated structural circumstances under which personal moods could disperse or emerge in parallel among network members.

Sensitivity analysis

Our online diary platform allowed us to obtain contact information reported by diary keepers, not by the network members themselves, which raises a big concern about how accurately the diary keepers judged a network member's mood during a specific contact. To address this potential issue of the diary keeper's error in judging a network member's mood, we checked the extent to which such judged mood scores were accurate and reliable by matching part of them with the mood scores rated by the network members themselves, as with the case with cross-checking alter-to-alter ties. As a result, we were able to compare how other diary keepers estimated the mood of these 74 network members during 2,368 contacts with how these 74 network members rated their own mood for each of the identical contacts in their own contact diaries.

Counting the original answering categories, concordant pairs accounted for only 49.1% of all mood rating pairs (Table 2). Of the 50.9% pairs that were discordant, however, 43.9% showed only a one-category difference (e.g., while a diary keeper rated a network member's mood as "excellent" during a specific contact, that member rated her own mood during that contact as "very good," which accounted for 14.0% of all 2,368 pairs). Therefore, about 93% of these score differences between the moods rated by diary keeper and network member on the identical contact fell between -1 and 1 .

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

We performed a sensitivity analysis to check for the potential effects of the diary keeper’s judgement error on the parameter estimates of the mixed-effects model, using a perturbation approach with the initial findings of such cross-checking. Specifically, we simulated 200 datasets of mood scores that deviated from the observed scores by 0, 1, and 2 with probabilities 0.5, 0.43, and 0.07, respectively. Fitting the same mixed-effects model with mood scores from the b^{th} simulated dataset, we obtained estimates and standard errors of the l^{th} model parameters, denoted by $\hat{\theta}_l^{(b)}$ and $\hat{s}_l^{(b)}$, respectively.

Taking the influence of judgement error into account, we then estimated the parameter θ_l by the average of these 200 $\hat{\theta}_l^{(b)}$ with the standard error equal to the square root of the sample mean of these 200 $\hat{s}_l^{(b)}$ squared, plus the sample variance of these 200 $\hat{\theta}_l^{(b)}$. It is clear that the augmented judgement errors increased the standard errors of the parameter estimates, as shown in Table 3. Consequently, most of the covariate effects were no more significant. A person’s mood score, however, still significantly increased about 0.088 ($p = .018$) for every additional point scored in the average mood of fellow network members with one degree of separation. The effect size was about 0.079 for those members separated by two degrees, though the strength reduced to marginal significance ($p = .049$). Such effects eventually diminished for those network members at degree three and beyond. The average mood score of diary keepers remained highly influential, with an estimated size of 0.507 ($p < .0001$), while other covariates were irrelevant to how network members’ moods varied.

Discussion

With higher quality data collected through a new method (online diary) than those collected from one-shot survey data, our results about concurrent mood over

interpersonal contacts in daily life are consistent with those of the Framingham Heart Study, which analyzed 20 years of historical data.⁴ Unlike other studies of egocentric networks, however, we analyzed how personal moods of more than 10,000 network members were linked to one another in 133 complete contact networks, based on the information our diary keepers provided.

To be more consistent with how we collected and analyzed the data, we focus on “concurrent mood” in our subsequent discussions. We actually analyzed a hybrid construct derived from the average mood for each ego-alter pair, even though conceptually, we have relied on ego’s perceptions of both ego’s and alter’s moods during social interactions. Because each diary keeper judged her own mood and the mood of each contacted person at the same time, it is reasonable to assume that the two mood scores are strongly associated. To take potential “raters’ effects” into account, more precisely, we took the essential step to include the diary keepers’ mood scores when in contact with specific members in the model, which were most influential, as shown in Table 1.

We tried various analytic strategies and model selections before achieving our final models. For instance, some network members were dropped out of the analysis mainly because they lacked any neighbors separated by one or two degrees. We included the current covariates in the final mixed-effects model by standard variable selection procedures, which also took into account how factors on the individual, tie, and contact levels might be linked to how diary keepers rated the mood scores. The conclusions from our analyses varied slightly when we used somewhat different criteria to choose network members and covariates for the modeling.

Among members’ characteristics, age and gender turned out not to be significant in the exploratory data analysis. Since more than 65% of the members were female, however, we kept gender in the final model for adjustment. Because tie strength and

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

the relationship between the member and the diary keeper were believed to be influential factors, furthermore, we retained the relationship in the model even though it lacked strong significance. Compared to other factors, contact features were supposed to be more influential. We have included contact mode, purpose, duration, and diary keeper’s own mood during each contact, after excluding variables that were not statistically significant, such as when and where the contact took place. In addition to these covariates, it is possible that some other observed or unobserved factors also might be relevant to diary keepers’ features but not included in the model. To reduce such potential impacts on the estimated coefficients, we added random components representing the variation among diary keepers. Specifically, our mixed-effects model showed that the average mood of network members linked by only one step contributed significantly to a given member’s mood, as did the average mood of those members separated by two degrees. Those members who were separated by three degrees or more did not show a clear association.

We reached the findings by a special longitudinal design that followed up 133 participants with online contact diaries for seven months. With the advantage of a web application, the ClickDiary program offers a friendly interface to collect detailed information about an egocentric network, the estimated relationships among all network members, and the mood status of both parties during each contact. To minimize recall bias, the program allowed diary keepers to record only the main contacts that occurred with the same person only within the past 24 hours. That is, ClickDiary encouraged participants to enter the information about their daily contacts as soon as possible.

While there is no gold standard to evaluate the extent to which these diary entries are valid, it would help to cross check the basic entries against similar studies of contact diaries. Participants in this study, for example, recorded an average of 12

contacts per day, which was very close to, although slightly fewer than, the number of contacts in compatible social surveys and more conventional paper-pencil diary studies in China, Hong Kong, and Taiwan.^{21 22} As discussed earlier, one major difference of the ClickDiary lies in the stricter criteria about what counts as a contact. Most previous diary studies either included all fleeting contacts or used “two to three words in exchange” as the minimum requirement for enlisting contacts, whereas the ClickDiary asked for only the contacts that involved at least “three sentences.” This last unique feature may also justify the validity of the basic profile of the findings from the ClickDiary.

By cross-checking the mutual ratings of a network member’s mood, we were able to evaluate the judgement errors from the way diary keepers rated how others felt during a specific contact. The same rare data also enabled us to evaluate another major concern: how accurately the diary keepers judged the strength of ties among their network members. As in large probability sampling surveys on self-reported egocentric networks, which always involved a risk of informant inaccuracy, we asked the focal persons to judge the strength of ties among those surrounding them. Unlike those surveys where the respondents rated the degree of acquaintanceship among a small number of confidants²³⁻²⁵, however, our diary keepers tended to have a tough task, because their contact networks usually stretched far and beyond such core networks. As a result, they often needed to estimate how well any two individuals on their contact lists knew each other even though they did not know either individual well enough in the first place. Although some diary keepers were likely to report “don’t know” under such a circumstance, it remains critical to seek a validity criterion to cross-check their responses.

Another option for conducting a concordant pairs analysis would be to use an intraclass correlation coefficient or Kappa coefficient. Since concordant pairs

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

accounted for only 49.1% of all mood rating pairs, however, and only 74 of all diary keepers' contact persons also rated their own moods for the same contacts, the agreement is poor, with a small weighted Kappa value of about 0.096. Instead of relying solely on a summary index of agreement, we further investigated the disagreement structure. Given that among the 50.9% cases that showed disagreement, 43.9% revealed only a one-category difference, we assume that about 93% of these score differences (between the moods rated by diary keeper and network member on the identical contact) would fall between -1 and 1 , and 7% of the differences would be 2. The cross-checking results provided us an opportunity for performing a sensitivity analysis to assess robustness of our main findings.

As in a recent similar diary study,²⁶ thus, we were able to validate the alter-to-alter ties of constructed contact networks from these 74 diary keepers. It is noteworthy, however, that these data may not represent all the ties among the 12,070 members in 133 contact networks. Thus, constructing "complete contact networks" out of the diary keepers' evaluations on how their network members were tied to each other remains a limitation. Were all of these network members also involved in keeping a ClickDiary, the estimated ties among them could have been verified by the extent to which they actually contacted one another during the same study period. The task of asking 2,070 network members to record every contact they made for seven months, however, would have been too costly, enormous, and unfeasible.

As a more realistic, alternative strategy, our design of constructing "proxy" complete contact networks by relying on 133 diary keepers has facilitated a rare analysis of mood correlations among members in different positions in egocentric networks. Such an approach could be further justified if other conditions also satisfied the assumption that two linked members (that is, any two alters who knew each other according to diary keeper's judgement) indeed made contacts with each other during

the study period. To check this assumption, we have also analyzed our subset data, which showed that any two members in a contact network who knew each other had a 90.5% chance to contact each other at least once during the seven months. If the chance to contact each other turned out to be lower in some contact networks, however, we would have overestimated network members' effect on concurrent mood. In addition, other unmeasured external factors also might affect the mood scores. We treated such unmeasured effects as random variations among different persons.

From the model estimates shown in Table 1, we found that diary keepers tended to rate the contact persons with higher mood scores when the contacts were face-to-face or lasted for a longer period of time. According to the seminal study on "emotional contagion," which we outlined earlier, the underlying mechanism could be facial mimicry during everyday encounters.⁸ When people had a chance to contact each other for a longer time or in person, their concurrent mood would have been more obvious.

In addition, the tendency of mood spread also could have been linked to personality traits. Extraversion, for example, could induce positive mood, and neuroticism could help predict negative mood.²⁷ Even though our study measured diary keepers' Big-5 personality traits, however, we could not examine how such personality traits influence concurrent mood among the contact persons. In other words, we asked diary keepers to judge their own Big-5 personality traits, but not those of the contact persons, who were the main actors of the study. We assume that diary keepers' Big-5 traits would be correlated with their own mood and that these traits would have directly or indirectly affected how they perceived the mood of their contact persons (network members). To adjust for the effects due to various characteristics of these diary keepers, we included relevant covariates and random components in the mixed-effects models. The models indeed showed that the

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

covariate of diary keeper’s mood with a contact person had a very large effect on the contact person’s mood score.

To address the issue of potential response bias from diary keepers, we have further polished our models. For instance, diary keepers’ dispositional mood would have strong effects on their own mood scores and those of contacted persons (or network members). The network members’ mood scores would be affected by the tie strength and relationships between the members and diary keepers. Diary keepers’ ratings also may vary by contact attributes, such as contact mode and duration of each one-on-one contact. To take these potential effects into account, we included covariates of these tie and contact factors in the mixed-effects model. To account for the various effects among the diary keepers on network members’ mood scores that are not fully adjusted by the covariates in the model, we further added random components to increase the accuracy of the estimates.

In this study, we demonstrated that two members connecting with each other had a high probability of making actual contact during the study period. Future studies should further benefit by recording when a pair of network members actually contacted each other. With such information about the exact timing of each contact among the network members, it would be more feasible to identify the direction of contagion or diffusion of personal mood within the network.

It would have been ideal to model how network members’ moods associate with one another, if the temporal observations of the contacts among members had been more complete. Diary keepers, in practice, rated and recorded a member’s mood when they actually contacted the member, which occurred about eight times, on average, for each member during the study period. The limited number of observations of the network members’ mood somewhat prevented us from directly analyzing mood spread within a network. Under the circumstances, we were restricted to examine

concurrent mood among neighboring members, with the assumption that a pair of acquaintances had contacted each other at least once during the study period of seven months.

We were able to identify about 2.12% of the network members who had contacts with multiple diary keepers. We believe that the real percentage would be small, even though that percentage might have been underestimated, because it is nearly impossible in real life to identify exactly how many network members actually came into contact with multiple diary keepers. In addition, two diary keepers may have assigned different names for identical persons, or know them by different names. That circumstance might pose another limitation, but the effect on our major results would be minimal. Furthermore, with the current information about the contacted persons' personal background, our model cannot fully adjust for the effects of homophily,²⁸ most notably the similarities in personality traits, as well as other relevant risk factors. In particular, some pairs of daily contacts tend to be those who resemble each other in that they systematically rate their own and other people's moods in a similar manner.

In our study we retrieved contact records from 133 eligible diary keepers. During the study period, 259 other volunteers also registered but turned out to be ineligible because they failed to comply with the requirements of the diary keeping. About 60% of these ineligible volunteers were college students recruited from several classes, who quitted after a few tries, while many others only visited the platform once. To align with the routine practices of empirical studies, we have treated these volunteers as ineligible or "not applicable" cases, whose diary entries were largely incomplete or too scarcely completed to qualify for any network analysis.

Collecting diary data in this prospective study was not an easy task, because diary keeping has proved to be highly demanding for many participants, even with financial incentives. Such a heavy burden prevented some participants from recording

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

online diaries as required, such as keeping diaries at least three times a week and at least 10 times a month. To have a better grasp of interpersonal contacts in everyday life, we have actually lowered our requirements for qualifications by including participants who had recorded online diaries for at least 30 days during the study period of 7 months. To achieve the minimum requirements for statistical analysis within an egocentric contact network, we also included only those who had contacted at least 30 unique individuals.

In our mixed-effects model, we removed network members who had either no one-degree separation neighboring members or no two-degree neighbors. There was no extra information available to impute from neighboring members. We also excluded strangers for three reasons, as described earlier. Mainly, interactions with strangers carry very different implications in studies of network diffusion. Conceptually, strangers are by default not part of one's personal networks. Even though we asked participants to record contacts with all individuals, it was actually unusual or unnatural for participants to judge how a stranger was connected with their network members, which would require a somewhat different research framework and analytic strategies. It would be intriguing to explore, in some extended studies, whether and how interactions with strangers would bring about somewhat unique patterns of emotional contagion.

Like most other studies with a small sample size, the subjects who participated in the ClickDiary study volunteered without a strict sampling procedure. The resulting sample of diary keepers is thus skewed towards female, younger, and better-educated subjects. As common in other diary studies that rely on a small sample of subjects, however, the main goal of our study was not using a representative sample to make an inference to the general population.¹¹ Rather, we used the detailed information about all contacts and ties to build 133 sophisticated complete contact networks, some of

which intertwined with one another, which allowed us to examine how personal mood may occur concurrently in everyday life. After conducting detailed analyses of the complicated contact networks and cross-checking both parties' reports on personal moods involved in the identical contacts, nonetheless, we do observe clear patterns of concurrent mood in the first- and second-order social ties. Our unique diary approach shows significant network autocorrelation of personal mood among the network members, even though we are unable to claim causal effects or a clear direction of mood association between ego and alters. Our approach, data, and findings are particularly useful, in sum, given that observational studies can hardly provide sufficient empirical evidence as to how personal moods may spread to friends' friends through contacts. Furthermore, although our current observational data do not support claims of mood contagion, continuously improved designs in similar diary approaches could enhance the potential for addressing some of the hard questions about the causal effects of network contagion.

Conclusions

In line with earlier studies about how emotions and moods emerge concurrently among network members, we aim to make a substantive contribution to the literature by extending the investigation to the mood averaged from a series of contacts between two individuals in everyday life. While the literature has focused on how emotions and moods transmit at the tie level, our study relies on a bottom-up approach that first scrutinizes how such moods may vary at the contact level before aggregating the mood scores for each pair of a diary keeper and a contact person. We achieved this approach by collecting data with an improved version of contact diaries.

In addition to recording key contact features, as well as how each contact person was linked to the diary keeper, our participants also judged each contact person's

mood during each specific contact and estimated how well each pair of contact persons knew each other. Not only does such a version of contact diaries yield all ego-alter ties in an egocentric network, but it also generates nearly all alter-alter ties, which essentially enable us to construct comprehensive network structures surrounding each diary keeper. This methodological innovation, in turn, enhances our efforts to make the substantive contribution to the literature about social networking and emotional contagion.

As with most other social network studies, it is relatively easy to collect empirical data about the ties between a focal person and those surrounding him/her (or “ego-alter ties”), which are key indicators to understand the structure of an egocentric network.²⁹ It becomes highly challenging, however, to collect helpful information about, or reconstruct from any sources, the relationships among network members, which allows researchers to analyze the structures of a complete network. One convenient and flexible design in this study relied on some incentives and the sampling strategy to help diary keepers evaluate and confirm how well any pairs of their network members knew each other. In particular, our system assigned an “absent tie” as the default value of the alter-alter tie (meaning the pair did not know each other), which was the case in about 78.4% of all alter-alter ties. When rating these ties, the diary keeper only needed to either confirm such an absent tie or change the option to either “knew each other well” or “knew each other, but not well.” With a median of 76 alters per ego, an average diary keeper managed to evaluate the strength of $76 \times 75/2 = 2,850$ alter pairs within the study period. Being better motivated to report and confirm such alter-alter ties, as a result, diary keepers in this study completed and verified about 99.97% of all ties, which allowed us to analyze concurrent mood among nearly all network members in egocentric networks.

Using special study designs in ClickDiary, we have been able to cross-check

both the network members' moods and the tie strength among members by matching the diary keepers' estimates and some of the network members' own ratings. Future studies could make the best use of all network members' own reports to reconfirm the strength of ties with one another in complete networks. Such ultimate validity criteria would further verify, in a more comprehensive manner, how accurately diary keepers had judged the ties among the members in their personal networks, even though the distributions of such estimates were similar to those of previous paper-pencil diary studies. Most notably, our findings imply that similar personal mood can occur simultaneously, to varying extents, among the friends, relatives, and other acquaintances clustered around different locations within personal networks. Applying the core concepts of network diffusion and richly designed contact-by-contact data to the inquiries about personal well-being, the current study sheds new light on how social network perspectives can help explain the ways individuals express their personal moods concurrently during social interactions in everyday life.

Acknowledgments

This research was supported by a grant from the Academia Sinica, Taiwan (AS-103-TP-C03). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We would like to thank Nan Lin for valuable comments and Jie-Yu Sung for administrative support.

Data sharing statement

The data (<https://figshare.com/s/f2223e95335f5dac44df>; DOI: 10.6084/m9.figshare.5208703) and the codes

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(<https://figshare.com/s/f7247dbcf3fade138859>; DOI:10.6084/m9.figshare.5208706)
used in this study are all public available from figshare.com.

Conflicts of Interest

The authors declare that we do not have any competing interests related to this study.

Contributors

JSH and YCF proposed the research project to and obtained funding from Academia Sinica, Taiwan. TCC, TJY, and JSH designed the study. TCC, TJY, THH, and JSH were involved in data collection. TJY, THH, and JSH provided statistical expertise. THH analyzed the data. TCC wrote the first draft of the manuscript. YCF and JSH revised and reorganized the manuscript extensively. All authors contributed to the interpreting the results and critically revising the manuscript for important intellectual content and approved the final version of the manuscript.

References

1. Hatfield E, Cacioppo JT, Rapson RL. Emotional Contagion. *Current Directions in Psychological Science* 1993;2(3):96-100. doi: 10.1111/1467-8721.ep10770953

2. Hill EM, Griffiths FE, House T. Spreading of healthy mood in adolescent social networks. *P Roy Soc B-Biol Sci* 2015;282(1813) doi: ARTN 20151180 10.1098/rspb.2015.1180

3. Hill AL, Rand DG, Nowak MA, et al. Emotions as infectious diseases in a large social network: the SISa model. *P Roy Soc B-Biol Sci* 2010;277(1701):3827-35. doi: 10.1098/rspb.2010.1217

4. Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *Bmj* 2008;337:a2338. doi: 10.1136/bmj.a2338

5. Rosenquist JN, Fowler JH, Christakis NA. Social network determinants of depression. *Mol Psychiatr* 2011;16(3):273-81.
6. Centola D. The Spread of Behavior in an Online Social Network Experiment. *Science* 2010;329(5996):1194-97.
7. Coviello L, Sohn Y, Kramer AD, et al. Detecting emotional contagion in massive social networks. *PloS one* 2014;9(3):e90315. doi: 10.1371/journal.pone.0090315
8. Hatfield E, Cacioppo JT, Rapson RL. Emotional Contagion. Cambridge: Cambridge University Press 1994.
9. Deng H, Hu P. Matching Your Face or Appraising the Situation: Two Paths to Emotional Contagion. *Front Psychol* 2018;8 doi: Artn 2278 10.3389/Fpsyg.2017.02278
10. Dezechache G, Conty L, Chadwick M, et al. Evidence for unintentional emotional contagion beyond dyads. *PloS one* 2013;8(6):e67371. doi: 10.1371/journal.pone.0067371
11. Kramer AD, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America* 2014;111(24):8788-90. doi: 10.1073/pnas.1320040111
12. Shalizi CR, Thomas AC. Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. *Sociological methods & research* 2011;40(2):211-39. doi: 10.1177/0049124111404820
13. Fu YC, Ho HC, Chen HM. Weak ties and contact initiation in everyday life: Exploring contextual variations from contact diaries. *Social Networks* 2013;35(3):279-87.
14. Kahneman D, Krueger AB, Schkade DA, et al. A survey method for characterizing daily life experience: The day reconstruction method. *Science* 2004;306(5702):1776-80. doi: DOI 10.1126/science.1103572
15. Fu YC. Contact diaries: Building archives of actual and comprehensive personal networks. *Field Methods* 2007;19(2):194-217. doi: 10.1177/1525822x06298590
16. Chan TC, Yen TJ, Fu YC, et al. ClickDiary: Online Tracking of Health Behaviors and Mood. *Journal of medical Internet research* 2015;17(6):e147. doi: 10.2196/jmir.4315
17. Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. *J Res Pers* 2003;37(6):504-28. doi: Doi 10.1016/S0092-6566(03)00046-1
18. Chang YH, YC F. The evolution of the Taiwan Social Change Survey. In: Japanese

General Social Survey, ed. JGSS Symposium 2003: Birth of JGSS and Its Fruit: Social Surveys in Different Countries and Areas and JGSS. East Osaka City: Institute of Regional Studies, Osaka University of Commerce, 2004:149-60.

19. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. *Proceedings Biological sciences / The Royal Society* 2014;281(1785):20140268. doi: 10.1098/rspb.2014.0268

20. nlme: Linear and Nonlinear Mixed Effects Models [program]. R package version 3.1-128 version, 2016.

21. Fu Y-c. Measuring personal networks with daily contacts: a single-item survey question and the contact diary. *Social Networks* 2005;27(3):169-86.

22. Fu Y-c, Wang DW, Chuang JH. Representative Contact Diaries for Modeling the Spread of Infectious Diseases in Taiwan. *PloS one* 2012;7(10) doi: ARTN e4511310.1371/journal.pone.0045113

23. Burt RS. Network items and the general social survey. *Social Networks* 1984;6(4):293-339. doi: [http://dx.doi.org/10.1016/0378-8733\(84\)90007-8](http://dx.doi.org/10.1016/0378-8733(84)90007-8)

24. Marsden PV. Network data and measurement. *Annu Rev Sociol* 1990;16:435-63.

25. Marsden PV. Survey Methods for Network Data. In: Scott J, Carrington PJ, eds. *The Sage Handbook of social Network Analysis*. London: Sage Publications 2011:370-88.

26. Fu Y-c., Velema, T.A., Hwang, J-S. Upward contacts in everyday life: Benefits of reaching hierarchical relations in ego-centered networks. *Social Networks*. 2018;54:266-278.

27. Gomez R, Cooper A, Gomez A. Susceptibility to positive and negative mood states: test of Eysenck's, Gray's and Newman's theories. *Pers Indiv Differ* 2000;29(2):351-65. doi: Doi 10.1016/S0191-8869(99)00198-1

28. Bollen J, Goncalves B, van de Leemput I, et al. The happiness paradox: your friends are happier than you. *Epj Data Sci* 2017;6 doi: ARTN 4 10.1140/epjds/s13688-017-0100-1

29. Kaptein M, Nass C, Parvinen P, Markopoulos P, Nice to know you: Familiarity and influence in social networks. System Sciences (HICSS), 2013 46th Hawaii International Conference on; 2013 7-10 Jan. 2013.

Figure Legend

Figure 1. The clustering patterns of average personal mood in the complete contact networks of four diary keepers.

Each node represents a person, whose relationship with the diary keeper is displayed with a circle for family members, relatives, and good friends and a square for the others. The frame color of the node reveals the strength of tie to the diary keeper (brown for “know each other very well”; orange for “know each other, but not well”). Node color denotes average mood scores of the persons during the study period, with a color gradient ranging from green, which indicates the worst mood, to red, which indicates the best mood. The figure does not include the diary keeper, who is linked to everyone in respective contact network.

Table 1. Estimates of the effects associated with a network member’s mood score in the mixed-effects models using diary data during May-November 2014.

Variable	Value	Std. Error	t-value	p-value
Male [female]	0.0032	0.0052	0.6236	0.5329
<i>Tie strength with diary keeper</i>				
Knew well	0.0163	0.0062	2.6177	0.0089
[knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0003	0.0097	0.0330	0.9737
Good friend	0.0057	0.0101	0.5634	0.5731
Coworker/trade partner	0.0036	0.0094	0.3887	0.6975
Schoolmate/teacher/student	0.0301	0.0084	3.5603	0.0004
[Other]				
<i>Contact mode (%)</i>				
Face-to-face	0.0277	0.0097	2.8486	0.0044
Voice only	0.0202	0.0136	1.4877	0.1369
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0142	0.0073	-1.9522	0.0509
Daily routine	-0.0205	0.0105	-1.9516	0.0510
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0208	0.0079	2.6152	0.0089
1-4 hours	0.0385	0.0101	3.8141	0.0001
4 hours or more	0.0474	0.0149	3.1866	0.0014
Average mood of diary keeper	0.7427	0.0074	100.9489	0.0000
<i>Average mood of network neighbors</i>				
one degree of separation	0.1326	0.0149	8.8671	0.0000
two degrees of separation	0.0590	0.0191	3.0962	0.0020
All others	-0.0024	0.0167	-0.1460	0.8839

Note: Those listed in brackets are the base categories of the models.

Table 2. Pairs of mood ratings from the diary keepers and 74 of their network members who also rated their own moods during the same contacts

		Member's moods rated by the diary keeper			
		(1) Poor	(2) Good	(3) Very good	(4) Excellent
Network member's self-rated moods	(1)	0 (0.0%)	3 (0.1%)	1 (0.0%)	5 (0.2%)
	(2)	5 (0.2%)	42 (1.8%)	203 (8.6%)	76 (3.2%)
	(3)	5 (0.2%)	171 (7.2%)	877 (37.0%)	331 (14.0%)
	(4)	3 (0.1%)	75 (3.2%)	326 (13.8%)	245 (10.3%)

Table 3. Combined effect estimates of the same mixed-effects models fitted with 200 different simulated datasets of mood using a perturbation approach.

Variable	Value	Std. Error	t-value	p-value
Male [female]	0.0027	0.0159	0.1720	0.8635
<i>Tie strength with diary keeper</i>				
Knew well	0.0111	0.0199	0.5551	0.5788
[knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0008	0.0286	0.0266	0.9788
Good friend	0.0056	0.0299	0.1882	0.8507
Coworker/trade partner	0.0013	0.0264	0.0480	0.9617
Schoolmate/teacher/student	0.0233	0.0243	0.9580	0.3381
[Other]				
<i>Contact mode (%)</i>				
Face-to-face	0.0114	0.0306	0.3719	0.7099
Voice only	0.0101	0.0452	0.2236	0.8231
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0163	0.0218	0.7465	0.4554
Daily routine	-0.0123	0.0318	0.3879	0.6981
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0118	0.0257	0.4613	0.6446
1-4 hours	0.0207	0.0314	0.6607	0.5088
4 hours or more	0.0225	0.0455	0.4952	0.6205
Average mood of diary keeper	0.5074	0.0300	16.9135	0.0000
<i>Average mood of network neighbors</i>				
one degree of separation	0.0878	0.0371	2.3643	0.0181
two degrees of separation	0.0789	0.0401	1.9684	0.0490
All others	0.0217	0.0361	0.6026	0.5468

Note: Those listed in brackets are the base categories of the models.

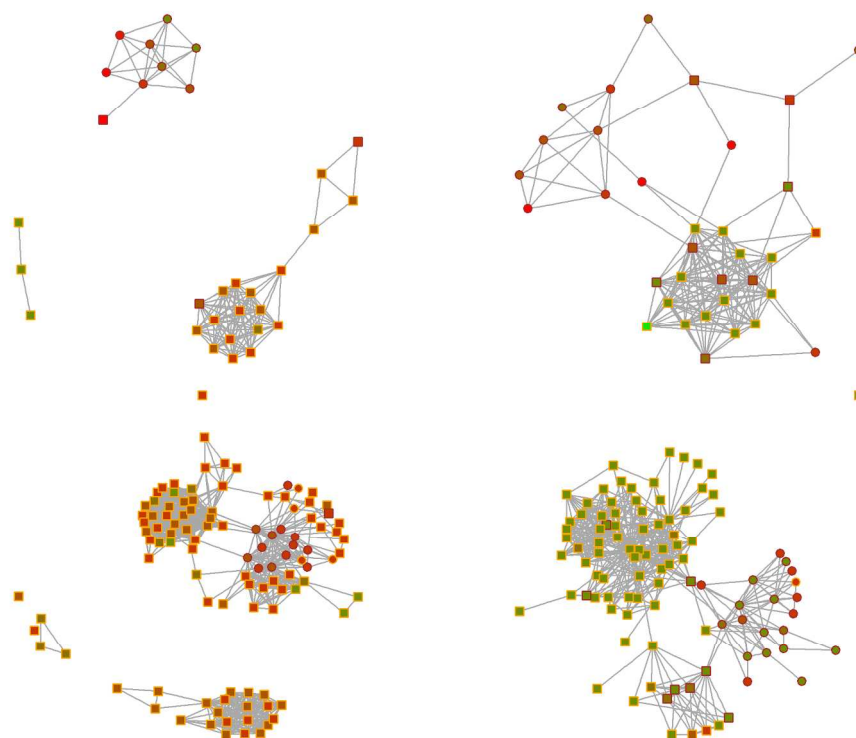


Figure 1. The clustering patterns of average personal mood in the complete contact networks of four diary keepers.

Each node represents a person, whose relationship with the diary keeper is displayed with a circle for family members, relatives, and good friends and a square for the others. The frame color of the node reveals the strength of tie to the diary keeper (brown for "know each other very well"; orange for "know each other, but not well"). Node color denotes average mood scores of the persons during the study period, with a color gradient ranging from green, which indicates the worst mood, to red, which indicates the best mood. The figure does not include the diary keeper, who is linked to everyone in respective contact network.

199x149mm (300 x 300 DPI)

Figure S1. Homepage of the ClickDiary platform



Figure S2. The interface of contact diary

我的日記

- Health diary 點選健康日記
- Contact diary 點選接觸日記
- Feedback 我的記錄圖
- Account 帳號資訊
- Profile info. 個人資料
- Contact list 接觸名單
- Instruction 使用說明
- Other questionnaires 其他計畫連結
- Logout 登出

Rank of scores 排行榜

- Today 本日排行
- This week 本週排行
- This month 本月排行
- Number of contacted persons 人脈排行
- 個人點數

Summary of the personal ranking

接觸日記 My contact diary

Person you have contacted before

請選擇：
☐ 曾經填寫過的接觸對象
☐ 新增接觸對象

※系統自動隱藏前一小時間內填寫過的接觸名單
 【方式二】快速搜尋 溫蒂核 確定

Add new contacted person

Quick search the name of contacted person

The familiarity with the contacted person before this contact

這次接觸以前，您和他/她熟悉程度？

☐ 認識很熟 ☐ 認識但不熟 ☐ 不認識

How likely you will discuss important matters with him/her ?

您可能和他/她討論重要的事情嗎？

☐ 非常可能 ☐ 可能 ☐ 不可能 ☐ 非常不可能

開始點選本次接觸

Please confirm the familiarity between the contacted person and five sampled alters from my contacted list

1. 溫蒂核和下列顯示的人彼此之間熟不熟？
 (這次填寫前從未確認過熟悉程度的人數有 0 人)

溫蒂核可能認識的人	熟悉程度
鄧市環	<input type="radio"/> 認識很熟 <input type="radio"/> 認識但不熟 <input type="radio"/> 不認識

When does this contact occur during the day?

2. 本次主要接觸時段？

☐ 上午 ☐ 中午 ☐ 下午 ☐ 晚上 ☐ 凌晨

Who initiates this contact?

3. 是誰主動接觸？

☐ 自己 ☐ 對方 ☐ 共同 ☐ 其他

How did you contact? (i.e. face-to-face physical contact, face-to-face non-physical contact, by text, by voice)

4. 本次主要接觸方式？

☐ 面對面 (有身體的接觸，如說) ☐ 面對面 (沒有身體的接觸) (手)
☐ 只有文字 (包含簡訊、e-mail、Line) ☐ 語音通話 (包含語音留言)

Where did you contact? (i.e. home, your school or working place, public place, or other places)

5. 接觸的時候您在哪裡？

☐ 現住所 (包括宿舍) ☐ 我的工作場所/學校
☐ 公共場所 ☐ 其他 (如：對方的住所/工作場所/學校)

Where does your contacted person locate? (i.e. northern, central, southern, eastern Taiwan, other places, or unknown)

6. 本次接觸的時候，對方在哪裡？

☐ 北部 ☐ 中部 ☐ 南部 ☐ 東部
☐ 其他地區 ☐ 不知道

How long did this contact last? (i.e. <1 minutes, 1-4 minutes, 5-14 minutes, 15-59 minutes, 1-4 hours, >4 hours)

7. 接觸時間？

☐ 少於1分鐘 ☐ 1-4分 ☐ 5-14分 ☐ 15-59分
☐ 1-4小時 ☐ 大於4小時

What is the main purpose of this contact? (school work/work, leisure activities, medical visits, transportation, social and gossip, daily routine, and others)

8. 主要接觸內容？

☐ 課業/工作 ☐ 休閒活動 ☐ 就醫 ☐ 通勤
☐ 社交與聊天 ☐ 日常作息/家務 ☐ 其他

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

How much do you feel the contact benefitted you? (note: the gain was not related to the emotion)

What is your mood during the contact ?

What is your contacted person's mood during the contact ?

Does the contacted person have the symptoms of the flu ?

9. 接觸過程中，您覺得收穫大不大? (*註:與心情/情緒無關的具體收穫)

☐ 很大☐ 有一點☐ 幾乎沒有☐ 有損失或額外付出

10. 本次接觸的時候，您的心情好壞?

☐ 非常好☐ 很好☐ 還好☐ 不好

11. 對方的整體心情好不好?

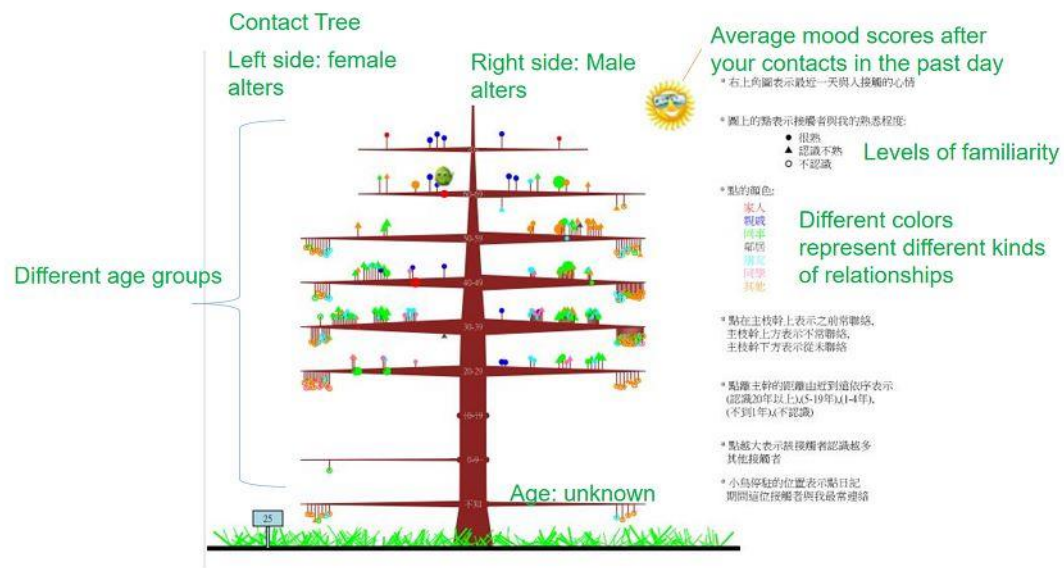
☐ 非常好☐ 很好☐ 還好☐ 不好
☐ 不知道

12. 對方是否有以下症狀?

☐ 沒有感冒症狀☐ 可能有感冒☐ 確定有感冒☐ 可能有登革熱

儲存

Figure S3. Contact tree: summary chart of contacted persons for a participant



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Table S1. Estimates of the effects associated with a network member's mood score in the mixed-effects models using diary data during April-October 2015.

Variables	Value	Std. Error	t-value	p-value
Male [female]	-0.0090	0.0045	-1.9745	0.0484
<i>Acquaintanceship with diary keeper</i>				
Knew well	0.0115	0.0057	2.0085	0.0446
[Knew, not well]				
<i>Relationship with diary keeper</i>				
Family member/relative	0.0017	0.0080	0.2094	0.8342
Good friend	0.0048	0.0091	0.5229	0.6011
Coworker/trade partner	-0.0005	0.0078	-0.0580	0.9538
Schoolmate/teacher/student	0.0057	0.0078	0.7335	0.4633
[Other]				
<i>Contact mode (%)</i>				
Face-to-face	0.0353	0.0083	4.2326	0.0000
Voice only	0.0212	0.0132	1.6103	0.1074
[Text only]				
<i>Contact purpose (%)</i>				
Work/school	-0.0265	0.0074	-3.5783	0.0003
Daily routine	-0.0499	0.0107	-4.6710	0.0000
[Other]				
<i>Contact duration (%)</i>				
[Less than 5 minutes]				
5-59 minutes	0.0093	0.0077	1.2072	0.2274
1-4 hours	0.0309	0.0094	3.2984	0.0010
4 or more hours	0.0359	0.0132	2.7261	0.0064
Average mood of diary keeper	0.6667	0.0074	89.9680	0.0000
<i>Average mood of network neighbors</i>				
One degree of separation	0.2180	0.0155	14.0250	0.0000
Two degrees of separation	0.1156	0.0195	5.9321	0.0000
All others	-0.0295	0.0164	-1.7921	0.0732

Note: Those listed in brackets are the base categories of the models.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary file 1: Analysis for the diary data collected during April-October 2015

We launched a second wave of data collection with the same incentives to promote the participation in April 2015. To verify the findings from the data collected in the seven months of 2014, we retrieved seven-month contact diaries from April 1 to October 31, 2015. During this second study period, 130 participants completed at least 30 days of contact diaries with at least 30 persons. Note that 54 of them had participated in the May-November 2014 study as well. Very similar to the data collected in the previous study period, these diary keepers recorded 156,892 contacts with 13,539 persons in 2015. The contact lists among the 130 participants averaged 104 persons, with a minimum, median, and maximum of 30, 57 and 1,372, respectively. Diary keepers recorded about 9 contacts per day in 2015, with a minimum of 3, a median of 7, and a maximum of 64. On average, the persons on the contact lists appeared 11 times during the 7 months, ranging from 1 to 211, with a median of 3 times. During this second study period, 8,680 contacts were missing mood scores. As a result, 365 persons were excluded for having no average mood scores. We also excluded network members whom the diary keepers did not know (12%), and the number of contacted persons was reduced to 11,604 with 145,813 contacts for modeling mood variation. Diary keepers confirmed nearly all (99.69%) of the ties between any two persons on the contact lists. The aggregated data showed that about 80.9% of the pairs did not know each other, 9.6% of them knew each other well, and 9.5% just knew each other. For a comparison with the diary data collected during 2014, we listed the model estimates of the 2015 data in Table S1.

STROBE statement: checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Checklist	Page #
Title and abstract				
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	Y	p2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	Y	p2
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Y	p5-p6
Objectives	3	State specific objectives, including any prespecified hypotheses	Y	p7
Methods				
Study design	4	Present key elements of study design early in the paper	Y	p8-p9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Y	p9-10
Participants	6	(a) Cohort study ? Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study ?Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross sectional study ?Give the eligibility criteria and the (b) Cohort study ? For matched studies, give matching criteria and number of exposed and unexposed Case-control study ?For matched studies, give matching criteria and the number of controls per case	Y	p8-p10
			NA	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Y	p11-p13
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Y	p11-p13
Bias	9	Describe any efforts to address potential sources of bias	Y	p20-p23
Study size	10	Explain how the study size was arrived at	Y	p9-p10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Y	p11-p13
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	Y	p11-p13
		(b) Describe any methods used to examine subgroups and interactions	Y	p11-p13
		(c) Explain how missing data were addressed	Y	p11-p13
		(d) Cohort study ? If applicable, explain how loss to follow-up was addressed Case-control study ? If applicable, explain how matching of cases and controls was addressed Cross sectional study ? If applicable, describe analytical methods taking account of sampling strategy	Y	p11-p13
		(e) Describe any sensitivity analyses	Y	Supplementary file
Results				
		(a) Report numbers of individuals at each stage of study? eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Y	p14-p15

Participants	13*	(b) Give reasons for non-participation at each stage	NA	NA
		(c) Consider use of a flow diagram	NA	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Y	p14-p15
		(b) Indicate number of participants with missing data for each variable of interest	Y	p14-p15
		(c) Cohort study ? Summarise follow-up time (eg average and total amount)	Y	p14-p15
Outcome data	15*	Cohort study ? Report numbers of outcome events or summary measures over time	Y	p14-p15
		Case-control study? Report numbers in each exposure category, or summary measures of exposure	NA	NA
		Cross sectional study? Report numbers of outcome events or summary measures	NA	NA
Main results	16	(a) Report the numbers of individuals at each stage of the study?eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Y	p14-p15
		(b) Give reasons for non-participation at each stage	NA	NA
		(c) Consider use of a flow diagram	NA	NA
Other analyses	17	Report other analyses done?eg analyses of subgroups and interactions, and sensitivity analyses	Y	p16-p18
Discussion				
Key results	18	Summarise key results with reference to study objectives	Y	p19-21
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Y	p20-p23
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other	Y	p20-p23
Generalisability	21	Discuss the generalisability (external validity) of the study results	Y	p23-p25
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Y	p25

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross sectional studies.