ORIGINAL ARTICLE

Comparing the measurement of different social networks: Peer nominations, online communication, and proximity data

T. J. van Woudenberg[1]* ⓘ, K. E. Bevelander[2], W. J. Burk[1], C. R. Smit[1], L. Buijs[1] and M. Buijzen[1,3]

[1]Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands; [2]Radboud Institute for Health Sciences, Primary and Community Care, Radboud University and Medical Centre Nijmegen, Nijmegen, The Netherlands; [3]Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands
*Corresponding author. Email: t.vanwoudenberg@bsi.ru.nl

Action Editor: Laura Koehly

Abstract

Background: Technological progress has enabled researchers to use new unobtrusive measures of relationships between actors in social network analysis. However, research on how these unobtrusive measures of peer connections relate to traditional sociometric nominations in adolescents is scarce. Therefore, the current study compared traditional peer nominated networks with more unobtrusive measures of peer connections: Communication networks that consist of instant messages in an online social platform and proximity networks based on smartphones' Bluetooth signals that measure peer proximity. The three social network types were compared in their coverage, stability, overlap, and the extent to which the networks exhibit the often observed sex segregation in adolescent social networks.

Method: Two samples were derived from the *MyMovez* project: a longitudinal sample of 444 adolescents who participated in the first three waves of the first year of the project (Y1; 51% male; $M_{age} = 11.29$, $SD_{age} = 1.26$) and a cross-sectional sample of 774 adolescents that participated in fifth wave in the third year (Y3; 48% male; $M_{age} = 10.76$, $SD_{age} = 1.23$). In the project, all participants received a research smartphone and a wrist-worn accelerometer. On the research smartphone, participants received daily questionnaires such as peer nomination questions (i.e., nominated network). In addition, the smartphone automatically scanned for other smartphones via Bluetooth signal every 15 minutes of the day (i.e., proximity network). In the Y3 sample, the research smartphone also had a social platform in which participants could send messages to each other (i.e., communication network).

Results: The results show that nominated networks provided data for the most participants compared to the other two networks, but in these networks, participants had the lowest number of connections with peers. Nominated networks showed to be more stable over time compared to proximity or communication networks. That is, more connections remained the same in nominated networks than in proximity networks over the three waves of Y1. The overlap between the three networks was rather small, indicating that the networks measured different types of connections. Nominated and communication networks were segregated by sex, whereas this was less the case in proximity networks.

Conclusion: The communication and proximity networks seem to be promising unobtrusive measures of peer connections and are less of a burden to the participant compared to a nominated network. However, given the structural differences between the networks and the number of connections per wave, the communication and proximity networks should not be used as direct substitutes for sociometric nominations, and researchers should bear in mind what type of connections they wish to assess.

Keywords: social networks; Bluetooth; nominations; communication; proximity; adolescents

1. Background

In social network analysis, studies have traditionally relied on self-reported nominations by participants (Eagle et al., 2009; Olguín-Olguín & Pentland, 2010; Wasserman, 1994). Usually, the nomination procedure involves asking participants to select or rank peers based on one or multiple questions. For example, "Who are close friends?" (de la Haye et al., 2011) or "Who do you look up to [in your class]?" (Campbell et al., 2008). Participants are either free to nominate as many classmates as they prefer or are limited to a specified number. However, asking participants about their relationships with peers has some disadvantages.

First, a common problem is nonresponse (De Lange et al., 2004). Nonresponse can be caused by a lack of time or motivation of the participant to provide the answers. In social network analysis, nonresponse is especially problematic because each missing nomination brings about an additional gap in the social network. That is, the missing data do not only relate to the participant that nominates but also to the participant who would otherwise be nominated (De Lange et al., 2004).

Second, peer-nomination or self-reporting could lead to social desirability and recall biases (Van de Mortel, 2008). For example, participants could underreport relationships with socially undesirable peers or overestimate interactions with participants who have a strong presence in the social network or with participants they met recently (Eagle et al., 2009).

Third, participants can have different interpretations of the questioned concept (Marin & Hampton, 2007). For example, adolescents might differ in their interpretation of what a friendship is, as indicated by an often observed finding that not all friendships are reciprocated (Hartup, 1996). That is, person A would interpret person B as a friend, but person B does not nominate person A as a friend.

Technological advances have enabled researchers to measure peer relationships without relying on peer nominations. For example, existing data of relationships (e.g., social media messages; Garton et al., 1997), or unobtrusive measures of interactions (proximity based on location data; Cho et al., 2011; Li & Chen, 2009), could be used to infer relationships between participants. However, it is unknown how these new types of measures relate to the gold standard of sociometric nominations in adolescents. Therefore, the current study investigates three types of social networks: *Peer nominated networks*, which are based on self-reported relationships by adolescents, *proximity networks*, which are based on connections between adolescents' Bluetooth devices, and *communication networks*, which are based on connections between senders and receivers of online instant messages. The aim of this study was to describe and compare the three types of social networks in order to identify similarities and differences between the networks and to better understand the various types of interactions in adolescents' social networks.

1.1 Communication network

A promising line of social network research has explored methods to measure peer interactions in mediated communication (Garton et al., 1997), such as phone conversations (e.g., Aiello et al., 2000; Onnela et al., 2007), e-mail (e.g., Ebel et al., 2002; Kossinets, 2006; Wen et al., 2019), or online social platforms such as Facebook (e.g., Del Vicario et al., 2017; González-Bailón et al., 2014; Wilson et al., 2012) and Twitter (e.g., González-Bailón et al., 2014; Takhteyev et al., 2012). Online social networks can be established based on online connections (e.g., friends on Facebook or followers on Twitter) or online interactions (e.g., messages sent via WhatsApp). Using online communication as a measure of connections for social network analysis involves some challenges. For example, there is a substantive difference between online and offline relationships (Salathé & Jones, 2010; Subrahmanyam et al., 2008), due to several features of online communication (e.g., anonymity, asynchronicity, and accessibility; Valkenburg & Peter, 2011). For example, the quality of online and offline friendships differs (Chan & Cheng, 2004), friendships on social networks does not necessarily mean that individuals meet offline (Salathé & Jones, 2010), and

the most important peers online are not necessarily the most important offline peers offline (Subrahmanyam et al., 2008).

In addition, online social networks express different network properties than offline social networks (Wilson et al., 2012). For example, online social networks are less centralized around a few influential individuals who ensure that all peers are more closely connected in a network. As a result, "small-world" properties (the principle that individuals are all linked by short chains of acquaintances) are less present in online than in offline social networks (Wilson et al., 2012).

Wilson et al. (2012) argue that online interactions (e.g., online conversations) are a more accurate representation of meaningful peer connections on social networks than online friendship status. In other words, online interactions with peers (e.g., the number of messages between peers) in a network seem more meaningful than the existence of an online connection (e.g., being a friend on Facebook).

A benefit of using interactions as ties in social networks is that they can have a value based on the number of interactions (valued ties). These weighted networks are more informative because they identify more and less active relationships. Previous research has demonstrated that adding valued ties in nominated networks does not strengthen associations with adolescent substance use (Bauman et al., 2007). However, the networks were weighted based on self-reported quality of the friendship (i.e., feeling of closeness to friend, visits to friend's home, interaction outside school, and parent involvement in the friendship) that might be affected by the previously discussed biases. On the other hand, online interactions are different in the sense that they measure the number of opportunities for contagion or exchange of information. Therefore, weighted online interactions measure different kinds of relationships compared to weighted nominated network and new research questions can be answered. For example, the frequency and content of email communications in an organizational setting can be used to expose top performers in a company that would not have been discovered based on nominated networks (Wen et al., 2019).

1.2 Proximity network

Another promising line of social network research has used spatial and temporal data to infer proximity between network actors (Salathé & Jones, 2010). These types of data produce social networks with a high resolution that systematically registers connections at a determined time interval. For example, GPS data can be used to construct location-based social networks (e.g., Cho et al., 2011; Li & Chen, 2009). Based on GPS data, researchers are able to examine whether people are in the same geographical location. In addition, Bluetooth signals can be used to measure when participants are in close proximity to another Bluetooth device. For example, a Bluetooth beacon could be placed in a fixed location (e.g., Kim et al., 2012), or participants carry around a Bluetooth device in the form of wearable sensors (e.g., Dong et al., 2012; Kim et al., 2012; Polastre et al., 2005) or smartwatches (e.g., Montanari et al., 2017). A study by Salathé & Jones (2010) used wearable sensors to measure close proximity interactions in adolescents on a typical school day. The networks based on these data were high in density and showed a relatively homogenous distribution of interaction partners and interaction time. This means that some participants had very few interactions, while others had many interactions, and some of the interactions were only measured once (>20 seconds), while other interactions lasted more than 60 minutes.

In addition, Eagle et al. (2009) combined the spatial data and the Bluetooth approach by collecting cell tower information and Bluetooth signals of the research telephones of other proximate subjects. In their study, 94 university students and staff members received a research smartphone with an application, that stored location data from cell towers and other proximate subjects from Bluetooth device discovery scans at five-minute intervals. In the study, the author compared the proximity data to self-reported survey data of peer proximity. The analyses showed some overlap between the two types of data but also indicated that the two networks were distinct from

each other. More specifically, most interactions measured with the automated assessment were not reported by the participants and when proximity was reported, the length of the interaction was typically overestimated.

A similar approach has been used in the current study to unobtrusively measure peer proximity in adolescents. We developed a research application, installed on a research smartphone, that measures peer proximity based on the same design as used by Eagle et al. (2009). That is, the research smartphones of participating adolescents scanned for other devices every 15 minutes of the day (between 7:00 and 19:30) in three short bursts. For every smartphone that was detected more than once in a single scan (to control for random passing by), each connection between two participants was stored on the server.

1.3 Approach and hypotheses

Although divergent approaches have been used to measure social interactions and relationships between actors, little is known how these types of networks relate to one another and whether objective network measures could substitute self-reported nominations. Therefore, the current study aimed to compare the three types of social networks that were derived from the same adolescents (i.e., nominated, communication, and proximity) measured in the *MyMovez* project. The networks were compared in terms of response rates, stability, and overlap between networks. Response rates were assessed by identifying how many participants provided data for each type of network, and how many connections were established between peers. The stability was measured by looking at the ratio of connections within each network that were observed across multiple assessments. The overlap between the three networks was assessed in pairwise fashion, by examining the ratio of connections that were present in both networks compared to the connections that were only present in one of the two networks. In addition, adolescents' social networks are known to be segregated by sex and the current study aimed to investigate to what extent the three social networks exhibit this pattern of connections.

No previous study has directly compared the response rates of participants for the different types of networks. Therefore, we based our first hypothesis on the description of participation in a previous study (De Lange et al., 2004) that indicated that nonresponse is a limitation of nominated networks. Therefore, we hypothesized that (H1a) more participants would be included in the proximity network than in the nomination network because no additional effort is required of the participants and therefore less nonresponse. In addition, we anticipated that fewer participants would be included in the communication network because not all participants would make use of the social platform. Therefore, we hypothesized that (H1b) fewer participants are included in the communication network than in the nomination network.

One of the advantages of proximity and communication networks is that they provide a continuous measure of the strength of relationships between network actors (Eagle et al., 2009; Salathé & Jones, 2010; Salathé et al., 2010; Wang et al., 2013). Based on the observations of Eagle et al. (2009), more connections were measured in the proximity network than in the nominated network, we can infer that on average proximity networks produce more data points per participant than nominated networks. In addition, we expected that the proximity network captures those interactions with peers that are in proximity every day, but also those interactions with acquaintances that are only in proximity a small number of times. Therefore, (H2a) we hypothesized that more connections would be measured in the proximity network than in the nominated network. Likewise, participants are not only able to send multiple messages to peers per day (Wilson et al., 2012) but also send a sporadic message to a peer that would not be considered as someone they spend time with during the break. Therefore, (H2b) we hypothesized that more connections would be measured in the proximity network than in the nominated network.

Less is known about the differences in stability between these networks. However, there are some conceptual differences between nominated networks and the proximity network (Eagle

et al., 2009; Salathé et al., 2010). More specifically, we expected that participants would nomi-nate peers with which they generally would interact, irrespective of the observable interactions during a measurement period. On the other hand, in the proximity network, more opportunities for change occur compared to nominated networks. That is, the peers that are in close proxim-ity to the participants will differ more from day to day, or week to week. Therefore, taking two samples of interactions in a measurement period will result in two more different social networks, compared to the nominated networks. Therefore, we hypothesized that (H3) stability is higher in the nominated network than the proximity network.

Lastly, we did not have any specific expectations about the overlap between the different net-works. Therefore, we explored the overlap between the three types of networks, to inform us whether one type of network could be used to substitute another. In addition, this study inves-tigated the often observed sex segregation in the networks. Previous studies have shown that adolescents' social networks based on nominations are segregated by sex (Camarena et al., 1990; McPherson et al., 2001; Mercken et al., 2009). That is, males tend to nominate other males, and females tend to nominate other females more often. We hypothesized that (H4) the nominated network and the communication network were highly segregated by sex because it is more likely that adolescents spend time or communicate with others of the same sex. However, we expected that the proximity network would be less sex segregated because being in close proximity is a less deliberate decision than nominating or talking to peers. Therefore, connections in the proximity network are less bounded by cultural conventions or friendship preferences.

2. Methods

2.1 Participants

The study used data of the *MyMovez* project (Bevelander et al., 2018), which investigated adoles-cents' health behaviors (i.e., nutrition, media use, and physical activity) and their social networks for three years. The first year (data collection waves 1, 2, 3) and the second year (wave 4) marked the first phase of the project in which the health behaviors of adolescents were monitored with-out intervening. The data of the second year could not be included due to the high attrition rate and participants who left the schools. The third year (waves 5, 6, 7) marked the second phase of the project in which four different types of interventions were tested to promote either water consumption or physical activity.

As a result of the different phases of the project, the number of included classrooms and participants in the sample varied between waves and years. In addition, the social platform was added to the research application in the third year. Therefore, the current study used two distinct samples. For both samples, we used the inclusion criterion that at least 60% of the students in each classroom had to participate in the project to obtain representative samples of the social networks within each classroom (Marks et al., 2013).

2.1.1 Y1 sample

The longitudinal Y1 sample contained the first three waves in the first year of the project. Each wave contained a week of measurement with an eight-week time interval in between assessments: February–March 2016 (W1), April–May 2016 (W2), and May–June 2016 (W3). The Y1 sample included 444 participants (51.35% male, $M_{age} = 10.49$ years, $SD_{age} = 1.25$) in 25 classes.

2.1.2 Y3 sample

The cross-sectional Y3 sample contained the data of the fifth wave of the project, which was assessed in the third year (February–March 2018). In this year, new classrooms were added to the project because this wave served as the baseline measure for the project's interventions. The Y3

Table 1. Overview of measured social networks in the two samples.

	Y1			Y3
	Wave 1	Wave 2	Wave 3	Wave 5
Nominated network	Y1 sample	Y1 sample	Y1 sample	Y3 sample
Proximity network	Y1 sample	Y1 sample	Y1 sample	Y3 sample
Communication network	–	–	–	Y3 sample

sample included 774 participants (47.16% male, $M_{age} = 10.96$ years, $SD_{age} = 1.22$) in 43 classes. Due to the set-up of the project, 10 classrooms ($n = 212$; 20.62% of all participants) were part of both the Y1 and Y3 samples. For an overview of the included social networks per sample see Table 1.

2.2 Procedure

Parents/legal guardians of the adolescents in the participating classrooms received information about the project and could enroll their children by providing active consent. For more details see Bevelander et al. (2018). On the first day of the project, participants received instructions on the procedure and the use of the *MyMovez Wearable Lab*: A smartphone with a tailor-made research application and a wrist-worn accelerometer. Before receiving the *Wearable Lab*, participants signed for assent to participate in the project. Subsequently, participants were instructed to wear the accelerometer at all times (the device was water resistant) and take the smartphone with them as much as possible. The smartphone was equipped with a research application by which daily questionnaires were administered (e.g., peer nomination questions). Beginning in wave 5, the app contained a social platform in which the participants could communicate with each other. The smartphone also connected to the accompanying accelerometer and other research smartphones via Bluetooth.

2.3 Measures

2.3.1 Nominated network

Participants received a sociometric question (i.e., "With whom do you hang out during the breaks?") at a random moment of the day during the measurement week. The app provided a list of all classmates as well as lists of students from the other participating classrooms of that school. In addition, participants could search for names in the provided search field and were required to nominate at least one peer from the same grade (self-nominations were impossible). Each nomination resulted in a connection (*edge*) going from the nominee (*ego*) to the nominated participants (*alter*), resulting in a directed network. In this study, nominations outside of the same classroom were excluded.

2.3.2 Communication network

The communication network was derived from the *Social Buzz*, the social platform that was incorporated in the *MyMovez* application in the third year. In the *Social Buzz*, participants could post messages on the message board of the classroom or send private messages to classmates. The latter type of message was used to create the communication network. For every message that a participant (*ego*) sent to a classmate (*alter*), the communication network would assign an edge from the sender to the receiver, again resulting in a directed network. Because multiple messages between participants could be sent during one measurement period, the communication networks
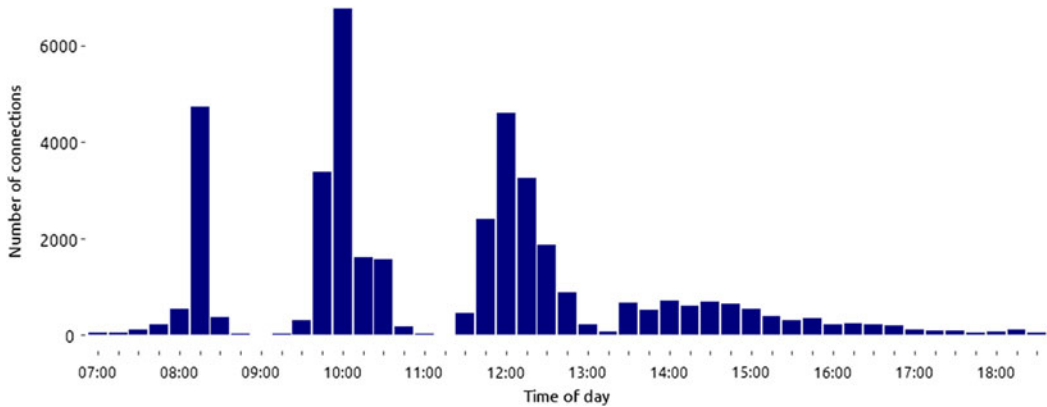
Figure 1. The number of Bluetooth connections per 15 minutes of the day.

could consist of multiple edges between participants. The first and last days of the measurement were excluded because on these days the smartphones were handed out by the researchers or the participants had to hand in the materials. Participants sent between 1 and 221 messages per day ($M = 16.53$, $SD = 28.99$). Similar to the nomination network, only edges within the same classroom were included.

2.3.3 Proximity network

The smartphones scanned for other research smartphones every 15 minutes between 7:00 AM and 7:30 PM, except school hours (when the participants were forced to be in close proximity). One scan existed of three bursts. For every time that two smartphones were within Bluetooth range (approximately 10 meters or 32 feet) for two or more bursts, the proximity network would assign a connection (*edge*) between the two participants. The edge was assigned from the smartphone of the participant (*ego*) to the peer (*alter*). There was a small difference between the timestamps of the research smartphones. Therefore, some scans were registered as within the measurement time on one device, but measured as outside of the measurement time on another device. As a result, not all edges were reciprocal by default. In fact, 92% of all edges in Y1 and 100% of all edges in Y3 were reciprocal. Again, multiple edges per day could be assigned between two participants, and the first and last days of the measurement were excluded, and only nominations within the same classroom were included. As can be seen in Figure 1, most of the edges were accumulated before and after school hours and during the breaks (around 10:00 and 12:00 am).

2.4 Strategy of analysis

Based on the sociometric nominations, communication, and Bluetooth data, social network graphs were created by using the Igraph package (Csardi & Nepusz, 2005) in R (R Core Team, 2019). For each classroom in both samples and each type of network, a graph was created. This resulted in: 25 nominated networks for the Y1 sample and 43 nominated networks for the Y3 sample, 43 communication networks for the Y3 sample, and 25 proximity networks for the Y1 sample and 43 proximity networks for the Y3 sample. All analyses were performed on the classrooms of the Y1 and Y3 samples separately. Figure 2 shows an example of three graphs based on the three different network types for one of the classrooms.

First, response rates were evaluated by looking at the number of participants and number of edges for each of the social networks per wave. The differences between the number of included
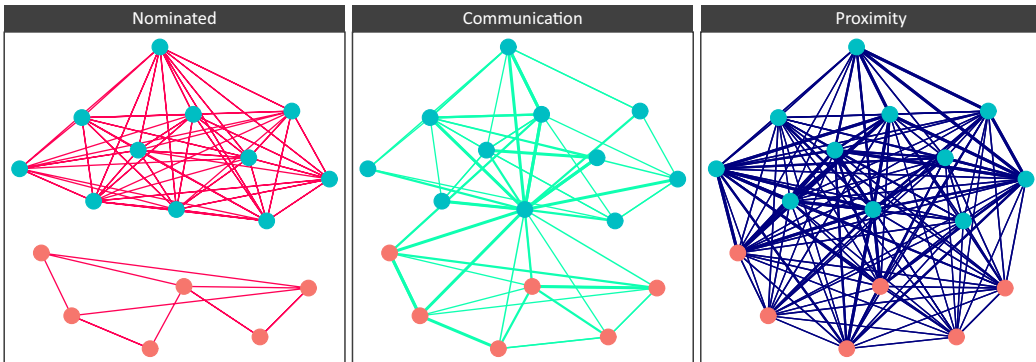
Figure 2. Example graph of one classroom for the three networks.
Note. The dots represent the adolescents (blue is male, pink is female). The lines represent the connections between adolescents. The thickness of the lines indicates the number of edges per wave.

participants and the number of edges were tested by using mixed-effects models (Bates, 2010). In the mixed-effects models, the type of network was the independent variable and number of participants, or the number of edges, per class was the dependent variable. Random intercepts per classroom and per wave (only for the Y1 sample) were added to account for the clustering of data per wave and per class.

Next, the stability of the nominated and the proximity network over the waves in Y1 was assessed by looking at the amount of change in connections between the waves. This was examined with the Jaccard index: the number of edges that is present in both networks at the two time points (intersection), divided by the total number of edges that is present in at least one of the networks at the two time points (union; Hamers, 1989). Previous research has shown that in adolescents' nominated social networks, between 50% and 65% (Jaccard index between 0.50 and 0.65) of the friendship connections are stable over time (Berndt, Hawkins, & Hoyle, 1986; Berndt & Hoyle, 1985), but this can vary based on multiple factors (e.g., time between measures, the number of peers that participants can nominate, or network boundaries). In addition, the Jaccard index was also used to assess the overlap between the different networks in Y1 and Y3. Again, the differences between the networks were tested by using mixed-effects models. The type of network was the independent variable, and the Jaccard index was the dependent variable. Again, random intercepts per classroom and wave were added to the mixed model.

Lastly, we investigated sex segregation in the three networks by looking at the ratio of same-sex edges in the networks and compare this to the expected ratio of same-sex edges if all edges were random. The differences between the ratio of same-sex connections were tested by using mixed-effects models (Bates, 2010). The type of network was the independent variable, the ratio of same-sex nominations was the dependent variable, and random intercepts were added per classroom and wave.

3. Results

3.1 Coverage

The first set of analyses investigated the coverage of the three networks. More specifically, we looked at how many participants and how many connections were included in the social networks in Y1 and Y3, to determine whether the more participants are included in the proximity network than in the nomination network, and fewer participants are included in the communication network than in the nomination network.

3.1.1 Y1 sample

In total, 444 participants, in 25 classes, were included in the longitudinal Y1 sample. Of those participants, 392 (88.29%) filled out the nomination question in at least one of the three waves. For 370 participants (83.33%), proximity data were available for at least one of the waves. On average, more participants filled out the nomination question per classroom per wave ($M = 13.6$, $SD = 5.07$) than participants providing proximity data per classroom per wave ($M = 10.1$, $SD = 5.04$). A mixed-effects model with a random intercept per wave and classroom showed that this difference was statistically significant, $b = -3.47$, $SE = 0.45$, $p < 0.001$. For an overview of the coverage per wave per classroom see Appendix A.

The participants in the Y1 sample provided a total of 7,932 edges in the nominated network and 9,814 in the proximity network. Despite fewer participants providing data for the proximity network, the proximity network provided more edges (3,271, range: 1,692–4,879) than the nominated network (2,644, range: 2,505–2,742) per wave. Looking at the classroom level, a similar number of edges were observed in the nominated ($M = 106$, $SD = 72.5$) and proximity ($M = 128$, $SD = 157$) networks per wave. A mixed-effects model with a random intercept per wave and classroom showed that there was no significant difference between the number of edges in the two networks, $b = 25.09$, $SE = 15.21$, $p = 0.10$. On an individual level, this means that on average participants nominate 6.12 ($SD = 2.57$) other peers per wave and 7.50 ($SD = 7.89$) peers are detected by the smartphones per wave (participants without any edges are included).

3.1.2 Y3 sample

In the cross-sectional sample, 774 participants, in 43 classrooms, were included. Of those participants, 720 (93.02%) filled out the nomination question. For 598 participants (77.26%), proximity data were available and for 518 participants (66.93%) communication data were available. A mixed effects model with a random intercept per classroom showed that on average, more participants filled out the nomination question per classroom ($M = 16.70$, $SD = 6.00$) than participants providing proximity data ($M = 13.90$, $SD = 6.66$; $b = -2.83$, $SE = 0.68$, $p < 0.001$) and communication data ($M = 12.00$, $SD = 6.40$; $b = -4.70$, $SE = 0.68$, $p < 0.001$). For an overview of the response rates per wave per classroom see Appendix B.

The participants in the Y3 sample provided fewer edges in the nominated networks (5,541) than in the proximity networks (16,068) and the communication networks (22,456). A mixed-effects model with a random intercept per classroom showed that per classroom, the participants provided a smaller number of edges in the nominated networks ($M = 129$, $SD = 78.5$) compared to the proximity networks ($M = 374$, $SD = 366$), $b = 244.81$, $SE = 99.02$, $p = 0.015$. Also, the number of edges in the nominated network was smaller than the number of edges in the communication networks ($M = 522$, $SD = 872$), $b = 393.37$, $SE = 99.02$, $p < 0.001$. On the individual level, this means that on average, participants nominate 6.77 ($SD = 2.63$) other peers, send 26.80 ($SD = 41.80$) messages to peers, and 19.70 ($SD = 17.90$) peers are detected by the participants' smartphone (again, participants without any edges are included).

In short, both samples showed that fewer participants were included in the communication and the proximity networks than the nominated networks. However, the nominated networks had the lowest number of edges and the communication networks produced the most edges. Only the difference in the number of edges between the nominated networks and the communication networks was significant. Therefore, we have found no support for the first hypotheses (H1a) that more participants are included in the proximity network than in the nomination network. We did find support for the hypothesis (H1b) that fewer participants are included in the communication network than in the nomination network. Looking at the number of edges in the different networks, we did not find support for the second hypothesis (H2a) that more connections are measured in the proximity network than in the nominated network, but did find support for the hypothesis (H2b) that more connections are measured in the proximity network than in the nominated network.

3.2 Stability

The second analysis investigated the stability of the nominated and the proximity networks over the first three waves. More specifically, we looked at the ratio of edges that were present in two consecutive waves compared to all edges in these two networks. The stability of the social networks could only be assessed in the longitudinal sample (Y1), and therefore the stability of the communication network was not assessed.

In the Y1 sample, some of the classrooms had low response rates in one of the networks in one or more waves (see Appendix A). In order to investigate the stability in representative social networks, two subsamples of the longitudinal sample were created. The first subsample included classrooms in which at least 50% of the participants provided data for both networks in waves 1 and 2. The second subsample included classrooms in which at least 50% of the participants provided data for both networks in waves 2 and 3. In both subsamples, classrooms were excluded in which fewer than eight participants provided information for one of the two types of networks. This resulted in subsamples of 16 and 7 classrooms, respectively. The two subsamples were used in the subsequent tests to examine the stability of both types of networks.

To test for the stability of the nominated and proximity networks, the Jaccard index was used to assess the degree of overlap within each network over time. Overall, more than half of the edges in the nominated network were stable from wave 1 to wave 2 (Jaccard index = 0.56) and from wave 2 to wave 3 (Jaccard index = 0.59). Around one in four edges in the proximity network was stable from wave 1 to wave 2 (Jaccard index = 0.26) and from wave 2 to wave 3 (Jaccard index = 0.28). In addition, the Jaccard indices were calculated per classroom for the two subsamples of the Y1 sample. A mixed-effects model with a random intercept per wave and classroom showed that on average, the Jaccard indices of the nominated network ($M_{W12} = 0.57$; $M_{W23} = 0.56$) were significantly higher than the Jaccard indices of the proximity network ($M_{W12} = 0.22$; $M_{W23} = 0.26$), $t(29.15 = 10.00, p < 0.001$ and $t(8.77) = 4.76, p = 0.001$. This means that in both subsamples, the stability of the nominated networks was similar to previous research (Chan & Poulin, 2007) and more stable over time than the proximity networks. In contrast, the stability of the proximity network was rather low, indicating more variability in the networks between the waves. Therefore, we found support for the third hypothesis (H3) that the nominated network would be more stable over time than the proximity network.

3.3 Overlap

The third set of analyses investigated the overlap between the three social networks. More specifically, we looked at the ratio of connections that is shared with another network compared to all connections in two networks.

3.3.1 Y1 sample

To assess the overlap between the nominated networks and the proximity networks in the Y1 sample, the Jaccard index was used to express the ratio of edges that were present in both networks. Over the three waves, a total of 2,276 edges were present in both networks, 1,670 edges were only present in the nominated networks, and 1,580 edges were only present in the proximity networks (Jaccard index = 0.41). This means that 41% of all edges were overlapping in both nominated networks and proximity networks. Also, the number of exclusive edges of the two networks was comparable, indicating that there was not one network oversampling the other. However, comparing the two networks per wave separately resulted in less overlap between the two networks. As can be seen in Table 2, the Jaccard indices per wave were considerably lower than the 0.41 (the Jaccard index when all waves are combined). Potentially, subsequent waves added more interactions that were not included in both networks in the first wave, complementing the social networks and thus increasing the overlap between the networks.

Table 2. The average number of exclusive and shared affiliated peers per wave and the total number of edges per wave for the nominated and the proximity networks.

	Nominated only	Proximity only	Both networks	Jaccard index
Wave 1				
Affiliating peers per participant	3.98 (*4.06*)	4.08 (*4.06*)	2.92 (*3.18*)	0.27
Total number of edges	1,446	1,480	1,059	
Wave 2				
Affiliating peers per participant	5.03 (*4.42*)	2.80 (*3.29*)	2.36 (*3.08*)	0.23
Total number of edges	1,866	1,040	876	
Wave 3				
Affiliating peers per participant	6.53 (*5.08*)	1.51 (*2.61*)	1.49 (*2.59*)	0.16
Total number of edges	2,187	507	498	

Note: The mean number of affiliating peers per participant is presented, followed by the standard deviation in brackets.

Next, we looked at the number of unique edges per participant in order to investigate the extent to which the networks measure the same affiliating peers. Over the three waves, the participants had on average 6.07 (SD = 6.42) affiliating peers that were observed in both networks, 13.70 (SD = 10.40) affiliating peers that were only observed in the nominated network, and 7.55 (SD = 7.16) affiliating peers that were only observed in the proximity network. This means that per participant, only 22% of the affiliating peers were observed in both networks. Again, this indicates a low overlap between the nominated network and the proximity network. The number of affiliating peers per participant per wave is displayed in Table 2.

3.3.2 Y3 sample
The same Jaccard indices were used to express the overlap between the nominated networks, the proximity networks, and the communications network in the Y3 sample. As can be seen in Table 3, the overlap between the different types of networks was comparable, ranging from 26% to 30% of the edges being present in both networks. Also, a slightly higher Jaccard index was found for the overlap between the nominated networks and the proximity networks in the Y3 sample (0.30) compared to the average Jaccard index of the Y1sample (0.22).

What stands out is that the nominated networks and the proximity networks had a high number of unique edges. In addition, the communication networks did not add many edges to the proximity networks (11% of unique edges), which suggests that participants did not have many online conversations with individuals that they hang out with during the breaks or are in close proximity to each other during the day.

We also looked at the number of unique edges between participants. On average, 2.98 (SD = 4.17, 21.91%) unique edges per participant were observed only in the nominated network, 4.54 (SD = 4.62, 33.38%) unique edges were observed only in the proximity network, 0.38 (SD = 1.09, 2.24%) unique edges were observed only in the communication network, and 1.75 (SD = 2.75, 12.87%) unique edges were observed in all three networks. The remaining edges were observed in a combination of two of the three networks. Again, this indicates a low overlap between the nominated and the proximity networks. Also, the communication network does not add many unique edges (less than 3%) that are not included in one of the other networks.

3.4 Sex segregation in the networks
The last analysis investigated to which extent the communication and the proximity networks express a similar level of sex segregation as is often observed in adolescents' nominated social networks.

Table 3. The number of shared and unique edges for the nomination, communication, and proximity networks in year 3.

Nominated only	Communication only	Proximity only	Both networks	Jaccard index
3,836	1,105	–	1,705	0.26
2,610	–	4,209	2,931	0.30
–	646	4,976	2,164	0.28

3.4.1 Y1 sample

Based on the number of males and females in the classrooms, we calculated the percentage of same-sex nominations in case that the selection of peers was not segregated by sex (50.20%, $SD = 4.65$). This number served as the baseline to investigate whether participants were more likely than random chance to have a connection with a peer of the same sex. In the Y1 sample, 72.80% ($SD = 12.20$) of the connections in the nominated networks per class was between same-sex participants, which is higher than the nonsegregated chance ($t(126.08) = 7.45$, $p < 0.001$). In the proximity networks, only 53.70% ($SD = 18.50$) of the connections was between same-sex participants, which is not higher than expected based on nonsegregation ($t(82.04) = -1.37$, $p = 0.17$). Post hoc comparisons indicated that sex segregation was more prevalent in the nominated network than in the proximity network ($t(136.08) = 7.45$, $p < 0.001$). This indicates that the nominated network shows segregation by sex, while this is not the case in the proximity network.

3.4.2 Y3 sample

The same baseline percentage of same-sex nominations was calculated for the Y3 sample (50.00%, $SD = 5.46$). In this sample, 76.90% ($SD = 12.7$) of the connections in the nominated networks were between same-sex participants, which was higher than expected based on random chance ($t(57.06) = -12.74$, $p < 0.001$). In addition, the percentage of connections in the proximity network (56.70%, $SD = 14.50$) and the communication network ($M = 70.80$, $SD = 23.00$) were also higher than the baseline percentage, $t(52.20) = -2.80$, $p = .007$ and $t(46.71) = -5.75$, $p < 0.001$, respectively. This means that sex segregation was prevalent in all three networks. Post hoc comparisons indicated that the percentage of same-sex connections was higher in the nominated network than in the proximity network ($t(81.04) = 6.82$, $p < 0.001$) but not the communication network ($t(65.31) = 1.51$, $p = 0.13$). These results show that sex segregation appears in all three networks, but that the proximity network shows less sex segregation than the nominated network, supporting the last hypothesis (H4) that the nominated network and the communication network are more segregated by sex compared to the proximity network.

4. Discussion

The current study compared and contrasted three different types of social networks of adolescents: nomination, communication, and proximity. This comparison was made to investigate the possibility of using unobtrusive and automatic measures of social interactions, to assess relationships, which have typically been measured using peer nominations, in order to reduce the burden for participants. The three networks were compared in their response rates, stability, and overlap. In addition, the study investigated to what extent sex segregation was present in the three networks. The results generally showed that these social networks are structurally different from each other.

More specifically, this study showed that relatively fewer participants provided information for the proximity network compared to the nominated network in both the longitudinal and cross-sectional samples. One explanation is that the participants did not take the smartphone with them all the time and left it in a fixed place where they filled out the questionnaires (e.g., at home). This is supported by our observations in the classrooms when collecting the research materials at the end of each wave. Some participants did not bring the research smartphones to school and told us that their parents did not allow them to take the smartphone outside of their home. Also, we observed

a decline in the number of edges per participant in the proximity in the course of the project, while this was not the case for the nominated network. Potentially, the first time that the participants receive the smartphone (the first wave of Y1 and Y3), they did adhere to the instruction to have the phone on them at all times. However, this adherence declined as adolescents participated for a second or third time (W2 and W3 of Y1). Another explanation is that the measured weeks also included weekend days. On these days, it is likely that participants were not in proximity to peers and therefore no edges were measured in the proximity network. However, participants were able to nominate peers at the weekend. Furthermore, the communication network included the lowest number of participants while having the highest number of edges per participant. This means that if participants used the online social platform, this produced more edges than the other two networks. An explanation for this is that not all participants wanted to use the social platform because they did not like the platform or preferred another social platform. To conclude, this means that the social networks based on peer nominations included the most participants, but lacked the specific of connections per wave compared to the social networks based on online communications or Bluetooth connections.

In addition, this study showed that the nominated network was more stable over time than the proximity network. That is, more than half of the connections in the nominated network were present in two subsequent waves. Only one in four connections in the proximity network was present in two subsequent waves. This could indicate that the nominated network measures a state of a relationship, whereas proximity network measures an event within a relationship. Both types of relationships can be relevant for different types of research questions, making one type of network not more important than the other. The idea that both networks measure different types of relationships is supported by the low overlap between the networks that were observed within the waves. In pairwise comparisons of the networks, on average only one in four connections was present in both the nominated network and the proximity network.

Lastly, this study investigated sex segregation within the networks. Our findings showed that in both the longitudinal Y1 sample and the cross-sectional Y3 sample, the nominated network was segregated by sex, but that the proximity network of Y1 was not. In addition, the nominated network and the communication network showed similar levels of sex segregation. This means that adolescents are more likely to nominate, or communicate online with, others of the same sex, compared to those who they interact with during the day.

4.1 Limitations

Because of the novelty of the current study, several limitations should be discussed before drawing conclusions on the differences between the methods for collecting social network data. First, we noticed that participants did or could not always follow our instructions to have the smartphone with them at all times. In addition, no extra steps were taken to double-check whether the Bluetooth was working on all the research smartphones during data collection. As a result, fewer participants may have been included in the proximity network, especially when the adolescents participated for a second or third time (in waves 2 and 3). Therefore, not all classes could be included in the analyses of the stability of the networks over time. In this paper, we stayed close to the observed data by excluding classrooms in which less than half of the participants (and more than eight participants) provided data for all the three waves. The exclusion of some of the classrooms might have affected the results. Recent advancements in social network analysis provide possibilities to use multiple imputations of missing data in social networks (Krause et al., 2018). Future research can use imputation strategies or explore ways to increase responses in the beacon network, for example, by investigating whether wearable sensors are better suited for studying proximity in adolescents (e.g., Montanari et al., 2017). This way, participants wear the beacon at all times and more proximity data will be available per participant.

Also, the communication network was derived from the Social Buzz in the MyMovez app, which may have not been their preferred social app. Especially older adolescents already used WhatsApp or Facebook. And given that participants were informed that the Social Buzz was moderated by the researchers, the communication networks in this study might not reflect the natural occurring online communication networks outside of this study. In addition, the boundaries of the communication network and the nominated network were by design the same (i.e. the classroom). However, it is conceivable that adolescents have online communication partners that are outside of the school, or that they do not often meet face-to-face in their daily lives. Therefore, future studies should explore how the personal smartphone and social apps of participants could be used to generate social networks of adolescents, broader than the class. Likewise, network scientists could employ the same methods as used in the current study to compare social networks in different contexts, for example, organizational (Dong et al., 2012) or healthcare (Antonacci et al., 2017) settings.

Second, as a first step in determining criterion validity of the proximity network and the communication network, we looked at the segregation of connections by sex. Additional measures of network properties, other sociometric questions, or linking the social networks to behaviors can be used to provide a more extensive test of validity. By making the data publicly available, we want to offer other researchers the opportunity to further investigate the different types of networks and their properties. Also other types of research setups could be used to validate the unobtrusive measures, such as observations or lab studies. For example, researchers could observe the interactions of adolescents on the schoolyard and compare the observations to the proximity data.

Third, every connection in the proximity network was treated as a relevant connection and we set the threshold for proximity connections at one connection. So every connection between two participants was included in the analyses. We did not systematically investigate how increasing the threshold would change the network and how this would relate to the nominated network. Also, not all participants provided equal amounts of data. For some, one interaction might be valuable because it is one of only three interactions that day. Yet, if another person has 300 interactions within their classroom, that one connection is only a small fraction of the total interactions and might be less meaningful (Salathé et al., 2010). The same applies to the communication network, in which the relative frequency in which an adolescent sends a message to a peer impacts the relationship between the two adolescents (Wang et al., 2013). Future studies could further investigate (personalized) thresholds for when proximity and communication interactions are meaningful. One study is already looking into optimizing the data of the proximity network and investigating how many days of proximity data are needed to reflect a social network based on nominations (Simoski et al., 2019).

4.2 Conclusion

Altogether, the findings of this study indicate that nominated, communication, and proximity networks capture distinct types of connections between adolescents, potentially measuring different types of relationships. The nominated network is a stable network that includes relatively the most participants, but lacks the specificity of day-to-day measures and do not distinguish in the number of interactions in the relationship. Therefore, the nominated network can be described as a trait type of relationships between participants that is least likely to fluctuate over time. The communication network is very specific social network that includes many interactions per wave but includes the lowest number of participants. However, the communication network does not add many unique edges to the other networks. Therefore, the communication network can be seen as a collection of events within relationships. That is, this network could be used to supplement nominated networks by giving weights to the edges depending on the frequency of interactions or determine the relative reciprocity between two individuals (e.g., Wang et al., 2013). Lastly, the proximity network measures many interactions per wave but includes fewer participants than

the nominated network. The results indicate that the proximity network measures a type of relationship that can be described as a measure of events or occurrences. Therefore, this type of social network is more relevant when applied to relationships or behaviors that vary heavily from day-to-day. Also, the proximity network might reflect a broader range than deliberate peer interactions. Therefore, depending on the research aim, using one type of network over the other might have preference. For example, proximity networks have better fit with a topic that involves contagions, such as diseases or innovations. Nominated networks have a better fit with topics that involves social relationships, such as friendships or health behaviors (Adams, 2010).

With this in mind, the communication and proximity networks seem promising unobtrusive measures of peer interactions, with the additional benefit of multiple connections between participants within a measurement period. However, given the structural difference between the networks, the communication and proximity networks should not be used as a direct substitute for sociometric nominations. Network scientists should bear in mind what type of connections they wish to assess and use the best fitting network or combination of networks.

Declarations

Ethics approval and consent to participate

Informed consent was obtained from one of the parents of the participants before the start of the trial. Study procedures were approved by the Ethics Committee of the Radboud University (ECSW2014-100614-222).

Availability of data and material

The datasets used in the current study are available on: https://doi.org/10.17026/dans-zz9-gn44.

Funding

Authors' contributions

TW is the primary researcher on the study and is also the corresponding author. TW, KB, WJB, CS, LB, and MB conceptualized the design of the study; and TW, KB, CS, LB provided the instructions at the location and collected the data. TW and WB analyzed the data; TW wrote the concept version of the manuscript. KB, WB, CS, LB, and MB critically reviewed the manuscript. All authors have given final approval of the manuscript and agreed to be accountable for the accuracy and integrity of any part of the work.

References

Adams, J. (2010). Distant friends, close strangers? Inferring friendships from behavior. Proceedings of the National Academy of Sciences, *107*(9), E29–E30. doi: 10.1073/pnas.0911195107.
Aiello, W., Chung, F., & Lu, L. (2000). *A random graph model for massive graphs*. Oregon, Portland: STOC.

Antonacci, G., Colladon, A. F., Stefanini, A., & Gloor, P. (2017). It is rotating leaders who build the swarm: Social network determinants of growth for healthcare virtual communities of practice. *Journal of Knowledge Management*, *21*(5), 1218–1239. doi: 10.1108/JKM-11-2016-0504.

Bates, D. M. (2010). Lme4: Mixed-effects modeling with R (Version 1.1-21). Retrieved from https://cran.r-project.org/web/packages/lme4/index.html.

Bauman, K. E., Faris, R., Ennett, S. T., Hussong, A., & Foshee, V. A. (2007). Adding valued data to social network measures: Does it add to associations with adolescent substance use? *Social Networks*, *29*(1), 1–10. doi: 10.1016/j.socnet.2005.11.007.

Berndt, T. J., Hawkins, J. A., & Hoyle, S. G. (1986). Changes in Friendship during a School Year: Effects on Children's and Changes in friendship during a school year: Effects on children's and adolescents' impressions of friendship and sharing with friends. *Child Development*, *57*(5), 1284–1297. doi: 10.2307/1130451.

Berndt, T. J., & Hoyle, S. G. (1985). Stability and change in childhood and adolescent friendships. *Developmental Psychology*, *21*(6), 1007–1015.

Bevelander, K. E., Smit, C. R., van Woudenberg, T. J., Buijs, L., Burk, W. J., & Buijzen, M. (2018). Youth's social network structures and peer influences: Study protocol MyMovez project—Phase I. *BMC Public Health*, *18*(1), 504. doi: 10.1186/s12889-018-5353-5.

Camarena, P. M., Sarigiani, P. A., & Petersen, A. C. (1990). Gender-specific pathways to intimacy in early adolescence. *Journal of Youth and Adolescence*, *19*(1), 19–32. doi: 10.1007/BF01539442.

Campbell, R., Starkey, F., Holliday, J., Audrey, S., Bloor, M. J., Parry-Langdon, N., . . . Moore, L. (2008). An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): A cluster randomised trial. *Lancet (London, England)*, *371*(9624), 1595–1602. doi: 10.1016/S0140-6736(08)60692-3.

Chan, A., & Poulin, F. (2007). Monthly changes in the composition of friendship networks in early adolescence. *Merrill-Palmer Quarterly*, *53*(4), 578–602. doi: 10.1353/mpq.2008.0000.

Chan, D., & Cheng, G. (2004). A comparison of offline and online friendship qualities at different stages of relationship development. *Journal of Social and Personal Relationships*, *21*(3), 305–320. doi: 10.1177/0265407504042834.

Cho, E., Myers, S. A., & Leskovec, J. (2011). *Friendship and mobility: User movement in location-based social networks*. 1082–1090. https://doi.org/10.1145/2020408.2020579.

Csardi, G., & Nepusz, T. (2005). The igraph software package for complex network research (Version 1.2.4.1). Retrieved from https://cran.r-project.org/web/packages/igraph/.

de la Haye, K., Robins, G., Mohr, P., & Wilson, C. (2011). How physical activity shapes, and is shaped by, adolescent friendships. *Social Science & Medicine*, *73*(5), 719–728. doi: 10.1016/j.socscimed.2011.06.023.

De Lange, D., Agneessens, F., & Waege, H. (2004). Asking social Network questions: A quality assessment of different measures. *Metodoloski Zvezki*, *1*(2), 351–378.

Del Vicario, M., Zollo, F., Caldarelli, G., Scala, A., & Quattrociocchi, W. (2017). Mapping social dynamics on Facebook: The Brexit debate. *Social Networks*, *50*, 6–16. doi: 10.1016/j.socnet.2017.02.002.

Dong, W., Olguin-Olguin, D., Waber, B., Kim, T., & Pentland, A. "Sandy." (2012). Mapping organizational dynamics with body sensor networks. *2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks*, 130–135. doi: 10.1109/BSN.2012.16.

Eagle, N., Pentland, A., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, *106*(36), 15274–15278. doi: 10.1073/pnas.0900282106.

Ebel, H., Mielsch, L.-I., & Bornholdt, S. (2002). Scale-free topology of e-mail networks. *Physical Review E*, *66*(3), 1–4. doi: 10.1103/PhysRevE.66.035103.

Garton, L., Haythornthwaite, C., & Wellman, B. (1997). Studying Online Social Networks. *Journal of Computer-Mediated Communication*, *3*(1), JCMC313. doi: 10.1111/j.1083-6101.1997.tb00062.x.

González-Bailón, S., Wang, N., Rivero, A., Borge-Holthoefer, J., & Moreno, Y. (2014). Assessing the bias in samples of large online networks. *Social Networks*, *38*, 16–27. doi: 10.1016/j.socnet.2014.01.004.

Hamers, L. (1989). Similarity measures in scientometric research: The Jaccard index versus Salton's cosine formula. *Information Processing and Management*, *25*(3), 315–318.

Hartup, W. W. (1996). The company they keep: Friendships and their developmental significance. *Child Development*, *67*(1), 1–13. doi: 10.1111/j.1467-8624.1996.tb01714.x.

Kim, T., McFee, E., Olguin, D. O., Waber, B., & Pentland, A. "Sandy." (2012). Sociometric badges: Using sensor technology to capture new forms of collaboration. *Journal of Organizational Behavior*, *33*(3), 412–427. doi: 10.1002/job.1776.

Kossinets, G. (2006). Empirical analysis of an evolving social network. *Science*, *311*(5757), 88–90. doi: 10.1126/science.1116869.

Krause, R. W., Huisman, M., Steglich, C., & Snijders, T. A. (2018). Missing Network Data A Comparison of Different Imputation Methods. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 159–163. https://doi.org/10.1109/ASONAM.2018.8508716.

Li, N., & Chen, G. (2009). Analysis of a location-based social network. *International Conference on Computational Science and Engineering*, *4*, 263–270. doi: 10.1109/CSE.2009.98.

Marin, A., & Hampton, K. N. (2007). Simplifying the personal network name generator: Alternatives to traditional multiple and single name generators. *Field Methods*, *19*(2), 163–193. doi: 10.1177/1525822X06298588.

Marks, P. E. L., Babcock, B., Cillessen, A. H. N., & Crick, N. R. (2013). The effects of participation rate on the internal reliability of peer nomination measures. *Social Development*, *22*(3609–622. doi: 10.1111/j.1467-9507.2012.00661.x.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444.

Mercken, L., Snijders, T. A., Steglich, C., & de Vries, H. (2009). Dynamics of adolescent friendship networks and smoking behavior: Social network analyses in six European countries. *Social Science & Medicine*, *69*(10), 1506–1514.

Montanari, A., Nawaz, S., Mascolo, C., & Sailer, K. (2017). A Study of Bluetooth Low Energy performance for human proximity detection in the workplace. 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom), 90–99. https://doi.org/10.1109/PERCOM.2017.7917855.

Olguín-Olguín, D., & Pentland, A. (2010). *Sensor-based organisational design and engineering.* https://doi.org/10.1504/IJODE.2010.035187.

Onnela, J.-P., Saramaki, J., Hyvonen, J., Szabo, G., Lazer, D., Kaski, K., . . . Barabasi, A.-L. (2007). Structure and tie strengths in mobile communication networks. Proceedings of the National Academy of Sciences, *104*(18), 7332–7336. doi: 10.1073/pnas.0610245104.

Polastre, J., Szewczyk, R., & Culler, D. (2005). *Telos: Enabling Ultra-Low Power Wireless Research.*

R Core Team. (2019). *R: A language and environment for statistical computing.* Retrieved from http://www.R-project.org/.

Salathé, M., & Jones, J. H. (2010). Dynamics and Control of Diseases in Networks with Community Structure. *PLOS Computational Biology*, *6*(4), e1000736. doi: 10.1371/journal.pcbi.1000736.

Salathé, M., Kazandjieva, M., Lee, J. W., Levis, P., Feldman, M. W., & Jones, J. H. (2010). A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences*, *107*(51), 22020–22025. doi: 10.1073/pnas.1009094108.

Simoski, B., Klein, M., Araujo, E. F. M., van Halteren, A. T., van Woudenberg, T. J., Bevelander, K. E., & Buijzen, M. (2019). Parameter Optimization for Deriving Bluetooth-based Social Network Graphs. Presented at the IEEE Internet of People, Leicester, England.

Subrahmanyam, K., Reich, S. M., Waechter, N., & Espinoza, G. (2008). Online and offline social networks: Use of social networking sites by emerging adults. *Journal of Applied Developmental Psychology*, *29*(6), 420–433. doi: 10.1016/j.appdev.2008.07.003.

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. *Social Networks*, *34*(1), 73–81. doi: 10.1016/j.socnet.2011.05.006.

Valkenburg, P. M., & Peter, J. (2011). Online Communication Among Adolescents: An Integrated Model of Its Attraction, Opportunities, and Risks. *Journal of Adolescent Health*, *48*(2), 121–127. doi: 10.1016/j.jadohealth.2010.08.020.

Van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, The*, *25*(4), 40.

Wang, C., Lizardo, O., Hachen, D., Strathman, A., Toroczkai, Z., & Chawla, N. V. (2013). A dyadic reciprocity index for repeated interaction networks. *Network Science*, *1*(1), 31–48. doi: 10.1017/nws.2012.5.

Wasserman, S. (1994). *Advances in social network analysis: Research in the social and behavioral sciences.* Thousand Oaks, CA: SAGE.

Wen, Q., Gloor, P. A., Fronzetti Colladon, A., Tickoo, P., & Joshi, T. (2019). Finding top performers through email patterns analysis. *Journal of Information Science*, 0165551519849519. doi: 10.1177/0165551519849519.

Wilson, C., Sala, A., Puttaswamy, K. P. N., & Zhao, B. Y. (2012). Beyond social graphs: User interactions in online social networks and their implications. *ACM Transactions on the Web*, *6*(4), 1–31. doi: 10.1145/2382616.2382620.