Beyond the Core: Who Has Larger Social Networks?

Bas Hofstra, *Stanford University*
Rense Corten, *Utrecht University*
Frank van Tubergen, *Utrecht University*

The sociological literature on social networks overwhelmingly considers the number of core social contacts. Social networks, however, reach far beyond this small number of social ties. We know little about individual variation in the size of such extended social networks. In this study, we move beyond core networks and explain individual variation in the extended social network size among youth. We use survey data of Dutch adolescents ($N = 5{,}921$) and use two state-of-the-art measurements to compute extended network sizes: network scale-up methods through Bayesian modeling and the observed number of contacts on Facebook. Among both measurements, we find that extended networks are larger among ethnic majority members, girls, and those who often engage in social foci. This highlights a crucial role for preferences and opportunity in the genesis of extended networks. Additionally, we find that differences between both network sizes (scale-up and Facebook) are smaller for girls and higher educated. We discuss the implications of these findings and suggest directions for future research.

Introduction

Why are some individuals socially isolated but others highly connected? Scholars have devoted substantial attention on this question with regard to the number of people with whom individuals closely relate (Parigi and Henson 2014; Marsden 1987; McPherson et al. 2006). People seem to have close ties with only a few others. This conclusion often originates from studies on people's "core discussion network," where individuals are asked with whom they discuss "important matters." Scholars occasionally explain variation in the number of these core discussion partners (Burt 1984; Marsden 1987; McPherson, Smith-Lovin, and Brashears 2006). Findings from 1984, 2004, and 2008 (McPherson et al. 2006; Hampton, Sessions, and Her 2011a) revealed that American adults had an average of two to three core ties.[1] Dutch adults report fewer than three discussion partners (Van Tubergen 2014). It was estimated that Americans on average had

Direct correspondence to Bas Hofstra, Graduate School of Education, Stanford University, 520 Galvez Mall, Stanford, CA 94305, USA. E-mail: bhofstra@stanford.edu.

Social Forces 99(3) 1274–1305, March 2021
doi:10.1093/sf/soaa031
Advance Access publication on 22 April 2020

17 alters with whom they had "trusting" relationships—close friends, discussion partners, or those trusted for advice or with money (DiPrete et al. 2011).

Current literature on variation in social network size often focuses on these core network contacts. Individuals' social circles, however, reach further than this small group of core or trusting ties (Gurevich 1961; Pool and Kochen 1978). Social networks consist of *network layers* (Dunbar 1998). The first two layers range from 2–3 *core* ties to 15–17 *trusting* social relationships (or *sympathy group*, Dunbar 2016). The two layers beyond those are hypothesized to have approximate sizes of 50 and 150 (Dunbar 2016).[2] Beyond that is something that is defined as the *extended social network*. This layer includes all former layers and is the prime focus of this study.

We know surprisingly little about individual differences in the size of this extended social network, including both strong and weaker ties. This lack of attention is remarkable from a *substantive* point-of-view, because the extended social network size relates to many societally relevant issues. For instance, acquaintanceship ties facilitate access to information embedded in social networks (Granovetter 1973), and larger networks (measured via distal proxies) are associated with better health and well-being, and more social support (Holt-Lunstad, Smith, and Layton 2010). Yet, recent work started to explain individual variation in the extended network size (DiPrete et al. 2011; Lubbers, Molina, and Valenzuele-García 2019).

From a *methodological* perspective this research lacuna is less surprising. It is not straightforward to measure social ties. It is challenging to measure the number of core contacts (see Bearman and Parigi 2004; Small et al. 2015).[3] It stands to reason that measuring the number of ties becomes increasingly challenging when moving from close to distant ties (see Gurevich 1961). A growing body of literature, however, develops methodologies for extended network size estimates (rather than explaining variation) (Killworth et al. 1998a; McCarty et al. 2001; McCormick et al. 2010). Earlier measurement strategies provide a wide range of network size estimates, ranging from 108 (Killworth et al. 1998a) to 5,520 (Freeman and Thompson 1989). Methods to arrive at these point-estimates vary: asking respondents whom they knew from randomly drawn pages from phonebooks (Pool and Kochen 1978); using summation methods counting how many people respondents indicated they knew from a list of given relationships (McCarty et al. 2001); or counting the number of Christmas cards respondents send out (Hill and Dunbar 2003). Discrepancies in definitions on what constitutes a social tie cause these vast differences.

We thus have little substantive knowledge on who has larger networks beyond core ties, and it seems difficult to measure extended networks. These issues are related: methodological issues lead to only few substantive studies. Here, we advance our understanding pertaining to both issues.

First, we answer a basic but fundamentally substantive question: what explains individual variation in the extended network size? As Kadushin (2012) puts it, we "do not as yet have a theory or a systematic study of the causes of these variations" (p. 72) in network size. Sociological literature mostly focuses on core networks and often explains its variation through individual preferences

(McPherston et al. 2001) or broader social structures (Blau 1977; Feld 1981). Psychological studies show how personality factors (e.g., extraversion) relate to network size (Selden and Goodie 2018; Assendorpf and Wilpers 1998; Roberts et al. 2008; Selfhout et al. 2010). Yet, prior work hardly systematically considers individual variation in the *extended network size*. Here, we contribute in particular to sociological work (see Zheng et al. 2006; DiPrete et al. 2011; Lubbers et al. 2019) that found differences in the extended social network size by gender, race, income, education, and social foci. We test existing sociological explanations for the core network size and examine to what extent they apply to extended networks. Specifically, we elaborate a set of sociological mechanisms and derive our hypotheses based on those. Note that we do not directly observe these mechanisms, but assume theoretically that they generate our outcome. We then confront these hypotheses with our data. Because theories on opportunities, preferences, and their interplay (Blau 1977; Feld 1981; McPherson, Smith-Lovin, and Cook 2001) are fundamental in core tie formation, we take these as our point-of-departure. Hence, we situate these theories in the extended network size literature. Additionally, we develop a hypothesis on the role of romantic partners, one that contrasts prior findings on the role of partners in core networks. And finally, we explore intuitions on differences in network size by education and gender (see Brashears, Hoagland, and Quintane 2016). As such, we contribute new evidence on individual variation in extended network size.

Second and uniquely, we use a set of similar covariates for two recent methods that provide extended network size estimates. The first metric uses recent, state-of-the-art developments in the "network scale-up method" for a measure of the extended network size. In this method, respondents are asked via surveys how many people they know from various subpopulations (Killworth et al. 1998a; Killworth et al. 1998b; Zheng, Salganik, and Gelman 2006; McCormick, Salganik, and Zheng 2010; DiPrete et al. 2011; Maltiel et al. 2015). One can then "scale-up" to a population to calculate the network size. Our second metric considers unobtrusive, behavioral data and measures individuals' number of online social media contacts (e.g., Gonçalves, Perra, and Vespignani 2011; Kanai et al. 2012; Pollet, Roberts, and Dunbar 2011; Dunbar et al. 2015; Dunbar 2016).

However, both methods have limitations. It is difficult to put scale-up findings into a broader context of network sizes when covariates' correlations with network sizes are not compared with other measures of this unobserved (and often ill-defined) concept. Here, we consider the network scale-up estimates' lack-of-broader-context as an intrinsically valuable feature. We build on work by Hampton, Sessions Goulet, and Purcell (2011b) and compare two methodologically distinct but potentially similar concepts (despite likely *absolute* differences) within the same set of respondents and using the same covariates. This enables comparisons of covariates among the two outcomes and explicitly modeling their differences. Hence, we empirically situate this measurement in discussions about the extended network size. The key limitation to measuring network size on social media is that they are selective: both in membership—who becomes member?—and in privacy—who shows networks online? When some groups

maintain more privacy online and become members less frequently (Hofstra, Corten, and Van Tubergen 2016a, 2016b), estimated differences in network size based on public profiles between groups might become biased. Through combining detailed survey data with behavioral data from social media, we take into account sample selections in privacy and membership to attempt to offset biases in extended network size estimates.

In sum, we contribute new evidence on individual variation in extended network size and compare two state-of-the-art methods to capture extended network size. We provide this new evidence based on new, large, and linked empirical data. We focus on a surprisingly understudied target population in relation to network sizes: adolescents. Specifically, we link survey data on Dutch adolescents ($N = 5{,}921$) in 2012 (Kalter et al. 2015) to two data sources collected in 2014: (1) to survey data measuring these respondents' network size using the scale-up method and (2) to these respondents' observed Facebook profiles to measure their network size online (Hofstra, Corten, and Van Tubergen 2015; Jaspers and Van Tubergen 2017).

Theory and Hypotheses

At what distance are ties included in the extended network size? Here, we use a pragmatic definition of extended social networks provided by McCarty et al. (2001: 29) and DiPrete et al. (2011: 1242). We consider "all the contacts whom individuals know on a first name basis to be part of the social network, such that they would have a friendly chat if they were to meet randomly." Substantively, this definition relates to the societal outcomes discussed before. These contacts may grant meaningful connections in terms of knowledge to otherwise unknown sub-cliques. Methodologically, this definition provides a convenient boundary for persons recalling their network contacts.

Opportunities and Homophily

Meeting opportunities are key for the genesis of core ties (Blau 1977; Feld 1981). One dimension of meeting opportunities are *foci*. A focus is a "social, psychological, legal, or physical entity around which joint activities are organized" (cf. Feld 1981: 1016). Typical foci are associations, neighborhoods, work places, or schools. Those individuals who share a focus will share activities, more so than individuals who do not share a focus. Sharing these activities facilitates positive interactions between people and brings them together in reciprocally rewarding situations (Feld 1981). Hence, sharing a focus increases the likelihood for (positive) ties to form.

Research suggests that both strong and weak ties are formed in some sort of focus (Wimmer and Lewis 2010). Many acquaintances are met, for instance, at associations or at parties among adults (Mollenhorst et al. 2008), on campus among students (Wimmer and Lewis 2010), at schools among adolescents (Hofstra et al. 2017), or at work among adults (Zheng et al. 2006). The

importance of meeting opportunities for tie formation crosscuts target populations. We use this idea and conjecture that those who engage more often in socially or recreationally orientated foci have more opportunities to get into contact and make acquaintances with other people compared to those who do so less often—i.e., the extended network is a function of the time individuals spent in foci. We examine five foci to capture adolescents' social life: going out, associations, concerts, family, religious meeting places, and potential workplaces. We hypothesize:

Hypothesis 1: *Individuals who spend more time (a) going out, (b) in associations, (c) visiting concerts, (d) with family, (e) at religious meeting places, and (f) have a job, have larger extended social networks.*

A substantial body of literature focuses on tie formation and its relation to social network segregation (Kalmijn, 1998; McPherson et al. 2001; Currarini, Jackson, and Pin 2010). Besides the role of meeting opportunities, a common explanation for network formation is homophily (Kalmijn 1998; McPherson et al. 2001). Results in this line of research are that racial-ethnic segregation in networks is a ubiquitous feature of social life—often caused by an interplay of homophily and opportunity (Currarini et al. 2010; Wimmer and Lewis 2010). Given that racial-ethnic segregation in social networks is pervasive, we argue that an interplay of homophily and opportunity also relates to the extended network size.

For clarity, we reserve the term *homophily* for the preference of individuals to form relationships to similar others (*choice homophily* in McPherson et al.'s [2001] terminology). This follows Wimmer and Lewis' (2010: 588) suggestion to use homophily exclusively for the tie-generating mechanism, not the network outcome.

A second dimension of opportunity—besides foci—is the size of groups relative to other groups (Blau 1977). Homophily and group size work in tandem and cause differences in extended network sizes between ethnic minority and majority members. When people belong to a larger group, they have many opportunities to select ethnically similar ties, whereas minority members have fewer possibilities to make such homophilous choices. Ties among ethnically dissimilar people are costlier and require higher initial investments to overcome cultural boundaries (Kalmijn 1998). Assuming that there is a limit on the investments one can make in terms of time or emotional commitment, fewer ties are formed if the pool of potential alters includes relatively more *dissimilar* people. As such, ethnic minorities establish fewer extended network ties compared to ethnic majority members. Furthermore, established ties—assuming that among ethnic minority groups ties are more often between dissimilar people—will be broken more frequently (cf. Smith, Maas, and Van Tubergen 2012).

Approximately 79% of the population are so-called "Dutch majority" members in the Netherlands (Statistics Netherlands 2015). Relatively much smaller groups are ethnic minorities with an immigrant background. Minorities with a Turkish or Moroccan background, for instance, cover ~6% of the Dutch population. Hence, Dutch majority members—who prefer befriending other Dutch majority members—have ample possibilities to choose similar ties. Members of

ethnic minorities have far fewer such opportunities and will therefore have fewer ties. The group from which they prefer selecting their contacts essentially is much smaller. The negative opportunity element limits their possibilities to exhibit a preference for similar others. Hence, relative differences in group sizes will result in an extended network size that is significantly larger for those people who belong to a majority ethnic background, than for those whose representation is smaller. Such an interplay between homophily and group size affects the number of Twitter connections among students (Halberstam and Knight 2016), and DiPrete et al. (2011) find that racial minorities have smaller extended social networks among adults too. Therefore, we hypothesize:

Hypothesis 2: *Dutch majorities have larger extended social networks than those from ethnic minority groups.*

Differences in the extended network size may also reflect ethnic segregation of social settings. As foci increase the likelihood of a tie emerging between two people sharing a focus, this likelihood will be higher when potential contacts in foci share an ethnic background. This is again based on the conjecture that individuals prefer befriending others with whom they share an ethnicity (McPherson et al. 2001). Hence, when there are many others sharing an ethnic background with in a focus, individuals will more likely form more ties as they have ample possibilities to make homophilous choices (see Hofstra et al. [2017] among adolescents). Here, we focus on the school setting as a key focus of tie formation among adolescents (McPherson et al. 2001). We do not have information on the ethnic composition of the social foci mentioned before, nor is adolescents' school-of-choice completely exogenous. Some of the relative group size effects may result from nonrandom sorting of adolescents over foci that we do not address. We do, however, investigate the number of potential alters in schools and school classes who ethnic backgrounds with the focal adolescent. We hypothesize:

Hypothesis 3: *Individuals who have a greater number of co-ethnic individuals in (a) their schools and (b) in their school classes will have larger extended social networks.*

Romantic Partners

Previous work suggests that individuals who are in romantic relationships have fewer strong relationships (Kalmijn 2003; Rözer, Mollenhorst, and Volker 2015). Features of strong relationships are such that they require time and emotional investments (Granovetter 1973). "Social withdrawal" after finding a partner may be related to individuals' limited resources to maintain core contacts—people initially have a limited amount of time and emotional capacity to maintain relationships with close contacts *besides* romantic partners (Slater 1963; Kalmijn 2003).

This pattern, however, differs for weaker compared to close ties. When you form a romantic relationship, a partner may introduce you to many new *acquaintances*. These new acquaintanceships are cheaper to maintain than strong ties, and they involve less time and less emotional investment. These new

acquaintances introduce you to again other people, and so on. This introduces a bandwagon effect where some people are in a better position to obtain new contacts compared to others. Romantic partners provide these opportunities to meet new contacts (when social circles do not entirely overlap), thus contrasting a "social withdrawal from strong ties" as distal contacts are cheaper to maintain. Recent results seem to align with such intuitions. Van Tubergen and colleagues (2016) found that those adults in a romantic relationship had more social contacts. (Note that those with many social ties may be more likely to enter in a romantic relationship as well, and that we are unable to identify such reverse causality.) We hypothesize:

Hypothesis 4: *Individuals who indicate being in a romantic relationship have larger extended social networks than adolescents who are not.*

Education and Gender

Research on "cultural omnivores" shows that some groups—those of higher status/education—pursue broader ranges of leisure activities than others (Peterson 1992). Some of this discrepancy is attributed to capabilities; some individuals are simply better capable of managing a broader range and more activities than others. Hence, some of the variation in the finding that those of high status and education engage in a broader range of activities is a byproduct of them being cognitively able to do so. This may be one reason for larger network sizes among higher-educated individuals—they engage in a wider range of foci and, therefore, have more opportunities to befriend others than lower-educated individuals. The set of social foci we addressed before may not completely adjust for this indirect association.

Additionally, cognitive abilities themselves may be a factor in forming/maintaining of social ties (Dunbar 1998). Specifically, cognitive abilities facilitate keeping track and maintaining all of the relationships one has—the better one is able to do so, the more contacts one has. As such, cognitive capabilities correlate with the extended network size. If we are to assume that educational attainment is somewhat correlated with cognitive ability, we could assume that there is variation in the capacity to maintain and keep into contact with many social contacts by educational level as well. Both mechanisms lead to our hypothesis:

Hypothesis 5: *Individuals in higher educational track levels have larger extended networks than individuals in lower educational track levels.*

Another consistent finding is that women's social networks differ from men's. Women's networks are generally found to be larger (Moore 1990; Bastani 2007; Hampton et al., 2011a; Van Tubergen 2014) and include more kin (Marsden 1987; Van Tubergen 2014). A notable exception to this pattern is DiPrete et al. (2011), who find no gender differences.

There are various mechanisms that may explain gender differences. First, women might cognitively be better equipped to manage networks than men. One indication of this is that women appear better capable than men in recalling contacts (Brashears et al. 2016). This may be a result of circumstances that shape women such they "develop a relatively greater ability to encode and

recall social networks" (Brashears et al. 2016: 82). This is consistent with research showing that those in low power situations have greater knowledge of their social networks (Simpson et al., 2011), under the assumption that these circumstances imply lower status, prestige, or power positions of women compared to men.[4] Second, men and women may differ in their sociality, may have different dispositions toward social ties (Brashears et al. 2016), or may differ in their level of social activities. Each of these mechanisms may explain why women may have larger extended networks than man. Note that we do not test these mechanisms directly but explore these intuitions. Hence, we hypothesize:

Hypothesis 6: *Girls have larger extended social networks than boys have.*

Data

We use data on Dutch adolescents from the "Children of Immigrants Longitudinal Survey in Four European Countries" (CILS4EU) (Kalter et al. 2015). We use the Dutch data, as our measures of interest are included in that section, but data were collected in Sweden, Germany, and England too. We use the second ($N_{WAVE2} = 5{,}921$) and fourth ($N_{WAVE4} = 4{,}073$) waves: the second wave contains the latest school-level data for the total set of respondents and the fourth wave measures the network size using the scale-up method and Facebook. The project followed 14- to 15-year-old adolescents for three years with a one-year time lag starting in 2010. The Dutch section continued for four more years as the "Children of Immigrants Longitudinal Survey in the Netherlands" (CILSNL) (Jaspers and Van Tubergen 2014, 2017). The surveys include many features such as leisure time activities, personal networks, and so forth. The data are stratified by the proportion of non-Western immigrants attending schools. In these strata, schools were selected with a probability proportional to the school size using the number of students at the relevant educational track level (further details on sample selection and sensitivity checks can be found in Supplementary Material A1).

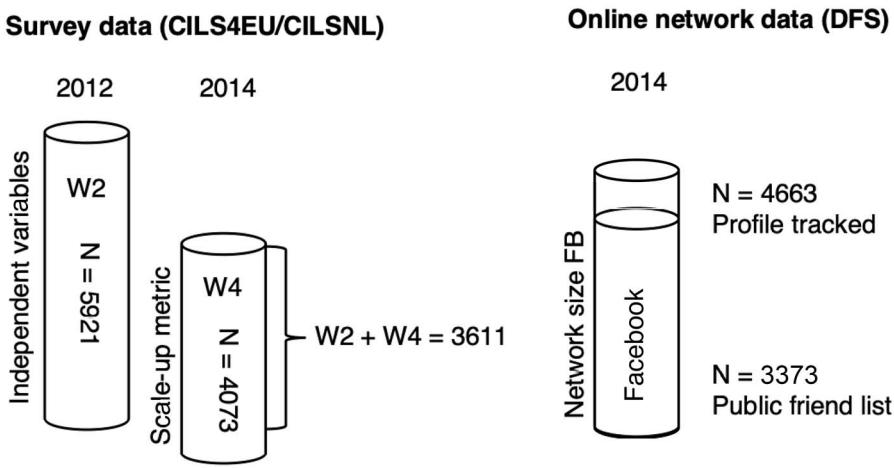
The Dutch Facebook Survey

We link the survey data to unobtrusively collected behavioral data from Facebook using the Dutch Facebook Survey (DFS; collected June–September 2014) (Hofstra et al. 2015). The DFS enriches the Dutch part of the CILS4EU and the CILSNL. Of the 4,864 respondents that indicated Facebook membership in waves 3 (2012–2013; $N = 3{,}423$) or 4 (2013–2014; $N = 3{,}595$) of the surveys, 4,473 (92%) were tracked on Facebook. For the respondents who kept a *public friend list*, we obtained friend lists ($N = 3{,}373$; 75.4% of respondents).[5]

Data Structure and Sample Selections

We link two waves of survey data with behavioral data from Facebook. The number of observations is 5,921 for wave 2, 4,073 for wave 4, and 3,373 for the

Figure 1. Schematic overview of the used data sources.



DFS (with a public friend list). Respondents keep their friend lists either private or public on Facebook (Hofstra et al. 2016b). We consider the respondents that have observable friend lists, as we want to measure their number of Facebook contacts in these lists. Potentially, we can compare 2,684 respondents' network sizes across the number of contacts on Facebook and the network scale-up measure. This is the number of respondents who participated in wave 4 of the survey and for whom we can observe their number Facebook contacts. Deletion of cases with missing values on the outcome variables leads to a set of 2,546 respondents. This is the number of cases we consider *descriptively*, and it captures the Facebook and scale-up network size measured during roughly the same period.

For our *inferential* analyses, we account for two types of selectivity: respondent attrition across waves 2 and 4 and selectivity in observable friend lists. We use Heckman selection models for both our network size measures, adjusting for respondents' gender, educational level, and ethnic background (outlined below) in the selection equations (detailed later on). As inclusion criteria for the inferential analyses, cases need to have nonmissing values on independent variables in wave 2 (N = 5,488 out of N_{WAVE2} = 5,921, 7.3% item nonresponse across independent variables). Figure 1 summarizes the data sources and their observations and Table 1 specifies conditions for inclusion in the analyses.

Independent Variables

Foci

Each independent variable is measured in wave 2 of the CILS4EU unless stated otherwise. Respondents were asked how often they spend time in foci. They

Table 1. Overview of the Used Criteria for Data Inclusion in the Analyses

	N
Inclusion criteria for descriptive analyses	
Public Facebook friend list + nonmissing network scale-up	2,546
Inclusion criteria for inferential analyses	
Nonmissing values independent variables in Wave 2	5,488

could indicate on a five-point scale (1-never, 2-less often, 3-once or several times a month, 4-once or several times a week, and 5-daily) how often they go out (bars/nightclub/etc.), spend time in associations (sport/music/etc.), visit concerts or DJs, spend time with family, and visit religious meeting places. Hence, we consider these five variables showing the time respondents spend in these foci. Additionally, respondents indicated whether they had a part-time job (yes/no).

Ethnic Background

We categorize respondents into ethnic background groups according to the country of birth of their biological parents, which is standard practice in scholarship on Dutch ethnic groups. When adolescents have one Dutch-born parent, they are categorized in the ethnic category of the parent not born in the Netherlands. When respondents have parents born in different non-Dutch countries, they are categorized in the mother's birth country. This categorization is regularly applied and used by Statistics Netherlands (2012). We categorize respondents into "Dutch ethnic majority" and "Ethnic minority" categories in line with Hypothesis 2 that contrasts these two groups (and because of few observations if we split into smaller groups). The "Ethnic minority" category includes adolescents of all non-Dutch ethnic origins (e.g., Dutch Caribbean, German, or Turkish youth).

Number of Co-ethnic in Class

We measured the number of students in a class who share ethnic backgrounds with respondents: the number of Dutch majority members for those of the Dutch ethnic majority, the number of Turkish origin for those of the Turkish ethnic minority, etc. We did use detailed categories for the "Other" ethnic background. For instance, we counted the number of Dutch Caribbean in a class for those of Dutch Caribbean ethnic background.

Number of Co-ethnic in School

Similarly, we calculated the number of *schoolmates* who share an ethnic background with the respondent (excluding classmates to separate effects of classmates from schoolmates). This variable was measured from secondary data obtained from the Dutch inspectorate of Education.

Table 2. Descriptive Statistics for the Independent Variables (*N* = 5,488)

	Min.	Max.	Mean	SD
Foci (H1)				
Going out	1	5	2.761	0.916
Associations	1	5	3.339	1.256
Concerts	1	5	1.932	0.762
Family	1	5	3.176	0.830
Religious meeting places	1	5	1.659	1.019
Job	0	1	0.563	–
Similarity of potential contacts (H2 + H3)				
Ethnic minority	0	1	0.294	–
Number co-ethnic class	0	28	12.451	7.869
Number co-ethnic school	0	2300	688.252	625.294
Romantic Partner (H4)	0	1	0.253	–
Educational track level (H5)				
Vocational	0	1	0.537	–
Senior general	0	1	0.255	–
University preparatory	0	1	0.208	–
Girls (H6)	0	1	0.510	–

Source: Survey data from the CILS4EU wave 2.

Romantic Partner

We measure whether the respondent indicated being in a romantic relationship (yes/no: "Do you have a boyfriend/girlfriend?").

Educational Track-Level

When Dutch adolescents transition to high school, they are placed in different educational tracks that differ in their type of education and level. We measured this categorization with an ordinal variable: 1-preparatory vocational education (Dutch: VMBO), 2-senior general (Dutch: HAVO), and 3-university preparatory education (Dutch: VWO).

Gender

We measure whether respondents indicated being a girl (1) or boy (0). Table 2 reports the descriptive statistics of the independent variables.

Confounding Factors

New or les-frequent Facebook users' number of contacts on Facebook is more remote from their overall number of connections compared to more-experienced

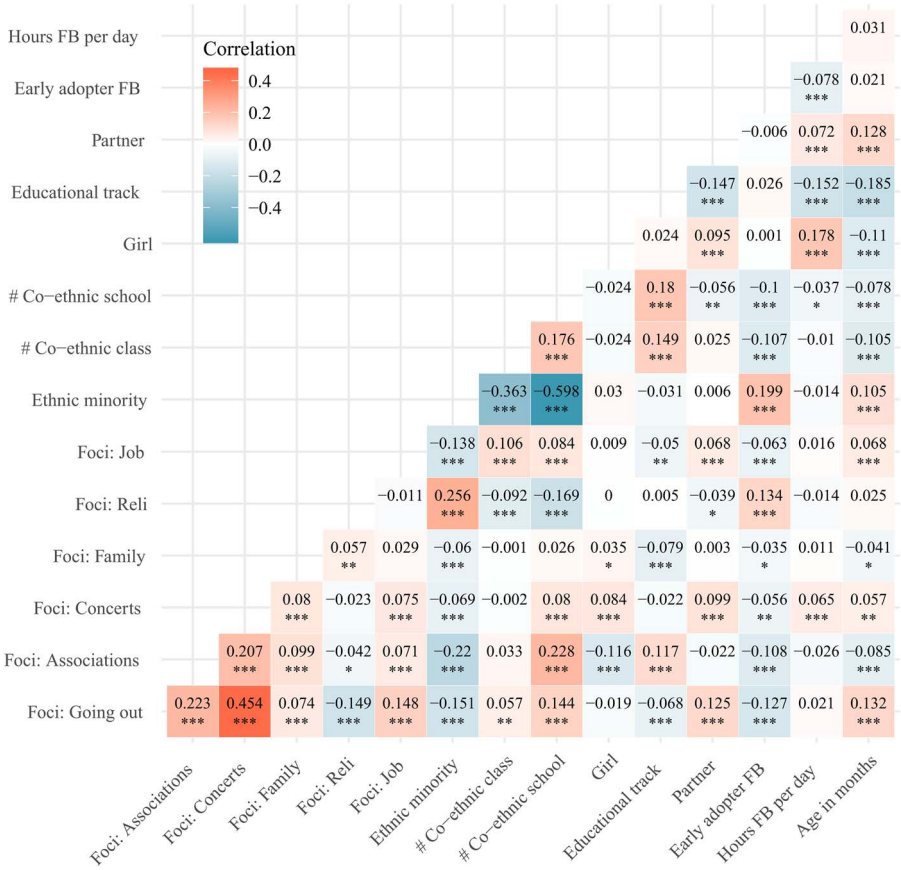
and more-frequent Facebook users. Therefore, we control for whether or not respondents were an early Facebook adopter in the Netherlands using a predefined classification—i.e., pre-2010 members were early adopters (Hofstra et al. 2016a)—and for the number of hours respondents spend each day on Facebook. The number of Facebook contracts is correlated with membership duration ($r = .268$; $p < .001$; Median = 2010) and with the amount of hours spent on Facebook per day ($r = .154$; $p < .001$; five categories, Median = 1 h or less). To help guide respondents and to reduce lack-of-response errors, the question for time spent on Facebook was asked using interval-censoring ("How much time per day do you spend on Facebook?": 1: ≤1 h to 5: ≥4 h). We categorize Facebook membership duration in years into two categories instead of keeping yearly categories to ensure large enough sample size in each cell (only seventy-five respondents became Facebook members in 2005–2007). Such early adoption of Facebook comes from the DFS, and the amount of hours spent on Facebook each day originates from wave 4 of the CILSNL. To maintain our sample size of 5,488 of nonmissing cases in wave 2, we include categories for respondents that did not participate in wave 2. Finally, we account for respondent's age in months—either when the Facebook (Mean = 223.708; SD = 7.633) or when the scale-up data (Mean = 218.905; SD = 7.815) were collected—because recent work suggests network sizes vary by age (Lubbers et al. 2019). Figure 2 depicts correlations between our covariates.

Measuring the Extended Social Network Size

The Number of Contacts on Facebook

On Facebook, members send/receive friendship invitations to/from others who accept/decline the invitation. When accepted, a Facebook tie within people's *friend list* shows an undirected, reciprocated friendship between two users. Using the DFS, we measure the number of contacts respondents have in their Facebook friend lists as the *network size on Facebook*. Note that this metric does not (nor do we intend to) distinguish between close friends or distal acquaintances. Yet, most empirical research suggests that Facebook contacts started out as and are likely meaningful, offline ties that may be categorized under our definition of extended network ties. Only .4% of online friendships are online-only among US college students (Mayer and Puller 2008), that primarily use online networks to strengthen and maintain their offline relationships (Ellison, Steinfield, and Lampe 2011) and befriend others on Facebook when they meet in offline foci (Wimmer and Lewis 2010; Hofstra et al. 2017). Additionally, about 80% of adolescents use online networks to maintain offline contacts (Subrahmanyam et al. 2008), and 77% of adolescents' online ties was formed offline (Van Zalk et al. 2014). Danish students consistently meet a large fraction of their Facebook friends offline (Sapiezynski et al. 2018) and US adults are overwhelmingly Facebook friends offline contacts (Duggan et al. 2015).

Figure 2. Correlations between covariates (∗$p < .05$; ∗∗$p < .01$;∗∗∗$p < .001$).



The Network Scale-Up Method

The network scale-up method (Killworth et al. 1998a; Killworth et al. 1998b) uses surveys to estimate individuals' extended social network size (McCormick et al. 2010). The method was developed to provide estimates of hard-to-reach populations (e.g., estimating the seroprevalence of HIV in a given target population). The method works as follows. Consider a population of size N. To estimate network size, one can ask respondents the number n randomly chosen members of the population they know. However, the larger the N is, the lower the likelihood that two randomly drawn persons know one another.

The network scale-up method circumvents this issue and asks individuals whether they know an entire set of people simultaneously. For instance, it asks "How many people do you know that are named Thomas?" instead of asking which of the ~40,000 people they know in the Netherlands are named Thomas (Meertens Institute 2016). When respondents then indicates that they know two

people named Thomas, one can estimate the total network size by assuming they (a) know 2/40,0000 of the entire population of persons named Thomas and (b) that this same proportion equally applies to the entire population (~17 million in the Netherlands),

$$\frac{2}{40,000} \times (17 \text{ million}) = 850. \tag{1}$$

The estimate's precision increases by averaging responses to different subpopulations (such as detainees). This yields the *basic scale-up estimator*:

$$\text{Scale-up degree}_i = \frac{\sum_{k=1}^{K} y_{ik}}{\sum_{k=1}^{K} N_k} \times N, \tag{2}$$

where y_{ik} is the number of people person i knows in subcategory k, N_k is the size of subcategory k, and N is the size of the population (cf. McCormick et al. 2010). Generally, the subpopulations that are prompted to respondents are occasionally referred to as "How many X's do you know?", where the X's refer to subpopulations.[6]

The CILSNL implementation shows respondents the following statement (translated from Dutch):

> *The next questions are about all the people you "know personally in the Netherlands." By knowing personally, we mean that you know the name of that person and that you would have a chat if you were to meet him or her on the street or in a shop.*

This implies reciprocal relationships, which makes it suitable for comparisons with the number of reciprocal contacts on Facebook. Respondents were prompted to recall contacts they know personally *in the Netherlands*. This phrasing is crucial as it allows to scale-up adolescents' recalled ties to the Dutch population and conveniently prevents two issues. First, many adolescents of immigrant background have transnational ties (Schimmer and Van Tubergen 2014) and, second, adolescents living close to national borders (e.g., Germany) may have many social ties across borders. Our statement thus prevents recalled contacts that we are unaware of and are unable to scale-up to because of unknown reference groups. Additionally, contacts in the Netherlands relate to "meaningful acquaintances" that provide help, support, and information in the national context. Our approach defines a clear and substantive boundary for our respondents, and one that is convenient methodologically.

Respondents indicated for *fifteen* populations how many contacts they had. Respondents were prompted with five names (Thomas, Kevin, Anne, Melissa, and Moham(m)ed) on the question "How many people do you know personally with the following name?" Response categories fell within the numerical ranges of 0, 1, 2–5, 6–10, 11–20, 21–50, or >50. To ease the answering process for

respondents and to reduce lack-of-response errors, the questions were asked using interval-censoring. This same strategy was used by DiPrete et al. (2011: 1251). We follow their strategy and take midpoints of the intervals. Following the strategy of Lubbers et al. (2019: 60), we take the value 11 as the largest number even when respondents indicate to know more individuals with a given name. This is to reduce recall errors that have been shown to grow significantly when respondents indicate to know many people in subpopulations. The typically Dutch names (first four) represent names of male and female genders from parents with either a higher or lower status background: Anne (girl, high status), Melissa (girl, low status), Thomas (boy, high status), and Kevin (boy, low status) (see Bloothooft and Onland 2011: 34). The fifth name represents the two most-prevalent versions in the Netherlands of the typical Islamic boy name Muhammed: Moham(m)ed (Meertens Instituut 2016). Including these different names will help mitigate the possibility that individuals from different societal strata, genders, and ethnicities know more or less of the prompted X's. Using the same interval-censoring, respondents indicated how many contacts they knew living in five medium-to-large cities in the Netherlands distributed equally among geographical regions (Groningen [north], Utrecht [center], Maastricht [south], Den Haag [south-west], Zwolle [north-center]). Three further questions noted how many contacts respondents had that were currently enrolled in one of three distinct levels of tertiary education in the Netherlands (MBO [tertiary lower-vocational], HBO [tertiary higher-vocational], University), again attempting to increase precision by prompting respondents to estimate their number of contacts across distinct Dutch educational strata. We take midpoints, and given that it is likely to know many individuals in any of these cities, or in any of the educational levels, we take fifty-one as the largest number. Finally, respondents indicated how many people they knew that were arrested in the last 12 months by the police or that were in jail. For the same reason as for the first names, we use midpoints and take eleven as the largest value (group population numbers in Supplementary Material A2).

Bayesian Estimation Using the Scale-Up Module.- Through our selection of X's we attempt to mitigate that some groups may know more people within specific subpopulations than other groups. This well-known issue is referred to as *barrier effects* (see Zheng et al. 2006; McCormick et al. 2010; DiPrete et al. 2011; Maltiel et al. 2015). We mitigate some of these barrier effects by carefully selecting a set of prompts that crosscut social strata and through sampling from one cohort of adolescents (i.e., lower age heterogeneity in reporting of known individuals). Yet, this phenomenon may still bias the basic scale-up estimator (Equation (2)). A second issue with the basic scale-up estimator is *transmission errors*: respondents may not always be perfectly aware of alter characteristics (e.g., whether a network contact lives in a given city).

Maltiel et al. (2015) propose to use Bayesian methods using Markov Chain Monte Carlo (MCMC) algorithms as an attempt to mitigate barrier effects and transmission errors. We follow their approach, and adjust the basic estimator in three ways. We allow for (1) a random effect for degree to regularize the estimates and reduce extreme value sensitivity, (2) the probability that a person

i knows someone in group *k* to vary randomly (instead of it being fixed) across individuals (accounting for barrier effects), and (3) a multiplier to the binomial proportion of people known in a subpopulation (accounting for transmission errors). For mathematical details, we refer to Maltiel et al. (2015: 1251) (model summary in Supplementary Material A3). We use the "NSUM" R package (Maltiel and Baraff 2015). For model priors, we follow Maltiel et al. (2015) and calculate starting-point parameters with the basic scale-up estimator, yielding a mean degree, a standard error (see Endnote 6), and an individual-level degree. We run 40,000 iterations of the MCMC algorithm (using closed-form Gibbs sampling), retain 4,000 of the chains, and calculate the average network size over these.

Ideally, we would directly fit a model of network size. Yet, available procedures require us to estimate the size of a hidden population even though we are aware of all fifteen population sizes in our module. To circumvent the limitations in the available methods and because we do not want to arbitrarily select one, we run the estimation process fifteen times. In each trial, we hold out one of the fifteen populations as the unknown one and run our inferential tests on each of the fifteen generated network sizes. Results are similar across all trials and the fifteen generated network sizes correlate ~.95. Given that the trials yield similar results, we select the trial with the "number of people known in prison" as the unknown one to present as the main analyses in this paper. Finally, we ran the Bayesian estimation process ten times where in each trial we randomly remove two of the fifteen populations to find out whether results are sensitive to different subset of prompts. Here too, results are mostly consistent to those presented.

Results

How Large Are Extended Social Networks?

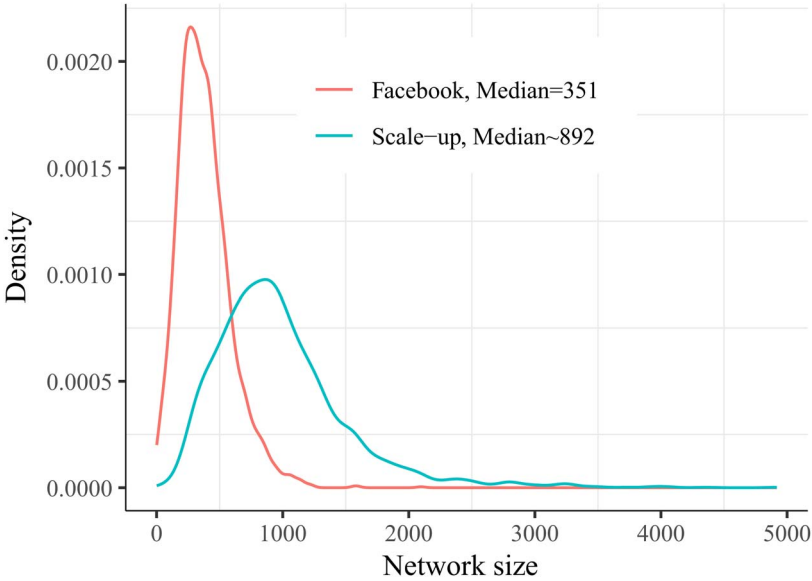
We first describe the obtained extended network sizes, then test our hypotheses, and finally consider differences between our two metrics. Table 3 compares the scale-up and Facebook network size. The median extended scale-up social network size is approximately 892, compared to a median number of Facebook contacts of approximately 351 and they correlate moderately ($r = .314$; $p < .001$). The scaling factor from the Facebook to the scale-up network size is 2.54, which approximates the 2.83 found by Hampton et al. (2011b). Figure 3 depicts density distributions for both metrics, resembling DiPrete et al.' (2011: 1254) plot of the number of acquaintances among Americans. The median network size of 892 of the scale-up method is substantially higher compared to that of estimates using similar methods. Prior estimates of extended network sizes using the scale-up method are in the range of 472–610 (Zheng et al. 2006; McCormick et al. 2010; DiPrete et al. 2011; Lubbers et al. 2019). Perhaps this is because our sample exclusively includes young adults (18–19 year olds) that might be more socially active across wider arrays of foci compared to adults. Findings using early behavioral data from before 2009—from the microblogging

Table 3. Descriptive Statistics on the Extended Social Networks ($N = 2{,}546$)[a]

	25%[b]	50%	75%	Mean	SD/SE
Number of contacts on Facebook	236	351	491	381.710	206.710
Scale-up estimator (Bayesian)	630.985	892.744	1295.287	980.387	529.244

[a]Numbers are based on observations for respondents who maintained public Facebook profiles as well as who completed the network scale-up questions.
[b]These statistics are the twenty-fifth percentile, median, and seventy-fifth percentile of these variables in these data.

Figure 3. Density distributions for the extended social network size using the scale-up method and the number of Facebook friends ($N = 2{,}546$).



website Twitter or from Facebook—show an average number of social media contacts of 180–200. As these platforms grew more popular, the number of contacts grew accordingly. Our average Facebook network size (~371) is on par with that of recent work, ~420 by Tulin, Pollet, and Lehmann-Willenbrock (2018).

Who Has Larger Extended Networks?

We use Heckman selection models to test our hypotheses (Heckman 1979). Formally, this takes the following form. The linear regression equation is

$$y_j = \beta_0 + \beta_1 X_j + \cdots + \beta_k X_j + \varepsilon_{1j}, \tag{3}$$

where β_0 represents a constant, $\beta_1 X_j + \cdots + \beta_k X_j$ represents a vector of independent variables, and ε_{1j} represents the error term. The selection equation is

$$z_j Y + \varepsilon_{2j} > 0, \tag{4}$$

where the errors terms of both equations are allowed to correlate,

$$\text{cor}\left(\varepsilon_{1j}, \varepsilon_{2j}\right) = \rho. \tag{5}$$

In regression Equation (3), we regress both network sizes on our covariates. We adjust for ethnic background, gender, and educational track level in selection Equation (4). We cluster-correct standard errors for adolescents' school cluster in wave 2 (to adjust for similarities between students in the same school), but results do not vary by using robust, noncorrected, or cluster-corrected standard errors. Selection (yes/no) means "surviving" from the second wave of the survey to being observed with a network size on (1) Facebook and (2) on the scale-up metric. We use Full-Information Maximum-Likelihood estimation. Heckman selection models produce regression weights that are corrected for selection effects. It also produces a correlation between the errors of both equations. These correlations are .959 (Facebook) and .981 (scale-up). A Wald test of independent equations for both outcomes shows that $\rho \neq 0$ ($p < .001$). Ethnic minorities (Facebook, scale-up), boys (scale-up), and lower educated (scale-up) are indeed less likely to be observed (see Supplementary Material A5). These results imply outcome selectivity and justify our modeling strategy. It also implies downward/upward biased estimates for some groups if we do not model this selectivity. A classic example of describing such patterns is men who do not work are likely to have low salaries if they would work. Hence, observed wages of working men are biased upward. We see a similar pattern here: ethnic minorities (or boys, or lower educated) who are less likely to be observed have lower network sizes. Not modeling that upwardly biases observed minorities' (or boys', or of lower educated) network size.

Because the distributions of both outcomes are not unusually skewed (Skewness Facebook = 1.172; Skewness Scale-up = 1.778), to allow for intuitive interpretation of coefficients, and following DiPrete et al. (2011), we use linear regressions in the substantive Heckman equation (see Supplementary Material A4 for uncorrected results and Supplementary Material A5 for selection coefficients). Table 4 depicts the results of these analyses.

Opportunities and Homophily

As depicted in Table 4, adolescents who spend more time going out, in associations, and going to concerts (all at least $p < .05$) have a larger number of contacts both on Facebook and in the scale-up measure. The magnitude of these correlations are substantial. Specifically, a one-unit increase (e.g., from

Table 4. Maximum-Likelihood Estimation Results of the Extended Network Size Measured via the number of Facebook Contacts and the Scale-Up Method Using Heckman Selections[a]

	Facebook			Scale-up		
	Coefficient	SE[b]	*p*-value[c]	Coefficient	SE[b]	*p*-value[c]
Constant	168.852	80.729	0.036	−92.775	178.841	0.604
Foci (H1)						
Going out	46.807	4.104	<0.001	56.933	8.362	<0.001
Associations	17.046	3.310	<0.001	22.763	5.820	<0.001
Concerts	12.300	5.446	0.024	28.896	12.145	0.017
Family	1.255	3.698	0.734	9.897	8.480	0.243
Religious meeting places	−3.150	3.849	0.413	18.516	7.244	0.011
Job (yes/no)	29.586	6.397	<0.001	6.117	14.347	0.670
Similarity of contacts (H2 + H3)						
Ethnic minority (reference majority)	−99.704	18.413	<0.001	−125.339	38.182	0.001
Co-ethnic Class	1.576	0.795	0.047	0.949	2.144	0.658
Co-ethnic School	−0.009	0.010	0.367	0.020	0.021	0.330
Romantic partner (yes/no) (H4)	29.741	7.384	<0.001	5.340	17.742	0.763
Education (H5)						
Vocational (reference)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Senior general	38.687	21.179	0.068	121.722	41.529	0.003
University prep.	−8.410	14.951	0.574	129.472	40.563	0.001
Girl (reference Boy) (H6)	30.768	9.619	0.001	144.340	20.572	<0.001

(Continued)

Table 4. ***Continued***

	Facebook			Scale-up		
	Coefficient	SE[b]	*p*-value[c]	Coefficient	SE[b]	*p*-value[c]
Confounders						
Early adopter of FB						
No (reference)	Ref.	Ref.	Ref.	–	–	–
Yes	53.946	5.533	<0.001	–	–	–
Nonparticipation Wave 2	32.401	19.978	0.105	–	–	–
Hours FB per day						
<1 h (reference)	Ref.	Ref.	Ref.	–	–	–
1–2 h	40.448	7.252	<0.001	–	–	–
2–3 h	53.702	11.204	<0.001	–	–	–
3–4 h	72.971	15.517	<0.001	–	–	–
>4 h	45.085	24.002	0.060	–	–	–
Nonparticipation Wave 2	9.015	8.087	0.265	–	–	–
Age in months	−1.317	0.352	<0.001	0.375	0.782	0.632
Observations	5,463			5,448		
Censored observations	2,776			2,328		
Uncensored observations	2,687			3,120		
Log pseudolikelihood	−21,214			−26,973		
ρ	0.961			0.981	–	

[a] In the selection equation we adjusted for ethnic background, gender, and educational level.
[b] Robust standard errors corrected for 112 (Facebook) and 113 (scale-up) school clusters.
[c] We take these *p*-values to use two-tailed tests when we consider the coefficients related to our hypotheses.

once a month to once a week) in going out, visiting associations, or going to concerts, adolescents gain approximately forty-six, seventeen, and twelve Facebook contacts, respectively. A one-unit increase in these same foci increases the number of contacts in the scale-up network with fifty-seven, twenty-three, and twenty-nine, respectively. Additionally, having a part-time job positively relates to the number of contacts on Facebook ($p < .001$), and going to religious meeting places more often positively relates to the scale-up network size ($p < .05$). These findings are mostly consistent with Hypothesis 1. Particularly leisure-orientated social foci relate to both outcomes, whereas Facebook ties relate to having a part-time job and people spending more time in religious places have larger scale-up network sizes. The latter finding is consistent with church attendance that positively relates to scale-up network sizes among US adults (DiPrete et al. 2011).

We also find evidence in support of our conjecture that Dutch majority members have larger social networks than ethnic minority members (H2). Specifically, ethnic minorities have 100 (Facebook) and 125 (scale-up) contacts less than Dutch majority members ($p < .01$).

We find moderate-to-no evidence for the conjecture that having more co-ethnic classmates (a) and schoolmates (b) positively relates to the extended network size (H3). When respondents have more co-ethnic classmates, their number of Facebook contacts is higher ($p < .05$).

Romantic Partners

We find that those who are in a romantic relationship have a larger extended network size on Facebook (H4). Those in a romantic relationship have approximately 30 Facebook contacts more than those we are not in a romantic relationship. We find no statistically significant relation between the scale-up extended network size and having a romantic partner.

Education and Gender

Those in higher educational tracks have larger networks compared to those in the lowest track in the scale-up measure of network size. Specifically, adolescents in the vocational educational track have approximately 122–130 contacts less compared to those in the Senior general or in the University preparatory track (both p-values $< .01$) as measured via the scale-up method. Overall, this is evidence in support of our conjecture that individuals in higher educational track levels have larger extended networks than those in lower educational track levels (H5), but only so among the scale-up network size.

The results suggest that girls have larger extended networks than boys, based on both on Facebook and in the scale-up measure, consistent with our intuition (H6). Specifically, girls seem to have approximately 31 (Facebook) and 144 (scale-up) contacts more than boys have (both $p < .001$).

Confounding Factors

We observe that early rather than late adopters of Facebook (<2010) and those who spend more time on Facebook have larger numbers of Facebook, consistent with our intuitions. Additionally, those adolescents who are younger have more Facebook contacts, consistent with prior findings (Lubbers et al. 2019).

Facebook Vis-à-Vis Scale-up

Next, we compare how our covariates relate to differences in the two metrics. The analyses presented in Table 5 regresses the difference between the Facebook and Scale-up network size (Facebook−Scale-up) on our covariates. Essentially, we show for which individuals the Facebook network size is closer to the scale-up network size. We present the Heckman selection model for parsimony, but an uncorrected, linear regression model provides similar results ($\rho \neq 0$ cannot be rejected). Some findings stand out compared to others. Specifically, the Facebook network size is closer to the size of the scale-up method for those with a part-time job ($p < .05$), girls ($p < .001$), those in higher educational tracks ($p < .01$), and early Facebook adopters ($p < .001$).

Conclusions

Social contacts lend support (Lubbers et al. 2019), advice (McPherson et al. 2006), and grant resourceful connections to unknown groups (Granovetter 1973). There are only few substantive papers on individual variation (see DiPrete et al. 2011; Lubbers et al. 2019) in the number of ties beyond the closest ones. We set out to address this and answered two questions. What explains individual variation in the extended network size? And can we provide some insight into the number of Facebook ties vis-à-vis the scale-up network size? Using large-scale data on thousands of Dutch adolescents and linking it to behavioral data from an online network, we answered these questions and contributed to recent work on variation in extended network sizes (DiPrete et al. 2011; Lubbers et al. 2019).

How large are extended social networks in our sample? We found a median of 351 Facebook contacts, whereas the median extended social network size using the scale-up method was 892. Our estimation of the scale-up network size is substantially higher compared to prior work using the scale-up method among adults, showing extended network sizes in the range of 472–610 (Zheng et al. 2006; McCormick et al. 2010; DiPrete et al. 2011; Lubbers et al. 2019). This discrepancy is perhaps due to our adolescent sample that may be more socially active. The scaling factor of the number of Facebook contacts to the network scale-up network size (2.54) approximates that by Hampton et al. (2011b) (2.83).

Beyond point estimates, however, investigating individual variation in network size provides insight into societal integration and group differences. So which adolescents have larger extended social networks? We turned to literature

Table 5. Maximum-Likelihood Estimation Results of the Differences between the Facebook and Scale-Up Network Size Using Heckman Selections[a]

	Facebook−Scale-up		
	Coefficient	SE[b]	*p*-value
Constant	154.340	383.446	0.687
Foci			
Going out	18.664	15.728	0.235
Associations	18.318	9.978	0.066
Concerts	31.500	20.042	0.116
Family	22.881	17.265	0.185
Religious meeting places	21.531	13.833	0.120
Job (yes/no)	−51.465	23.445	0.028
Similarity of cont.			
Ethnic minority (reference majority)	24.247	50.499	0.631
# Co-ethnic Class	−2.959	3.411	0.386
# Co-ethnic School	0.053	0.029	0.067
Romantic partner (yes/no)	−24.928	27.603	0.366
Education			
Vocational (reference)	Ref.	Ref.	Ref.
Senior general	−12.831	29.482	0.663
University prep.	−114.764	35.099	0.001
Girl (reference Boy)	−91.672	23.246	<.001

(Continued)

Table 5. ***Continued***

	Facebook—Scale-up		
	Coefficient	SE[b]	*p*-value
Confounders			
Early adopter of FB			
No (reference)	Ref.	Ref.	Ref.
Yes	52.364	19.577	0.007
Nonparticipation Wave 2	103.512	62.534	0.098
Hours FB per day			
< 1 h (reference)	Ref.	Ref.	Ref.
1–2 h	−4.972	25.173	0.843
2–3 h	15.767	38.052	0.679
3–4 h	21.966	72.334	0.761
> 4 h	98.922	72.804	0.174
Nonparticipation Wave 2	95.548	186.958	0.609
Age in months	1.268	1.634	0.438
Observations	5,468		
Censored observations	3,381		
Uncensored observations	2,087		
Log pseudolikelihood	−19,368		
ρ	−0.002		

[a]In the selection equation we adjusted for ethnic background, gender, and educational level.
[b]Robust standard errors corrected for 112 school clusters.

on tie-generating mechanisms among core ties (Blau 1977; Feld 1981). Consistent with fundamental prior work, we hypothesized and corroborated that those who spend more time in foci have larger extended networks, particularly with respect to social activities such as sporting associations or going out. We reaffirm the relevance of theories on participation in social space on the genesis of social ties among acquaintances.

In addition, we expected and found that adolescents who are ethnic minority members had smaller networks than those with Dutch backgrounds, consistent with prior work on adult racial minorities in the United States (DiPrete et al. 2011). Ethnic minorities seem to draw fewer rather than dissimilar ties from the opportunity set in their extended social networks. By and large, theories on the interplay between opportunities and homophily are central to explain the core as well as the extended network size.

Adolescents in a romantic relationship (only on Facebook), higher educated, and girls have larger network sizes than their counterparts, consistent with our hypotheses and previous work on adults (Van Tubergen et al. 2016). Differences in recall abilities or social activities may explain gender and educational differences in extended network sizes.

Finally, we identified differences between the scale-up and Facebook network size. Perhaps the two network sizes capture two distinct network layers in Dunbar's sense given their absolute difference (Dunbar et al. 2015) and some individuals' network layers overlap more than those of others. Girls, higher educated, those with a part-time job, and earlier Facebook adopters have network layers that are less distal. Some *post hoc* intuitions may substantiate these findings. Some factors (jobs, being on Facebook longer) may cause a larger part of the most-outer network layer (scale-up) to be considered important enough to be accumulated as Facebook contacts. Additionally, some groups (higher educated and girls) may find it important to have a larger part of that outer layer as Facebook contacts.

Several limitations to this study merit acknowledgement. First, data on a more-general target population would be beneficial. As of yet, however, we do not know of data that combine surveys—including a scale-up measurement—and data from online networks. Here, we tentatively assume that some network generation processes are equivalent between adults and adolescents. This may be reasonable, because our theoretical mechanisms are general in nature and not limited to either adult or adolescent populations. Because of this assumed equivalence, we view this study as a step between the few systematic studies on measuring and explaining the extended network size and future studies considering more-general target populations. Additionally, using a different set of scale-up prompts—for instance using Dutch names prevalent in older subpopulations—may decrease *point estimates* of the scale-up network size in our sample of youngsters (they know less of these X's).

Second, point estimates of the scale-up network size among adolescents were higher compared to those found among adult samples. Our conjecture is that this discrepancy exists because adolescents are more socially active than adults. Yet, because there are few studies that employ scale-up

methods among adolescents it is challenging to compare our estimate to an adolescents baseline estimate. Our high point estimate may result from some unknown issue with using scale-up modules among adolescents. In this study, we focused on variation in extended social network sizes. Even when an unforeseen scale-up/adolescent issue artificially alters our point estimate among adolescents, the variation between adolescents may persist qualitatively without that inflation. We commend future research into this issue.

Third, we have no longitudinal network data to model (potentially) causal relations or to unravel micromechanisms for the genesis of social ties. We commend future work that gathers multiple waves of behavioral data on the number of contacts on Facebook to study such processes more directly and is able to unravel explicit rather than implicit mechanisms; for instance, through gathering Facebook networks of adolescents and then applying exponential random graph modeling techniques (e.g., Wimmer and Lewis 2010).

Fourth, limiting scale-up contacts exclusively to persons "known in the Netherlands" ignores transnational ties. However, our approach considers the "meaningful acquaintances" in the national context and it is methodologically imperative to have a reference population. Future endeavors could include scale-up questions where in each subsequent set respondents have to recall contacts in increasingly wider geographical regions: neighborhood, city, focal country, neighboring countries, and so forth. One can then compare size estimates across regions to pinpoint how they vary.

Fifth, a growing body of literature studies personality factors (extraversion/agreeableness) and network size (Selden and Goodie 2018). Our data did not include questions on such factors, even though engagement in social foci may be exogenous to them. We commend future research on both a range of social foci and personality factors to disentangle their correlations with extended network sizes.

Finally, we mostly tested existing explanations for the number of core contacts in the context of extended social network sizes. Future research could identify unique explanations to the extended social network size that do not apply to core networks, or *vice versa*. One such endeavor could be to gather information about core and extended networks. One can then investigate which covariates relate to the core and not the extended network size.

Notes

1. Increases in social isolation in McPherson et al. (2006) are debated and refuted (McPherson, Smith-Lovin, and Brashears 2009; Fischer 2009; Paik and Sanchagrin 2013).
2. 150 is the maximum number of stable relationships humans are argued to be able to manage (Dunbar 1998; Kanai et al. 2012)—coined "Dunbar's Number." There is high variation among point estimates of the distinct

network layers (5, 15, etc.); Pollet et al. (2011) find that the support group size (mean = 7) ranges from 0 to 25.

3. People vary in what they consider important discussion topics (Bearman and Parigi 2004). The use of core discussion ties to measure strong ties is contested. Small et al. (2013, 2015) demonstrate that discussion partners are knowledgeable on the discussion topic or readily available in social contexts, but these persons are not always strong ties.
4. Hypotheses 5 and 6 may contradict when education is considered a power/status characteristic. Among teenagers power dynamics surrounding gender are experienced early and immediately matter during upbringing/adolescence. Educational strata, in contrast, only recently started to play a role among these adolescents. Moreover, status among adolescents might be higher among those in lower educational tracks. Hence, status/power differences along educational lines plays a role later in life and differently compared to gender.
5. The collection and use of the DFS was approved by an ethical review board for the social and behavioral sciences.
6. The standard error for the basic scale-up estimator is (McCormick et al. 2010: 60):

$$\mathrm{SE}\left(\text{Scale}-\text{up degree}_i\right) = \sqrt{\text{Scale}-\text{up degree}_i} \times \sqrt{\frac{1-\sum_{k=1}^{K} N_k/N}{\sum_{k=1}^{K} N_k/N}} \quad (6)$$

Supplementary Data

Supplementary data mentioned in the text are available to subscribers in *Social Forces* online.

About the Authors

Bas Hofstra is a Postdoctoral Research Fellow at Stanford University. He holds a PhD in Sociology from Utrecht University. His work orbits the study of diversity, stratification, and culture, often studied through computational methods and big data. It captures longitudinal systems of social and cultural exchange: from the gestation and birth of networks, careers, and ideas, to their use, up until their eventual cessation. His work appeared in, among others, *American Sociological Review* and *Social Networks*.

Rense Corten is associate professor at the Department of Sociology and the ICS. His research revolves around the themes of cooperation, trust, and (the dynamics of) social networks, with empirical applications including adolescent networks, social media, the sharing economy, online criminal networks, and laboratory experiments.

Frank van Tubergen is Professor of Sociology at the Department of Sociology, Utrecht University. He is a member of the ICS and the European Academy of

Sociology. His research topics are social networks, religion, and immigration. He recently published the textbook *Introduction to Sociology* (Routledge).

Acknowledgements

We acknowledge Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results. We thank Tyler H. McCormick for generously providing code for the network scale-up method from McCormick et al. (2010). Furthermore, early versions of this paper benefited from valuable feedback of Manja Coopmans, Jesper Rözer, Robert Krause, Jan Kornelis Dijkstra, Jeroen Weesie, and Vincent Buskens. We also thank seven Social Forces reviewers who provided insightful comments. Finally, this study benefited from discussions at the "1st International CILS4EU User Conference" in Mannheim.

Funding

This research benefited from the support of the NORFACE research program on Migration in Europe—Social, Economic, Cultural and Policy Dynamics and from the support of several grants from the Dutch Scientific Organization (NWO): NWO onderzoekstalent [406-12-004], NWO middelgroot [480-11-013], and NWO veranderingsstudies [481-11-004].

Conflict of interest

The authors declare no conflict of interest

References

Assendorpf, Jens B. and Susanne Wilpers. 1998. "Personality Effects on Social Relationships." *Journal of Personality and Social Psychology* 74(6):1531–1544.

Bastani, Susan 2007. "Family Comes First: Men's and Women's Personal Networks in Tehran." *Social Networks* 29:357–74.

Bearman, Peter and Paolo Parigi. 2004. "Cloning Headless Frogs and Other Important Matters: Conversation Topics and Network Structure." *Social Forces* 83(2):535–557.

Blau, Peter M. 1977. "A Macrosociological Theory of Social Structure." *American Journal of Sociology* 83(1):26–54.

Bloothooft, Gerrit and David Onland. 2011. "Socioeconomic Determinants of First Names." *Names: A Journal of Onomastics* 59(1):25–41.

Brashears, Matthew E., Emily Hoagland and Eric Quintane 2016. "Sex and Network Recall Accuracy." *Social Networks* 44:74–84.

Burt, Ronald S. 1984. "Network Items and the General Social Survey." *Social Networks* 6(4):293–339.

Currarini, S., Jackson, M.O., Pin, P. 2010. "Identifying the Roles of Race-Based Choice and Chance

in High School Friendship Network Formation." *Proceedings of the National Academy of Sciences* 107(11):4857–4861.

De Sola Pool, Ithiel and Manfred Kochen 1978/1979. "Contacts and Influence." *Social Networks* 1:5–51.

DiPrete, Thomas A., Andrew Gelman, Tyler H. McCormick, Julian Teitler, and Tian Zheng. 2011. "Segregation in Social Networks Based on Acquaintanceship and Trust." *American Journal of Sociology* 116(4):1234–1284.

Duggan, Maeve, Nicole B. Ellison, Cliff Lampe, Amanda Lenhard and Mary Madden 2015. *Social Media Update 2014* Washington, DC: Pew Research Center.

Dunbar, Robert I.M. 1998. "The Social Brain Hypothesis." *Evolutionary Anthropology* 6(5):178–190.

Dunbar, Robert I.M. 2016. "Do Online Social Media Cut through the Constraints that Limit the Size of Offline Social Networks." *Royal Society Open Science* 3:150292.

Dunbar, Robert I.M., Valerio Arnaboldi, Marco Conti and Andrea Passarella 2015. "The Structure of Online Social Networks Mirrors those in the Offline World." *Social Networks* 43:39–47.

Ellison, Nicole B., Steinfield, Charles, Lampe, Cliff. 2011. "Connection Strategies: Social Capital Implications of Facebook-Enabled Communication Practices." *New Media & Society* 13(6):873–982.

Feehan, Dennis M. and Matthew J. Salganik. 2016. "Generalizing the Network Scale-Up Method: A New Estimator for the Size of Hidden Populations." *Sociological Methodology* 46(1):153–186.

Feld, Scott L. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86(5):1015–1035.

Fischer, Claude S. 2009. "The 2004 GSS Finding of Shrunken Social Networks: An Artifact?" *American Sociological Review* 74(4):657–69.

Freeman, L. and C.L. Thompson 1989. "Estimating Acquantanceship Volume." In *The Small World*, edited by Kochen, M., pp. 147–58. Norwood, NJ: Ablex Publishing Corporation.

Gonçalves, Bruno, Nicola Perra, and Allessandro Vespignani. 2011. "Modeling Users' Activity on Twitter Networks: Validation of Dunbar's Number." *PLoS ONE* 6(8):e22656.

Gurevich, Michael 1961. *The Social Structure of Acquaintanceship Networks*. Cambridge: Massachusetts Institute of Technology, Dissertation.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78(6):1360–1380.

Halberstam, Yosh and Brian Knight 2016. "Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter." *Journal of Public Economics* 143:73–88.

Hampton, Keith H., Lauren F. Sessions, and Eun Ja Her. 2011a. "Core Networks, Social Isolation, and New Media." *Information, Communication & Society* 14(1):130–155.

Hampton, Keith H., Laurent F. Sessions Goulet, and Kristen Purcell, K. 2011b. "*Social Networking Sites and Our Lives*." (Pew Research Center), www.pewinternet.org/2011/06/16/social-networking-sites-and-our-lives/ (accessed September 19, 2019),

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153–161.

Hofstra, Bas, Rense Corten and Frank van Tubergen 2015. *Dutch Facebook Survey: Wave 1 [dataset and codebook]*. Department of Sociology/ICS: Utrecht University.

Hofstra, Bas, Rense Corten, and Frank van Tubergen. 2016a. "Who Was First on Facebook? Determinants of Early Adoption among Adolescents." *New Media & Society* 18(10):2340–2358.

Hofstra, Bas, Rense Corten and Frank van Tubergen 2016b. "Understanding the Privacy Behavior of Adolescents on Facebook: The Role of Peers, Popularity and Trust." *Computers in Human Behavior* 60:611–20.

Hofstra, Bas, Rense Corten, Frank van Tubergen, and Nicole B. Ellison. 2017. "Sources of Segregation in Social Networks: A Novel Approach Using Facebook." *American Sociological Review* 82(3):625–656.

Holt-Lunstad, Julianne, Timothy B. Smith, and J. Bradley Layton. 2010. "Social Relationships and Mortality Risk: A Meta-Analytic Review." *PLoS Medicine* 7(7):e1000316.

Hill, Russel A. and Robin I.M. Dunbar. 2003. "Social Network Size in Humans." *Human Nature* 14(1):53–72.

Jaspers, Eva and Frank van Tubergen 2014. "Children of Immigrants Longitudinal Survey in the Netherlands (CILSNL) - Wave 4." Full Version v4.0.0. DANS.

Jaspers, Eva and Frank van Tubergen 2017. "Thematic Collection: Children of Immigrants Longitudinal Survey in the Netherlands (CILSNL)." DANS.

Kadushin, Charles 2012. *Understanding Social Networks: Theories, Concepts, and Findings*. Oxford: Oxford University Press.

Kalmijn, Matthijs 1998. "Intermarriage and Homogamy: Causes, Patterns, Trends." *Annual Review of Sociology* 24:395–421.

Kalmijn, Matthijs. 2003. "Shared Friendship Networks and the Life Course: An Analysis of Survey Data on Married and Cohabiting Couples." *Social Networks* 25(3):231–249.

Kalter, Frank, Anthony Heath, Miles Hewstone, Jan O. Jonsson, Matthijs Kalmijn, Irena Kogan and Frank van Tubergen 2015. *Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU)* ZA5353, Data file version 2.1.0. Cologne [distributor]: GESIS Data Archive.

Kanai, Ryota, Bahador Bahrami, Rebecca Roylance and G. Geraint Rees 2012. "Online Social Network Size Is Reflected in Human Brain Structure." *Proceedings of the Royal Society B* 1–8.

Killworth, Peter D., Eugene C. Johnsen, Christopher McCarty, Gene Ann Shelley and H. Russel Bernard 1998a. "A Social Network Approach to Estimating Seroprevalence in the United States." *Social Networks* 20:23–50.

Killworth, Peter D., McCarty Christopher, H. Russel Bernard and Eugene C. Johnsen 1998b. "Estimation of Seroprevalence, Rape, and Homelessness in the United States Using a Social Network Approach." *Evaluation Review* 22:289–308.

Lubbers, Miranda J., José Luis Molina and Hugo Valenzuele-García 2019. "When Networks Speak Volumes: Variation in the Size of Broader Acquaintanceship Networks." *Social Networks* 56:55–69.

Maltiel, Rachael, Adrian E. Raftery, Tyler H. McCormick, and Aaron J. Baraff. 2015. "Estimating Population Size Using the Network Scale-Up Method." *The Annals of Applied Statistics* 9(3):1247–1277.

Maltiel, Rachael and Aaron J. Baraff. 2015. "NSUM: Network Scale Up Method. R package version 1.0." Retrieved from: https://CRAN.R-project.org/package=NSUM. (accessed July 5 2018).

Marsden, Peter V. 1987. "Core Discussion Networks of Americans." *American Sociological Review* 52(1):122–131.

Mayer, Adalbert, Puller, Steven L. 2008. "The Old Boy (and Girl) Network: Social Network Formation on University Campuses." *Journal of Public Economics* 92(1–2):329–47.

McCarty, Christopher, Peter D. Killworth, Russel H. Bernard, Eugene C. Johnsen, and Gene Ann Shelley. 2001. "Comparing Two Methods for Estimating Network Size." *Human Organization* 60(1):28–39.

McCormick, Tyler H., Matthew J. Salganik, and Tian Zheng. 2010. "How Many People Do you Know? Efficiently Estimating Personal Network Size." *Journal of the American Statistical Association* 105(489):59–70.

McPherson, Miller, Lynn Smith-Lovin and James M. Cook 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–44.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2006. "Social Isolation in America: Changes in Core Discussion Networks over Two Decades." *American Sociological Review* 71(3):353–375.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2009. "Models and Marginals: Using Survey Evidence to Study Social Networks." *American Sociological Review* 74(4):670–681.

Meertens Institute. 2016. "De Nederlandse Voornamenbank," http://www.meertens.knaw.nl/nvb/ (accessed September 6, 2016)

Mollenhorst, Gerald, Beate Volker, and Henk Flap. 2008. "Social Contexts and Personal Relationships: The Effect of Meeting Opportunities on Similarity for Relationships of Different Strength." *Social Networks* 30(1):60–68.

Moore, Gwen. 1990. "Structural Determinants of Men's and Women's Personal Networks." *American Sociological Review* 55(5):726–735.

Paik, Athony and Kenneth Sanchagrin. 2013. "Social Isolation: An Artifact." *American Sociological Review* 78(3):339–360.

Parigi, Paolo and I.I. Warner Henson 2014. "Social Isolation in American." *Annual Review of Sociology* 40:153–71.

Peterson, Richard A. 1992. "Understanding Audience Segmentation: From Elite and Mass to Omnivore and Univore." *Poetics* 21(4):243–258.

Pollet, Thomas V., Sam G.B. Roberts, and Robin I.M. Dunbar. 2011. "Use of Social Network Sites and Instant Messaging Does Not Lead to Increased Offline Social Network Size, or to Emotionally Closer Relationships with Offline Network Members." *Cyberpsychology, Behavior, and Social Networking* 114(4):253–258.

Pool, Ithiel de Sola and Manfrew Kochen 1978. "Contacts and Influence." *Social Networks* 1:5–51.

Roberts, Sam G.B., Ruth Wislon, Pawel Fedurek, Robin I.M. Dunbar. 2008. "Individual Differences and Personal Social Network Size and Structure." *Personality and Individual Differences* 44(4):954–964.

Rözer, Jesper Jelle, Gerald Mollenhorst and Beate Volker 2015. "Romantic Relationship Formation, Maintenance and Changes in Personal Networks." *Advances in life course research* 23:86–97.

Schimmer, Paulien and Frank van Tubergen. 2014. "Transnationalism and Ethnic Identification among Adolescent Children of Immigrants in the Netherlands, Germany, England, and Sweden." *International Migration Review* 48(3):680–709.

Selden, Mary and Adam S. Goodie 2018. "Review of the Effects of Five Factor Model Personality Traits on Network Structures and Perceptions of Structure." *Social Networks* 52:81–99.

Selfhout, Maarten, William Burk, Susan Branje, Jaap Denissen, Marcel van Aken, and Wim Meeus. 2010. "Emerging Late Adolescent Friendship Networks and Big Five Personality Traits: A Social Network Approach." *Journal of Personality* 78(2):509–538.

Simpson, Brent, Barry Markovsky and Mike Steketee 2011. "Power and the Perception of Social Networks." *Social Networks* 33:166–71.

Slater, Philip E. 1963. "On Social Regression." *American Sociological Review* 28(3):339–364.

Small, M.L. 2013. "Weak Ties and the Core Discussion Networks: Why People Regularly Discuss Important Matters with Unimportant Alters." *Social Networks* 35(5):470–483.

Small, Mario L., Vontrese D. Pamphile and Peter McMahan 2015. "How Stable Is the Core Discussion Network?" *Social Networks* 40:90–102.

Smith, Sanne, Ineke Maas, and Frank van Tubergen. 2012. "Irreconcilable Differences? Ethnic Intermarriage and Divorce in the Netherlands, 1995-2008." *Social Science Research* 41(5):1126–1137.

Statistics Netherlands. 2012. "Migranten, Vreemdelingen en Vluchtelingen: Begrippen op het Terrein Van Asiel en Buitenlandse Migratie," http://www.cbs.nl (accessed June 16, 2015).

Statistics Netherlands.2015. CBS Statline: Bevolking; kerncijfer, http://statline.cbs.nl (accessed March 23, 2015)

Subrahmanyam, Kaveri, Stephani M. Reich, Natalia Waechter, Guadalupe Espinoza. 2008. "Online and Offline Social Networks: Use of Social Networking Sites by Emerging Adults." *Journal of Applied Developmental Psychology* 29(6):420–33.

Van Tubergen, Frank. 2014. "Size and Socio-Economic Resources of Core Discussion Networks in the Netherlands: Differences by National-Origin Group and Immigrant Generation." *Ethnic and Racial Studies* 37(6):1020–1042.

Van Tubergen, Frank, Obaid A. Al-Modaf, Nora F. Almosaed and Mohamed B.S. Al-Ghamdi 2016. "Personal Networks in Saudia Arabia: The Role of Ascribed and Achieved Characteristics." *Social Networks* 45:45–54.

Tulin, Marina, Thomas V. Pollet and Nale Lehmann-Willenbrock 2018. "Perceived Group Cohesion Versus Actual Social Structure: A Study Using Social Network Analysis of Egocentric Facebook Networks." *Social Science Research* 74:161–75.

Wimmer, Adreas and Kevin Lewis. 2010. "Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook." *American Journal of Sociology* 116(2):583–642.

Van Zalk, Maarten H. W., Nejra van Zalk, Margaret Kerr, and Hakan Stattin. 2014. "Influences between Online-Exclusive, Conjoint and Offline-Exclusive Friendship Networks: The Moderating Role of Shyness." *European Journal of Personality* 28(2):134–46.

Sapiezynski, Piotr, Arkaduisz Stopczynski, David K. Wind, Jure Leskovec and Sune Lehmann 2018. "Offline Behaviors of Online Friends." *arXiv* arXiv.1811.03153.

Zheng, Tian, Matthew T. Salganik, and Andrew Gelman. 2006. "How Many People Do you Know in Prison?" *Journal of the American Statistical Association* 101(474):409–423.