American Sociological Review
2017, Vol. 82(3) 625–656

DOI: 10.1177/0003122417705656
journals.sagepub.com/home/asr

Sources of Segregation in Social Networks: A Novel Approach Using Facebook

Bas Hofstra,[a] Rense Corten,[a] Frank van Tubergen,[b] and Nicole B. Ellison[c]

Abstract

Most research on segregation in social networks considers small circles of strong ties, and little is known about segregation among the much larger number of weaker ties. This article proposes a novel approach to the study of these more extended networks, through the use of data on personal ties in an online social network. We illustrate this method's potential by describing and explaining the degree of ethnic and gender segregation on Facebook among a representative survey of adolescents in the Netherlands (N = 2,810; ~1.1 million Facebook friends). The results show that large online networks are more strongly segregated by ethnicity than by gender. Drawing on the same survey data, we find that core networks are more segregated in terms of ethnicity and gender than are extended networks. However, an exception to this pattern is personal networks of ethnic majority members, whose core networks are as segregated by ethnicity as their extended networks. Further analysis suggests this exception is due to their larger population size and the ethnic segregation of their social settings. We discuss the implications of these findings for the role of structural opportunities, homophily, and balance.

Keywords

segregation, ethnicity, gender, social networks, Facebook

One of the most consistent findings in sociological research is that strong-tie, core friendship networks tend to be homogeneously sorted (McPherson, Smith-Lovin, and Cook 2001). Network cleavages among strong ties are formed along ethnic, gender, religious, and social status lines. This finding appears in research on romantic relationships (Anderson et al. 2014; Feliciano, Robnett, and Komaie 2009; Kalmijn 1998; Lewis 2013; Potârcă and Mills 2015), core discussion networks (Marsden 1988; Smith, McPherson, and Smith-Lovin 2014), and personal friendship networks (Currarini, Jackson, and Pin 2010; Mouw and Entwisle 2006; Smith, Maas, and Van Tubergen 2014; Vermeij, Van Duijn, and Baerveldt 2009; Wimmer and Lewis 2010).

In contrast to the abundant literature on the segregation of core networks, little is known about the segregation of weaker ties, such as

[a]Utrecht University
[b]Utrecht University and King Abdulaziz University
[c]University of Michigan

Corresponding Author:
Bas Hofstra, Utrecht University, Department of Sociology/ICS, Padualaan 14, 3584 CH Utrecht, the Netherlands
E-mail: b.hofstra@uu.nl

those involving colleagues, neighbors, and acquaintances (DiPrete et al. 2011; Moody 2001). For various reasons, however, it is particularly important to study segregation among weaker ties, as they relate to a myriad of sociologically relevant issues. A classic argument is that such weak ties provide novel information on job openings and hence link to labor-market outcomes and the societal distribution of wealth (Granovetter 1973, 1983; Lin 1999). Second, when not only core networks but also people's extended networks are homogenous—whether in terms of ethnicity, race, religion, gender, or other characteristics—intergroup trust might be undermined (Fukuyama 1995; Gambetta 1988), and negative intergroup attitudes may prevail, as they are not challenged by personal encounters (Allport 1954). A rich body of literature suggests that even superficial contact (i.e., weak ties) between members of different ethnic groups has the potential to reduce intergroup prejudice (Pettigrew and Tropp 2006). Furthermore, the co-evolution of homophilous network selection and social influence can result in "echo chambers"—in which people are increasingly surrounded by like-minded people (Halberstam and Knight 2016)—and intergroup polarization of opinions and attitudes (Mäs and Flache 2013).

The lack of research on diversity among weaker ties is mainly due to methodological difficulties of gathering information that includes both strong and weak ties. To our knowledge, the only study on segregation among weak ties is by DiPrete and colleagues (2011). Using the 2006 General Social Survey (GSS), they found that Americans' "acquaintanceship" networks (i.e., weak ties) are approximately as segregated as their "trust" networks (i.e., core ties) along racial, political, and religious lines. They measured weak ties using network scale-up methods, in which respondents were asked to estimate the number of people with whom they are acquainted along racial, ideological, and religious lines in various contexts (e.g., neighborhoods). This method suffers from two limitations. First, although people are asked to think about their acquaintances in specific contexts, there may be selectivity in network recall ability (Brashears, Hoagland, and Quintane 2016). Second, the results may be affected by social desirability biases or "misperception or masking of behaviors and opinions that Americans think would be disapproved of by their associates" (DiPrete et al. 2011:1272).

We propose that the study of *online* social networks provides new opportunities to examine the segregation of large personal networks, including networks with stronger *and* weaker ties. An important advantage over the scale-up method is that online networks map networking *behavior* up to potentially thousands of contacts, without restrictions to specific contexts. As such, online networks are less prone to recall bias and other misperceptions.

Social media has experienced a remarkable rise to prominence over the past decade and is increasingly used to maintain interpersonal relationships (Ellison and boyd 2013). Initially, the popular belief was that such platforms, of which Facebook is the prime example, "would open up the vista of a social world that was intrinsically unlimited in size" (Dunbar et al. 2015:39). These online networks could supposedly be used to mitigate the social segregation typically found in offline friendship networks (Rainie and Wellman 2012; Robinson et al. 2015), as users are not restricted to exclusively befriending others from the schools they attend, from the neighborhoods in which they live, or from their offline activities.

Recent empirical research contradicts such beliefs about the promise of social media. Instead, this work shows that large online networks are indicative of an individual's complete set of *offline* relationships. For instance, early studies among U.S. college students found that only .4 percent of online friendships are online-only (Mayer and Puller 2008), and college students use online networks to maintain and strengthen their offline relationships but rarely to initiate new contacts (Ellison, Steinfield, and Lampe 2007, 2011). Moreover, 80 percent of social network site users state that they use such social platforms to stay in touch with their offline ties (Subrahmanyam et al. 2008). Similarly, 77 percent of adolescents report that they befriend others online only when they have met them offline (Reich, Subrahmanyam, and Espinoza 2008). In a study of Swedish adolescents, 77 percent of friends were friends both offline and online (Van Zalk et al. 2014). U.S. adults indicate that their online network friends are predominantly family, current and past friends, neighbors, and colleagues (Duggan et al. 2015). Furthermore, the structure and size of offline and online networks are similar (Dunbar 2016; Dunbar et al. 2015). Finally, studies conducted among U.S. college students find similar high levels of segregation by ethnicity and race on Facebook as is found on campuses (Lewis, Gonzalez, and Kaufman 2012; Lewis et al. 2008; Mayer and Puller 2008; Wimmer and Lewis 2010).

The overlap between online and offline social networks provides a number of opportunities for sociological research: it enables the study of a large portion of overall personal networks, including small samples of core ties and far larger numbers of weaker ties. Empirically, it is relatively easy to collect data on online interactions, such as the data documented on Facebook, as the platforms generate time-stamped, digital footprints of all their users' relationships (Golder and Macy 2014).

To illustrate this new approach to the study of segregation in social networks, we consider the Facebook networks of adolescents in the Netherlands. Because prior research shows that adolescents' strong-tie offline networks are highly segregated in terms of both *ethnicity* (e.g., Baerveldt et al. 2004; Vermeij et al. 2009) and *gender* (e.g., Lubbers 2003; Shrum, Cheek, and Hunter 1988; Smith and Schneider 2000), we first consider whether we find similar patterns of ethnic and gender segregation among the larger number of contacts in online networks.

Second, we examine the conditions under which ethnic and gender segregation of the extended network occur. Because previous research focuses exclusively on tie formation and segregation among core ties, there is little empirical evidence of the determinants of segregation among larger sets of network ties. In this study, we take the first steps to provide such evidence. As our theoretical point of departure, we engage classic network theories that are commonly used to explain segregation among core ties. Specifically, we consider the role of relative group size (Blau 1977a, 1977b), foci (Feld 1981, 1982, 1984), homophily (Byrne 1971; Lazarsfeld and Merton 1954), and balance (Heider 1946; Krackhardt and Handcock 2007), and we show their relevance in explaining segregation among hundreds of social relationships. As such, we contribute to the understanding of processes that underlie segregation in large networks while simultaneously testing existing, fundamental hypotheses in novel ways.

Third, we explore the differences in segregation between core networks and larger networks, because some scholars speculate there may be disparities in segregation among core and weaker ties (e.g., Blackwell and Lichter 2004; Granovetter 1973, 1983; Mollenhorst, Volker, and Flap 2008; Putnam 2000; Son and Lin 2012), although few studies have empirically studied this. Our study is among the first to elaborate and empirically test the conditions and mechanisms that create differences in the levels of segregation among core networks and larger networks.

ONLINE SOCIAL NETWORKS

We used a general survey of Dutch adolescents (Kalter et al. 2013, 2015) and linked these data to respondents' online Facebook networks in 2014. In doing so, we provide novel and detailed knowledge of respondents, their core networks, and their larger online networks. Previous studies of online social network segregation have used selective samples of Facebook friendships from U.S. colleges in 2005 and 2006 (Mayer and Puller 2008; Wimmer and Lewis 2010). At present, Facebook friends represent a wide range of social ties, such as family members, friends, and neighbors, and Facebook is the largest social network site worldwide, with approximately one billion daily users (Facebook 2015). Among adolescents in the Netherlands, Facebook is the most popular social network site; over 95 percent of Dutch adolescents have an account (Hofstra, Corten, and Van Tubergen 2016a).

Social networks emerge on Facebook when users send friendship invitations to other users, who can accept or decline the invitation. An accepted invitation shows an undirected, reciprocated friendship between two users. On Facebook, all relationships displayed on *friend lists* are indistinguishable with regard to tie strength (Lewis et al. 2008). Nevertheless, considerable evidence suggests that the number of "best" friends does not exceed 5 to 10 people. Research also suggests that a cognitive limit prevents a person from maintaining more than approximately five *close* relationships (e.g., Dunbar et al. 2015; Roberts et al. 2009; Zhou et al. 2005). In addition, 95 percent of Americans reported fewer than six confidants (i.e., core ties) in the 1984 and 2004 GSS (McPherson, Smith-Lovin, and Brashears 2006).[1] The average of 336 friends in our online network data thus suggests that most of the sample's online friends are weak rather than strong ties. At the very least, we capture a large portion of an individual's complete personal network; thus, we clearly go beyond the small number of core social ties. This assumption is supported by estimates suggesting that depending on the methods used, the average size of *overall* personal networks range from 150 (Hill and Dunbar 2003) to 750 (Zheng, Salganik, and Gelman 2006) contacts.[2]

THEORY AND HYPOTHESES

A substantial literature examines how offline social ties form and why network segregation occurs (e.g., Blau 1977a; Centola 2015; Feld 1981; Kalmijn 1998; Kossinets and Watts 2009; McPherson et al. 2001; Mouw and Entwisle 2006; Wimmer and Lewis 2010). Common explanations for the genesis of social segregation are relative group size, foci, homophily, and balance. We therefore focus on these factors.

Following Wimmer and Lewis (2010:588), we use the term *homophily* for the tie-generating mechanism, which indicates a preference for the selection of similar friends, and we use the term *segregation* to describe the composition of a network. For clarity, we use the term homophily to indicate what is commonly called "choice" homophily (e.g., McPherson and Smith-Lovin 1987; McPherson et al. 2001), that is, homophily net of meeting opportunities or other structural processes ("baseline" homophily). In this study, we examine segregation in social networks with regard to the ethnic and gender homogeneity found in personal social networks.

Meeting Opportunities: Relative Group Size and Foci

Theoretically and empirically, meeting opportunities are important in predicting strong tie formation (e.g., Kalmijn and Flap 2001; Mollenhorst, Volker, and Flap 2008, 2014; Mouw and Entwisle 2006; Smith, McPherson, and Smith-Lovin 2014; Vermeij et al. 2009; Wimmer and Lewis 2010). Two key dimensions of meeting opportunities are relative group size and foci (Blau 1977a, 1977b; Feld 1981); we consider these factors because we expect they drive segregation in large online networks.

Relative group size. The relative size of a group is an important factor in friendship formation (Blau 1977a, 1977b). Levels of personal network segregation may reflect the distribution of social categories in a population. For instance, when a society consists of 20 percent minority members and 80 percent majority members, individuals' network contacts—of both majority and minority members—will consist of 20 percent minority and 80 percent majority members under the condition of random mixing.

In the Netherlands, ethnic groups' relative size varies, whereas the distribution of men and women is approximately 50/50 (Statistics Netherlands 2015). Approximately 79 percent of people are "Dutch majority" members (Statistics Netherlands 2015). In contrast, ethnic minority groups, who have an immigrant background, are much smaller in size. For instance, minority members with a Moroccan background compose approximately 3 percent of the Dutch population.

Given these differences, we first compare ethnic and gender segregation in online networks. If relative group size is important in explaining segregation, we would expect ethnic segregation in large personal networks to be higher than gender segregation, as the distribution of majority and minority populations is more unequal than the gender distribution. Large online networks will thus reflect these unequal distributions in the population. Considering these disparities at the level of the population at large, we propose the following hypothesis:

> *Hypothesis 1a:* Larger online networks are more homogeneous by ethnicity than they are by gender.

Second, we compare ethnic segregation between ethnic majority and minority members. When people belong to a large group, they have ample opportunities to meet members of their own group, whereas members of smaller groups are likely to develop many ties outside their own group (likely from the majority group). Therefore, we expect ethnic segregation in online networks is higher for members of the Dutch majority group than for members of an ethnic minority. We thus propose the following hypothesis:

> *Hypothesis 1b:* Members of the ethnic majority have more homogeneous online networks than do members of the ethnic minority.

Foci. Along with groups' relative size in a population, we consider *foci* and their role in friendship formation (Feld 1981) and segregation. A focus is defined as "a social, psychological, legal, or physical entity around which joint activities are organized" (Feld 1981:1016). Social contexts can be represented as sets of different foci and individuals. Individuals engage in a number of different foci but not in all of them. Two individuals who engage in the same focus are thus more likely to share activities than are two individuals who do not share a focus. Sharing foci creates "positive sentiments indirectly through the generation of positively valued interaction" (Feld 1981:1017). Foci bring people together in mutually rewarding situations, and individuals form ties among others on whom they spend resources, such as time and emotions. Sharing a focus therefore increases the likelihood for a (friendship) tie to emerge (i.e., in the consideration of positive ties).

What aspects of foci foster dyadic similarity between individuals? Foci themselves are segregated because there is selectivity of specific groups to participate and enroll in particular foci (Feld 1981; Feld and Carter 1999). Hence, whereas a group's size relative to other groups is an important factor in friendship formation, these groups spread and organize in social settings in a nonrandom way. Therefore, personal networks will resemble the structural features of foci; that is, people who develop ties within foci will likely resemble one another.

Many empirical accounts illustrate that foci are segregated. In the United States and Europe, schools and school classes vary in their racial-ethnic compositions (Mouw and Entwisle 2006; Smith, Maas, and Van Tubergen 2014; Vermeij et al. 2009), and U.S. and European neighborhoods and cities tend to be racially and ethnically segregated (Lichter, Parisi, and Taquino 2015; Semyonov and Glikman 2009). Accordingly, scholars have found not only that many relationships are formed in the context of some sort of focus (e.g., Grosetti 2005), but also that homogeneity in foci fosters segregation in personal networks (e.g., Feld 1982, 1984; Kalmijn and Flap 2001; Mollenhorst et al. 2014). This is occasionally called "inbreeding" homophily (McPherson et al. 2001).

We consider schools and classrooms to be major foci for tie formation among adolescents (McPherson et al. 2001), because adolescents spend a considerable amount of their time in these settings. These two settings do not capture *all* the foci of adolescents, and parts of the relative group size effect may be attributed to the nonrandom sorting of adolescents over foci that we do not capture in this study. Nevertheless, not *all* relationships originate from foci, because people may "meet 'by chance' or as a result of adjacency along some continuum" (Feld 1981:1018).

Foci effects are often found in segregation among core networks (e.g., Feld 1982, 1984), and foci have similar effects on the dyadic similarity of friend *and* acquaintanceship networks that are measured by name-generating questions (Mollenhorst et al. 2008). Therefore, we assume that the foci mechanism does not vary by tie strength. We thus derive the following hypothesis:

> *Hypothesis 2:* Ethnic and gender homogeneity in schools and classrooms predicts the ethnic and gender homogeneity of online networks.

The Interplay between Meeting Opportunities, Homophily, and Balance

Some scholars have suggested that core networks are more strongly segregated than are extended networks. Granovetter (1973:1362), for example, states that "the stronger the tie connecting two individuals, the more similar they are" and "homophilous ties are more likely to be strong" (Granovetter 1983:210). Putnam (2000:20) similarly speculates that strong relationships, which constitute "bonding" social capital, are more likely to exist among similar people, whereas weak ties, which create "bridging" social capital, are more likely to exist among dissimilar people. Son and Lin (2012:602) argue that people with "stronger ties are more likely to share . . . commonalities" and as ties become weaker, "the ties' characteristics become dissimilar—more diverse." What follows is an explanation of the conditions that create differences in homogeneity among stronger and weaker ties, with a focus on meeting opportunities, homophily, and balance.

Two mechanisms suggest that dyadic similarity correlates with tie strength, and hence that core networks are more segregated than extended networks. The first is the homophily mechanism. According to the homophily argument, people generally prefer to befriend others similar to themselves (Byrne 1971; Lazarsfeld and Merton 1954; McPherson et al. 2001). Homophily exists along multiple dimensions, such as gender, ethnicity, race, education, or religion. Individuals may develop a *psychological* preference for similar friends (Byrne 1971), which represents an enhanced degree of psychological attraction between two similar entities (Lewis 2015). Homophily may be driven by shared cultural norms and beliefs (Smith, Maas, and Van Tubergen 2014), because shared norms can decrease the costs of investing in relationships (it takes less time to get to know one another) and increase returns on the investment (it becomes easier to interact) (Kalmijn 1998).

Given that people have ample opportunities to select same-gender and same-ethnic friendships, homophily may be more pronounced among core ties than among weaker ties (Mollenhorst et al. 2008). One reason for this phenomenon is that stronger ties are "costly" (Windzio and Bicer 2013), because strong ties involve more time, emotional intensity, intimacy, and reciprocal services (Granovetter 1973:1361), whereas "cheaper," weaker ties deplete fewer such resources. Therefore, if possible, individuals are more likely to strengthen their relationships with similar rather than dissimilar others (Leszczensky and Pink 2015; Windzio and Bicer 2013). Individuals perceive relationships in which they share commonalities with others to be more rewarding and less risky. People expect stable returns on investments in such relationships: it is easier to interact, and it takes less time to get to know one another because there are fewer (cultural) boundaries to overcome. Hence, dyadic similarity promotes tie strength.

The second reason why dyadic similarity would be associated with tie strength comes from the network balance mechanism. Assuming that similar dyads are more likely to be strongly connected than are dissimilar dyads, triadic closure (when *A* is friends with

B, and *A* with *C*, then *B* and *C* are likely to connect) may occur more often among similar than among dissimilar individuals.[3] When ties are strong, unbalanced network configurations produce psychological strain for actors (Heider 1946), which leads them to close the "forbidden" triad (Granovetter 1973). Furthermore, an individual who has two strong ties in a triad provides opportunities for the unconnected pair to befriend each other (Feld 1981; Mollenhorst, Volker, and Flap 2011). Additionally, the dyadic survival of a relationship in an "isolated dyad" is lower than that of a dyad embedded in a triad, due to group identity formation, group pressure, and conflict control. These group dynamics are more likely to emerge within an embedded dyad, which creates an increased probability of triadic closure (Feld 1997; Krackhardt and Handcock 2007). Among embedded dyads characterized by a strong relationship, these dynamics may be even stronger, as these actors may more strongly call upon the group's identity and norms.

The composition of an individual's core friendship network, which is often limited to approximately five persons (e.g., Marsden 1988; McPherson et al. 2006; Smith, McPherson, and Smith-Lovin 2014), may thus be affected by homophily and balance, more so than weaker ties (which are initially formed by opportunity). In the opportunity set of network contacts, a person may have at least five similar available people with whom relationships can be strengthened. Individuals' larger networks, however, will be more likely to reflect the structural features of the population and foci. Initially, network ties mirror the features of meeting opportunities. Over time, however, ties characterized by dyadic similarity may transition into stronger ties, whereas dissimilar dyads may remain in their existing state of loosely connected weaker ties.

To examine this empirically, we first consider the number of friends individuals have in their online network. We assume that when people create online social network accounts, they start by adding close friends and contacts. This process is similar to name generators, in which people mention their closest ties first (Marin 2004). Additionally, Facebook promotes network closure: it prompts people to become friends with the friends of their friends, which also makes it more likely that an individual's first friends on Facebook will be strong ties. When the number of online social network friends increases, an increasing number of them will likely be weaker ties.

These factors should result in lower levels of ethnic and gender homogeneity in larger online networks, because the relative number of strong ties is lower. However, we expect that the negative association between the number of online network friends and ethnic homogeneity pertains only to members of ethnic minorities. The opportunity set of potential contacts is often shaped such that ethnic majority members are overrepresented in public life and in foci. Therefore, members of larger ethnic majority groups have limited opportunities to befriend people from smaller groups: there are fewer such persons in the population and in the foci. Among ethnic majorities, this means core networks and larger networks largely comprise majority group members. Minority group members, in contrast, meet many dissimilar others (likely of the majority group). Although they may strengthen their relationships with the few similar minority members whom they meet (because of homophily and balance), their larger network will continue to resemble the structural features of the meeting opportunities. We therefore expect individuals with larger online networks to have lower levels of ethnic and gender homogeneity—the exception being ethnic majority groups, as the opportunities for meeting co-ethnics are so widespread for this group. Specifically, we propose the following hypotheses:

Hypothesis 3a: As online network size increases, ethnic homogeneity decreases, but only among ethnic minorities.

Hypothesis 3b: As online network size increases, gender homogeneity decreases.

We also provide a different test of the same arguments by directly contrasting ethnic and gender homogeneity among small, self-reported core networks with ethnic and gender homogeneity found in large online personal networks. Instead of examining the number of connections only in online networks, we compare core and large online networks directly. We thus hypothesize the following:

> *Hypothesis 3c:* Core networks have more ethnic homogeneity than do larger online networks, but only among ethnic minorities.

> *Hypothesis 3d:* Core networks are more gender homogeneous than the larger online networks.

DATA AND MEASURES

We use the second wave of survey data on adolescents in the Netherlands, which is part of a larger project titled "Children of Immigrants Longitudinal Survey in Four European Countries" (CILS4EU) (Kalter et al. 2013, 2015).[4] Although data were collected in the Netherlands, Sweden, Germany, and England, the measures we are interested in are included only in the Dutch data. In the CILS4EU, adolescents age 14 to 15 years were followed for three years, starting in 2010, with a one-year time lag. The survey included data on many individual characteristics, attitudes, and information about the individuals with whom respondents associated with in their leisure time. The survey also contained sociometric data on friendships within classrooms (~22 pupils in a classroom). The sample was stratified by the proportion of non-Western immigrants within a school. Within these strata, schools were chosen with a probability proportional to their size (based on the number of pupils at the relevant educational level).

In wave 1 (2010 to 2011), two classes were randomly selected within the schools, which resulted in 118 schools, 252 classes, and 4,963 Dutch pupils participating in the survey.[5] Because changes in class compositions between grades are common in the Netherlands, respondents were distributed among different classes in wave 2 (2011 to 2012) that were not part of the original sampling frame. To ensure that many wave 1 pupils also participated in wave 2, schools were asked to include more than the two classes initially sampled in wave 1 when respondents from wave 1 were in classes other than the previously sampled classes. Consequently, 2,118 new pupils were interviewed, and 3,803 of wave 1 respondents were surveyed again in wave 2 (76.6 percent; total $N = 5,921$). We used the second wave of the CILS4EU because it is the latest licensed data including sociometric classroom information.

The Dutch Facebook Survey

The Dutch Facebook Survey (DFS) enriched the Dutch part of the CILS4EU survey (Hofstra, Corten, and Van Tubergen 2015).[6] Data were collected between June and September 2014. Of the 4,864 respondents who indicated Facebook membership in wave 3 (2012 to 2013; $N = 3,423$) or 4 (2013 to 2014; $N = 3,595$) of the CILS4EU, 4,463 were tracked on Facebook. For respondents who kept a *public friend list*, we downloaded their complete Facebook friend lists ($N = 3,252$; 72.8 percent). There is selectivity in the downloaded friend lists: some respondents kept their lists private, others kept public friend lists. Girls, ethnic minority members, and unpopular adolescents are somewhat underrepresented, because they more often keep private friend lists (Hofstra, Corten, and Van Tubergen 2016b). Various Heckman selection-model specifications (Heckman 1979) show that our results are insensitive to these selection biases.[7] The 3,252 respondents have a combined total of 1,158,227 friends, and 2,810 (86.4 percent) of the respondents whose complete friend list we downloaded also participated in wave 2 of the CILS4EU.[8] This is the number of respondents for whom we present results.[9] Table 1 summarizes the data sources and our method of arriving at the final number of respondents.

Table 1. Overview of the Relevant Data Sources and Selections

	N	%
Survey data (CILS4EU)		
Wave 2 total number of respondents	5,921	100
Wave 2 respondents participated in wave 1	3,803	64.2
Wave 2 respondents who are newcomers	2,118	35.8
Online network data (DFS)		
Respondents indicated being on Facebook in waves 3 or 4 of the survey	4,864	100
Respondents whose profiles were tracked on Facebook	4,463	91.8
Respondent kept a public Facebook friend list	3,252	66.9
Conditions for inclusion in the final number of cases to analyze		
Participation wave 2 + Tracked on Facebook + Kept a public Facebook friend list	2,810[a]	

[a]Various Heckman selection model specifications show that our results are insensitive to selection biases.

Measuring Ethnic and Gender Homogeneity in Online Networks

There is no direct measure for friends' ethnic background and gender in the Facebook network. We predicted friends' gender and ethnic background based on their first names,[10] using the Dutch Civil Registration data (hereafter, DCR) for the entire Dutch population in 2010 (N = 15,785,208; Bloothooft and Schraagen 2011). We obtained (1) the fraction of the name carriers and (2) the fraction of the name carriers' fathers and (3) mothers who were born in the Netherlands, Turkey, Morocco, the Dutch Caribbean, other Western countries, or other non-Western countries. Additionally, we obtained the percentage of women among the name carriers.

We matched first names in the DCR to first names in the second wave of the CILS4EU survey as a training dataset. In the CILS4EU, we measured respondents' ethnic background by classifying them into one of the six largest ethnic groups in the Netherlands (Castles, De Haas, and Miller 2013): Dutch majority, Turkish, Moroccan, Dutch Caribbean, other Western (European or English speaking), and other non-Western. Moroccan and Turkish adolescents are children of immigrants from the low-educated labor force that was recruited by the Netherlands in the 1950s and 1960s. Dutch Caribbean adolescents originate from post-colonial countries in the Dutch Caribbean (e.g., Aruba and Suriname). Western and non-Western adolescents originate from neighboring countries such as Germany or conflict areas such as Afghanistan; these immigrant groups are relatively similar across Western European countries (Smith, Maas, and Van Tubergen 2014).

We classified respondents according to their biological parents' country of birth, which is standard practice in research on Dutch ethnic minority groups (cf. Smith, Maas, and Van Tubergen 2014; Stark and Flache 2012; Vermeij et al. 2009). When students have one parent who was born in the Netherlands, the student is classified into the ethnic background of the parent who was not born in the Netherlands; if a student's parents were born in different non-Dutch countries, the student is classified according to the mother's birth country. This definition is regularly applied and used by Statistics Netherlands (Statistics Netherlands 2012).

Combining the DCR and the CILS4EU, we developed an algorithm to estimate gender and ethnic segregation based on people's first names, which yields high correlations between the predicted and actual ethnicity and gender (this method is outlined in Part A of the online supplement). We calculated the percentage of women and the percentage of each of the six ethnicities in respondents' online networks. For each respondent, we assigned the percentage of same-gender friendships (i.e., the percentage of women for girls and percentage of men for boys) in their online networks.[11] Finally, we assigned each respondent the percentage of co-ethnic ties in their online networks (e.g., the percentage of Dutch majority members among online network friends for the Dutch majority adolescents).

Homogeneity in Core Friendship Networks

Wave 2 of the CILS4EU has two measures that capture ethnic homogeneity and one measure that captures gender homogeneity in core friendship networks: a name generator for the five best friends *in general* (only for ethnicity), and a name generator for the five best friends *in class* (not necessarily the same friends as the former).

First, we measured the actual number of friends of Dutch, Turkish, Moroccan, Dutch Caribbean, or another ethnic background using a name-generator question. Respondents could nominate their best friends (with a maximum of five) and provide ethnic background information. From these data, we calculated the percentage of co-ethnic friends among all the close friends (co-ethnic$_{\text{FRIENDS IN GENERAL}}$). We consider *ethnically similar* friends among best friends in general. Respondents may be more accurate in reporting ethnicities of ethnically similar than ethnically dissimilar friends. Furthermore, respondents were asked to report the ethnicities of their *best friends*. Respondents may more accurately report the ethnicities of their best friends than those of acquaintances. Therefore, respondents' misreporting of alter characteristics is likely reduced to a minimum.

Second, we measured the number of best friends in a class (with a maximum of five) who were girls (which is the only core-network measure available for measuring gender homogeneity) and those who were of Dutch, Turkish, Moroccan, Dutch Caribbean, other Western, or other non-Western ethnic backgrounds. We calculated the percentage of co-ethnic friends and the percentage of same-gender friends among all friends in a class (co-ethnic$_{\text{FRIENDS IN CLASS}}$ and same-gender$_{\text{FRIENDS IN CLASS}}$). Because these friends themselves were respondents in the survey, they self-reported their gender and ethnicity. We constructed gender and ethnic homogeneity with respect to best friends in a class with these self-reports, and hence they do not suffer from respondents' misperceptions in alter characteristics.

Homogeneity in Meeting Opportunities and Number of Online Network Friends

We constructed various measures to capture ethnic and gender homogeneity in two adolescent opportunity structures, the class and the school. First, using the CILS4EU, we measured the number of *classmates* who are female and those with the six ethnic backgrounds mentioned above, excluding best friends who are mentioned in the class and respondents themselves. We calculated the percentage of same-gender and co-ethnic classmates, and we excluded the respondent and the number of best friends who are mentioned. We excluded best friends because they are included in the core-network measure, and we do not want to double-count best friends across variables. With this approach, we are better able to separate the effects between variables (same-gender$_{\text{IN CLASS}}$ and co-ethnic$_{\text{IN CLASS}}$).

Second, we measured the number of female pupils in a school (aggregated from the classes surveyed) and the number of pupils in the school from a Dutch, Turkish, Moroccan, Dutch Caribbean, other Western, or other non-Western ethnic background (measured from secondary data obtained from the Dutch inspectorate), excluding best friends who are mentioned, other classmates, and the respondent. We calculated the percentage of same-gender schoolmates (excluding the respondent, the number of best friends who are mentioned, and the number of classmates) (same-gender$_{\text{IN SCHOOL}}$). We also measured the percentage of co-ethnic schoolmates (excluding the respondent, the number of best friends, and the number of classmates) (co-ethnic$_{\text{IN SCHOOL}}$). We measured these two variables using the CILS4EU.

We also calculated the number of online network friends from respondents' Facebook friend lists using the DFS. The distribution of the number of online network friends, if it is plotted, strongly resembles the distribution plot reported by DiPrete and colleagues (2011:1254) of the number of acquaintances reported by Americans.

Kinship Ties in Online Networks as a Confounding Factor

An issue with online versus offline friendship networks is that we restricted respondents to name *friends* in their self-reported core networks offline, whereas Facebook networks likely include *kin*. Therefore, when we contrast core networks with online networks, we compare two data sources of different sampling frames. Kinship ties in online networks might pull ethnic and gender homogeneity in different directions. Kin likely have a similar ethnicity as the respondent, whereas the gender distribution in families is likely to be 50/50. On the one hand, the presence of kin in online networks overestimates ethnic homogeneity; on the other hand, the presence of kin among Facebook friends might lead us to underestimate gender homogeneity (see McPherson et al. 2001:431).

We identify kinship ties in online networks in two ways, using the DFS. First, Facebook allows members to show kinship tags on their profiles. We tracked the number of kinship tags on Facebook profiles and calculated the percentage of kinship tags in the Facebook network (mean = 1.1 percent). We considered realistic tags (e.g., no granddaughters, given that we study adolescents). Individuals might not tag each family member on Facebook. Therefore, we calculated the percentage of friends in the Facebook network who share a surname with the respondent (mean = 1.7 percent). Non-kin friends may have a similar surname, which makes our analyses more conservative, because individuals with similar surnames are likely of the same ethnicity. Nevertheless, we may miss kin in online networks who are not tagged and who have different surnames. We mention where we remove kin from the online networks to avoid sampling mismatches (descriptive comparisons between core networks and the larger online networks) and where we control for these two variables (statistical tests of the hypotheses).[12]

Table 2 shows the descriptive statistics for ethnic and gender homogeneity in the large online networks (including kin), core networks, opportunity structures, and kinship ties in online networks, along with the distributions of boys and girls and ethnic groups in the data.

Additional Confounding Factors

We adjust for the year in which respondents joined Facebook using the DFS (median = 2010). Respondents who were members for shorter periods may have been more selective in their online network friendships. Facebook membership duration and the number of Facebook friends are positively correlated ($r = .250$; $p < .001$).

Using the CILS4EU, we also control for educational track in high school, because such a track may be related to ethnic prejudice (Lancee and Sarrasin 2015). When adolescents transition to high school in the Netherlands, they are placed into different tracks, which differ in their level and type of education. We measured this categorization using three dummy variables: preparatory vocational education (N = 1,358; Dutch: VMBO), senior general (N = 750; Dutch: HAVO), and university preparatory education (N = 586; Dutch: VWO). We also control for respondents' social attractiveness, which may be correlated with ethnicity (Wimmer and Lewis 2010). We measured social attractiveness by *popularity* (i.e., incoming popularity nominations from other classmates) (mean = 9.357; SD = 14.566). We calculated *popularity* by dividing the total number of classmates' received nominations for popularity by the total number of students in the class minus one multiplied by 100.[13]

Table 2. Descriptive Statistics of Ethnic and Gender Homogeneity in Large Online Networks, Opportunity Structures, Kinship Ties on Facebook, and the Distribution of Boys and Girls and Ethnic Background

	Min.	Max.	Mean	SD	*N*
Online networks[a]					
Co-ethnic$_{\text{FACEBOOK}}$	0	100	76.577	32.099	2,810
Same-gender$_{\text{FACEBOOK}}$	0	100	56.087	9.745	2,809
% Female	0	100	49.453	11.475	2,810
% Dutch	0	100	86.200	15.670	2,810
% Turkish	0	100	2.304	7.649	2,810
% Moroccan	0	59.460	1.729	5.015	2,810
% Dutch Caribbean	0	54.237	1.347	3.097	2,810
% Other Western	0	57.142	3.176	2.566	2,810
% Other non-Western	0	75.676	4.234	5.025	2,810
Core networks					
Co-ethnic$_{\text{FRIENDS IN GENERAL}}$	0	100	76.218	33.730	2,810
Co-ethnic$_{\text{FRIENDS IN CLASS}}$	0	100	67.525	38.249	2,677
Same-gender$_{\text{FRIENDS IN CLASS}}$	0	100	83.175	30.227	2,677
Opportunity structures					
Co-ethnic$_{\text{IN CLASS}}$	0	100	65.710	31.743	2,690
Co-ethnic$_{\text{IN SCHOOL}}$	0	100	67.038	30.926	2,763
Same-gender$_{\text{IN CLASS}}$	0	100	50.212	22.163	2,690
Same-gender$_{\text{IN SCHOOL}}$	0	100	47.428	18.328	2,638
Number of online network friends	1	1,067	336.853	177.702	2,810
Kinship ties on Facebook					
% Kinship ties declared	0	20	1.081	1.555	2,794
% Similar surname on Facebook	0	100	1.689	3.303	2,794
Girl	0	1	.515		2,809
Ethnic background					2,810
Dutch	0	1	.804		2,258
Turk	0	1	.020		57
Moroccan	0	1	.015		42
Dutch Caribbean	0	1	.023		65
Other Western	0	1	.088		247
Other non-Western	0	1	.050		141

[a]These estimates of homogeneity in Facebook networks include kin.

We adjust for ethnic out-group attitudes because they may be related to ethnic homogeneity in online networks. With the survey question, "Please rate how you feel towards the following groups…" respondents used a scale ranging from 0 (negative) to 100 (positive), with 10-point intervals, to rate how positively they feel toward groups of Dutch, Turkish, Moroccan, and Dutch Caribbean ethnic backgrounds. We constructed ethnic out-group attitudes by taking the mean positivity score—on a scale from 0 to 10—of respondents' answers to this question while excluding the respondent's own ethnic group (mean = 5.011; SD = 1.997). This variable is significantly negatively related to the percentage of co-ethnic friends online ($r = -.213$; $p < .001$).

We accounted for respondents' attitudes toward gender roles when we considered gender homogeneity in online networks. We captured respondents' progressiveness toward gender roles by counting (from zero to four) how many times respondents indicated that both men and women (instead of men *or* women) should take care of children, cook, earn money, and clean (Davis and Greenstein 2009) ($\alpha = .73$; mean = 2.689; SD = 1.352). This variable is significantly negatively

Table 3. Ethnic Homogeneity in Large Online Networks and in Core Networks

	Min.	Max.	Mean	SD	*N*
Co-ethnic$_{\text{FACEBOOK}}$[a]	0	100	75.974	32.099	2,792
Co-ethnic$_{\text{FRIENDS IN GENERAL}}$	0	100	76.218	33.729	2,810
Co-ethnic$_{\text{FRIENDS IN CLASS}}$	0	100	67.525	38.249	2,677

[a]This estimate excludes kinship ties on Facebook

related to the percentage of same-gender friends online ($r = -.072$; $p < .001$).[14, 15]

RESULTS

Ethnic and Gender Homogeneity in Online Networks

In 2014, Dutch adolescents' online social networks had, on average, 76.6 percent co-ethnic friends. If everyone connected at random on Facebook in the Netherlands, the average personal network would consist of 78.6 percent Dutch, 2.4 percent Turks, 2.2 percent Moroccans, 2.9 percent Dutch Caribbean, 9.5 percent other Western individuals, and 4.3 percent individuals with other non-Western ethnic backgrounds. However, on average, the online networks in our sample consist of 86.2 percent Dutch, 2.3 percent Turks, 1.7 percent Moroccans, 1.4 percent Dutch Caribbean, 3.2 percent other Western individuals, and 4.2 percent individuals with other non-Western backgrounds (see Table 2).

Table 3 shows the ethnic homogeneity of core networks and online networks, and Table 4 shows these results broken down by ethnicity (both tables exclude kinship ties in online networks). The percentage of co-ethnic friends online is 76 percent; in core networks it is 76.2 percent for friends in general (co-ethnic$_{\text{FRIENDS IN GENERAL}}$) and 67.5 percent for friends in a class (co-ethnic$_{\text{FRIENDS IN CLASS}}$). The correlations between co-ethnic friendships in core and larger online networks are high: .784 (co-ethnic$_{\text{FRIENDS IN GENERAL}}$) and .677 (co-ethnic$_{\text{FRIENDS IN CLASS}}$). Dutch majority members have the highest ethnic homogeneity online, 91.7 percent (co-ethnic$_{\text{FACEBOOK}}$), which resembles the homogeneity in core networks (co-ethnic$_{\text{FRIENDS IN GENERAL}}$ = 88.4 percent). The ethnic homogeneity of Turkish adolescents (co-ethnic$_{\text{FACEBOOK}}$ = 40.6 percent) is slightly higher than that of Moroccan adolescents (co-ethnic$_{\text{FACEBOOK}}$ = 28.5 percent). Ethnic homogeneity in online networks (~336 friends) mirrors ethnic homogeneity in core networks (~5 friends).

Table 5 shows the gender homogeneity of core networks and online networks broken down by gender (excluding kinship ties). On average, respondents have 56.3 percent same-gender friendships online. If everyone connected at random, such that the percentage of same-gender friendships on Facebook reflected the gender composition at the societal level, this number should be approximately 50 percent (Statistics Netherlands 2015). On average, adolescents reported 83.2 percent same-gender friends in a class. Boys had slightly more same-gender friendships online than did girls (boys = 57.1 percent; girls = 55.5 percent), but boys had approximately the same percentage of same-gender friendships in a class as did girls (boys = 83 percent; girls = 83.4 percent).

Meeting Opportunities: Relative Group Size and Foci

Relative group size. We begin by examining the role of relative group size in homogeneity in online networks. We first evaluate the extent to which ethnic and gender homogeneity estimates in online networks differ from one another (Hypothesis 1a). Figure 1 shows the kernel density smoothed distributions for ethnic and gender homogeneity in the online networks, suggesting that online networks are more segregated by ethnicity than they are by gender.

We estimated an intercept-only multilevel model in which the intercept is the sample

Table 4. Ethnic Homogeneity in Large Online Networks and in Core Networks, Broken Down by Ethnicity

	Dutch	Turkish	Moroccan	Dutch Carib.	Other Western	Other non-Western
Co-ethnic$_{\text{FACEBOOK}}$[a]	91.569	40.604	28.455	9.176	3.167	13.397
Co-ethnic$_{\text{FRIENDS IN GENERAL}}$	88.412	54.503	45.198	27.692	14.899	28.759
Co-ethnic$_{\text{FRIENDS IN CLASS}}$	79.843	31.730	19.228	12.769	11.364	19.975

Note: For the percentages of specific ethnic backgrounds within online networks broken down by respondents' ethnic background (e.g., the percentage of Moroccans in Facebook networks of Dutch majority members), see Figure A1 in the Appendix.
[a]These estimates exclude kinship ties on Facebook.

Table 5. Gender Homogeneity in Large Online Networks and in Core Networks, Broken Down by Gender

	Min.	Max.	Mean	SD	*N*
Same-gender$_{\text{FACEBOOK}}$[a]	0	100	56.313	10.041	2,791
Same-gender$_{\text{FRIENDS IN CLASS}}$	0	100	83.175	30.227	2,677
Boys					
Same-gender$_{\text{FACEBOOK}}$	0	100	57.132	10.153	1,356
Same-gender$_{\text{FRIENDS IN CLASS}}$	0	100	82.952	30.144	1,299
Girls					
Same-gender$_{\text{FACEBOOK}}$	0	100	55.538	9.876	1,435
Same-gender$_{\text{FRIENDS IN CLASS}}$	0	100	83.384	30.314	1,378

[a]This estimate excludes kinship ties on Facebook.

mean difference between the percentage of co-ethnic and same-gender friends online. This model can be specified as follows:

$$Y_{ijk} = \beta_{000} + s_{0k} + c_{0jk} + p_{0ijk} \tag{1}$$

where Y_{ijk} is the difference in ethnic and gender homogeneity in online Facebook networks (co-ethnic$_{\text{FACEBOOK}}$ − same-gender$_{\text{FACEBOOK}}$) for respondent *i* from class *j* and school *k*; $s_{0k} \sim (0, \sigma^2_{s_{0k}})$ is the error term at the school level; $c_{0jk} \sim (0, \sigma^2_{c_{0jk}})$ is the error term at the class level; $p_{0ijk} \sim (0, \sigma^2_{p_{0ijk}})$ is the error term at the pupil level; and β_{000} is the sample mean difference in this *intercept only model* (the syntax for the analyses is found in Part C of the online supplement).[16,17] With these models, we control for class and school tendencies in the difference between ethnic and gender homogeneity in online networks (Snijders and Bosker 2012).

The intercept of the intercept-only model significantly deviates from zero (intercept = 16.8; $p < .001$; see Table A1 in the Appendix), suggesting that ethnic homogeneity is approximately 16.8 percent higher than gender homogeneity in online networks (Hypothesis 1a).

Second, we evaluate how the size of the ethnic majority group relative to the minority groups relates to ethnic segregation in online networks (Hypothesis 1b). We specified several multilevel regression models in order to estimate the percentage of same-ethnic friends within respondents' online networks.[18,19,20] We delete the missing values of the variables presented in these analyses listwise and lose 9.1 percent ($N = 261$) of cases in the analyses. These models can be specified as follows:

$$Y_{ijk} = \beta_{000} + \beta_{00x} X_i + s_{0k} + c_{0jk} + p_{0ijk} \tag{2}$$

where Y_{ijk} is the percentage of co-ethnic friendships in the online network for pupil *i* from class *j* and school *k*; s_{0k}, c_{0jk}, p_{0ijk}, and β_{000} specify similar terms as in Equation 1; and β_{00x} is a vector for the independent variables at the pupil level (e.g., ethnic background). Table 6 shows the model that estimates the percentage of co-ethnic friends

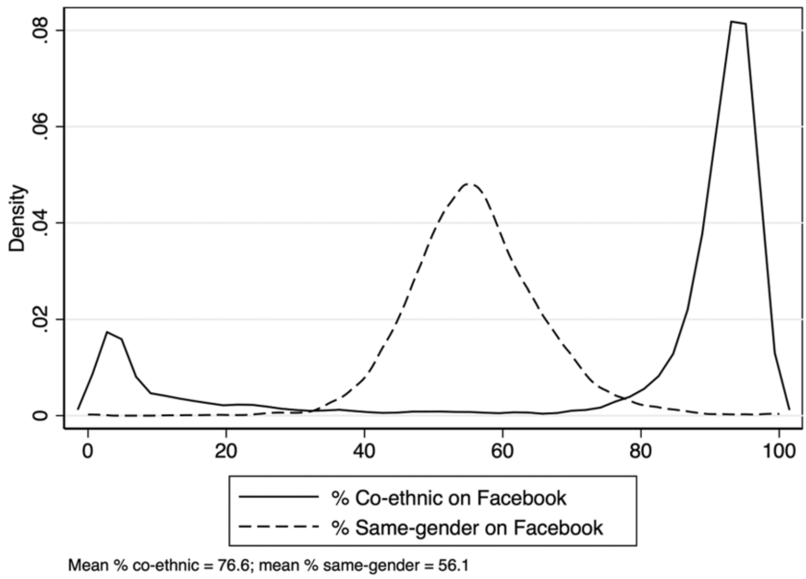

Figure 1. Density Plots of Ethnic and Gender Homogeneity in Large Online Networks

online for all respondents and separately for Dutch majority and ethnic minority members (while controlling for kinship ties in the online networks).

We find relatively large effects of ethnicity on ethnic homogeneity in online networks (while controlling for ethnic homogeneity in the class and school setting). Dutch majority-group adolescents seem to have at least 31 percent more co-ethnic friendships online than do students of other ethnic backgrounds ($p < .001$). For instance, adolescents of Moroccan ethnic background have 43.5 percent fewer, and those of Turkish descent 31.8 percent fewer, co-ethnic online networks friendships than do Dutch majority members. Among ethnic minority members, students of Turkish ethnic background exhibit more ethnic segregation in online networks than do all of their ethnic minority counterparts: they have at least 9 percent more co-ethnic friends in their online networks ($p < .01$). These variables reflect the propinquity of co-ethnic individuals in the population and relate to co-ethnic friendships in the large online networks. The results thus show that the larger majority group has significantly higher levels of ethnic homogeneity than do ethnic minority group members (consistent with Hypothesis 1b).

Foci. We now consider the extent to which the homogeneity of foci relates to homogeneity in online networks. We ask whether the percentage of co-ethnic and same-gender peers in class and in school is related to homogeneity in online networks (Hypothesis 2). In addition to the results shown in Table 6, we estimated a multilevel regression for the percentage of same-gender ties in online networks. This model takes the form of Equation 2, but here, Y_{ijk} specifies the percentage of same-gender friendships online. We delete the missing values of the variables presented listwise in this analysis and lose 7.8 percent ($N = 212$) of the cases. Table 7 shows results of this model (with kinship ties as control variables).

Table 6 shows that a two-standard-deviation increase in the percentage of co-ethnic classmates increases the percentage of co-ethnic friends online by 2.1 percent ($p < .01$). This relationship seems to be driven by Dutch majority members, because this variable is statistically significant for majority members ($p < .05$) but not for ethnic minority members ($p > .05$). Additionally, the percentage of co-ethnic schoolmates is associated with the percentage of co-ethnic friends in online networks for all respondents. A two-standard-deviation increase in the percentage of co-ethnic

Table 6. Multilevel Model Estimating the Percentage of Co-ethnic Friends in Online Networks

	All Respondents			Only Dutch Majority			Only Non-Dutch Minorities		
	Coef.	SE	*p*	Coef.	SE	*p*	Coef.	SE	*p*
Fixed part									
Intercept	67.935	(2.946)	***	68.491	(3.482)	***	32.998	(5.848)	***
Core-network									
Co-ethnic$_{\text{FRIENDS IN GENERAL}}$	.083	(.010)	***	.052	(.009)	***	.120	(.019)	***
Co-ethnic$_{\text{FRIENDS IN CLASS}}$	.008	(.004)	*	.008	(.004)	*	.004	(.020)	
Opportunity									
Co-ethnic$_{\text{IN CLASS}}$	.033	(.013)	**	.019	(.009)	*	.055	(.050)	
Co-ethnic$_{\text{IN SCHOOL}}$	.212	(.031)	***	.216	(.034)	***	.256	(.076)	***
Ethnicity									
Dutch	Ref.	Ref.	Ref.						
Turkish	–31.820	(3.617)	***				Ref.	Ref.	Ref.
Moroccan	–43.554	(3.249)	***				–9.398	(3.557)	**
Dutch Caribbean	–57.771	(2.760)	***				–22.427	(3.149)	***
Other Western	–61.696	(2.431)	***				–25.620	(3.093)	***
Other non-Western	–56.070	(2.044)	***				–21.088	(2.904)	***
Number of Facebook friends	–.002	(.001)	*	–.001	(.001)		–.007	(.003)	***
Facebook membership (ref.: 2013)									
2012	–1.836	(1.946)		–.472	(1.211)		–.426	(4.625)	
2011	–2.297	(1.783)		.417	(1.150)		–5.468	(4.113)	
2010	–2.018	(1.737)		.578	(1.123)		–4.313	(3.919)	
2009	–2.266	(1.756)		.166	(1.144)		–3.585	(4.044)	
2008	–1.447	(1.759)		.091	(1.158)		–.866	(4.085)	
2007	–2.571	(1.881)		–.446	(1.413)		–3.136	(4.652)	
2006	–1.141	(3.028)		–.569	(3.221)		2.532	(5.132)	
Girls (ref.: boys)	.043	(.250)		.124	(.149)		–.661	(.953)	
Educational track (ref.: lower voc.)									
Senior general	–.103	(.575)		–.265	(.480)		–1.370	(1.019)	
University preparatory	–1.188	(.578)	*	–.333	(.522)		–1.937	(1.086)	*
Indegree popularity	–.007	(.010)		–.010	(.008)		.015	(.028)	
Ethnic out-group attitudes	–.188	(.050)	***	–.195	(.042)	***	–.037	(.210)	
% Kinship ties declared	.063	(.118)		.004	(.067)		.202	(.416)	

(continued)

Table 6. *(continued)*

	All Respondents			Only Dutch Majority			Only Non-Dutch Minorities		
	Coef.	SE	*p*	Coef.	SE	*p*	Coef.	SE	*p*
% Similar surname on Facebook	.280	(.152)	*	.091	(.041)	*	.906	(.266)	***
Random part									
$\sigma^2_{s_{0k}}$ (school level)	5.246	(2.597)		7.027	(2.930)		.000	(.000)	
$\sigma^2_{c_{0jk}}$ (class level)	.000	(.000)		.108	(1.025)		.000	(.000)	
$\sigma^2_{p_{0ijk}}$ (pupil level)	31.552	(4.468)		13.878	(1.738)		91.807	(15.139)	
Number of schools	112			101			106		
Number of classes	309			278			233		
Number of pupils	2,549			2,053			496		
Log likelihood		–8092.940			–5726.875			–1824.677	

Note: Robust standard errors, adjusted for the school-identifier.
*p < .05; **p < .01; ***p < .001 (one-tailed tests).

Table 7. Multilevel Model Estimating the Percentage of Same-Gender Friends in Online Networks

	Coefficient	SE	p
Fixed part			
Intercept	58.901	(3.020)	***
Core-network			
Same-gender$_{\text{FRIENDS IN CLASS}}$	.036	(.005)	***
Opportunity			
Same-gender$_{\text{IN CLASS}}$	.039	(.010)	***
Same-gender$_{\text{IN SCHOOL}}$	.031	(.013)	**
Ethnicity (ref.: Dutch)			
Turkish	11.142	(1.668)	***
Moroccan	3.079	(2.985)	
Dutch Caribbean	–.942	(1.134)	
Other Western	.418	(.555)	
Other non-Western	4.090	(.830)	***
Number of Facebook friends	–.016	(.002)	***
Facebook membership (ref.: 2013)			
2012	–3.323	(2.855)	
2011	–3.092	(2.645)	
2010	–3.139	(2.635)	
2009	–3.345	(2.711)	
2008	–2.442	(2.795)	
2007	–2.500	(3.029)	
2006	–3.038	(4.052)	
Girls (ref.: boys)	–1.075	(.588)	*
Educational track (ref.: lower voc.)			
Senior general	1.062	(.529)	*
University preparatory	.006	(.563)	
Indegree popularity	–.014	(.013)	
Gender role attitudes	–.273	(.155)	*
% Kinship ties declared	.121	(.157)	
% Similar surname on Facebook	–.211	(.151)	
Random part			
$\sigma^2_{s_{0k}}$ (school level)	1.225	(.641)	
$\sigma^2_{c_{0jk}}$ (class level)	.000	(.000)	
$\sigma^2_{p_{0ijk}}$ (pupil level)	77.957	(4.366)	
Number of schools	109		
Number of classes	302		
Number of pupils	2,598		
Log likelihood		–9361.166	

Note: Robust standard errors, adjusted for the school-identifier.
$^*p < .05$; $^{**}p < .01$; $^{***}p < .001$ (one-tailed tests).

schoolmates increases the percentage of co-ethnic friends in online networks by 13.1 percent ($p < .001$). A two-standard-deviation increase in the percentage of same-gender classmates increases the percentage of same-gender friends online by 1.7 percent ($p < .001$). Additionally, a two-standard-deviation increase in the percentage of same-gender schoolmates increases the percentage of same-gender ties in online networks by 1.1 percent ($p < .001$). Given these results, we can conclude that the ethnic and gender composition of foci has a positive effect on the ethnic and gender homogeneity found in large online networks (consistent with Hypothesis 2).

The Interplay between Meeting Opportunities, Homophily, and Balance

Number of online network friends. We can now compare the difference in ethnic and gender homogeneity between strong versus weaker ties (Hypotheses 3c and 3d). Before doing that, we first consider the relationship between network size and ethnic and gender homogeneity in online networks (Hypotheses 3a and 3b).[21]

Table 6 shows that ethnic minority adolescents who have larger online networks also have a lower percentage of co-ethnic friends. For each 100 extra online network friends, the percentage of co-ethnic friends decreases by .7 percent. A two-standard-deviation increase in the number of friends decreases the percentage of co-ethnic online network friends by approximately 2.5 percent. For majority-group adolescents, the number of friends and the percentage of co-ethnic friends are not related ($p > .05$). The negative association between the number of friends and the percentage of co-ethnic friends is significantly stronger for minority members than for Dutch majority adolescents ($p < .001$; tested as the product of a dichotomous variable for Dutch/non-Dutch ethnic background and the number of friends). This result suggests that a larger online network coincides with lower ethnic homogeneity only among ethnic minority members (consistent with Hypothesis 3a).

Table 7 shows that when the number of online network friends increases, gender homogeneity decreases. A two-standard-deviation increase in the number of online network friends decreases the percentage of same-gender friends online by 5.7 percent ($p < .001$), which suggests that gender homogeneity is stronger among smaller online networks than among larger online networks (consistent with Hypothesis 3b).

Figures 2 and 3 depict the relationships between network size and ethnic and gender homogeneity in online networks. Figure 2 shows that the percentage of co-ethnic friendships online decreases when the number of friends online increases for ethnic minority adolescents. For ethnic minority adolescents, the number of friends in online networks is negatively correlated with the percentage of co-ethnic friends online ($r = -.343$; $p < .001$). Figure 3 shows that network size and the percentage of same-gender friendships in online networks is negatively correlated for both boys and girls ($r = -.312$; $p < .001$).

Self-reported core networks and online networks. Next, we contrast ethnic and gender homogeneity in self-reported core networks and larger online networks (Hypothesis 3c). To do so, we subtract ethnic homogeneity in online networks from ethnic homogeneity among friends in general (co-ethnic$_{\text{FRIENDS IN GENERAL}}$ – co-ethnic$_{\text{FACEBOOK}}$; results do not vary if we use co-ethnic$_{\text{FRIENDS IN CLASS}}$) and estimate the differences across ethnic groups between these factors in a multilevel regression model (see Table A2 in the Appendix). Members of minority groups have at least 26 percent more co-ethnic friends among their core friends than among their online network friends than do Dutch majority members ($p < .001$; we excluded self-reported core networks because these are part of the dependent variable in this model). Hence, we find evidence supporting our argument that ethnic minority members have higher levels of ethnic homogeneity in their core networks than in their larger networks, whereas we find no such association for ethnic majority members.

We also consider whether gender homogeneity in core networks is higher than in online networks (Hypothesis 3d). We estimate a multilevel model where the dependent variable is the difference between the percentage of same-gender friends in the core and online networks (same-gender$_{\text{FACEBOOK}}$ – same-gender$_{\text{FRIENDS IN CLASS}}$). The intercept statistically deviates from zero (intercept = 29.995; $p < .001$; see Table A3 in the Appendix; we excluded core networks because these are part of the dependent variable), which suggests that gender homogeneity among core ties, as measured from the survey, is approximately 30 percent higher than gender homogeneity in online networks,

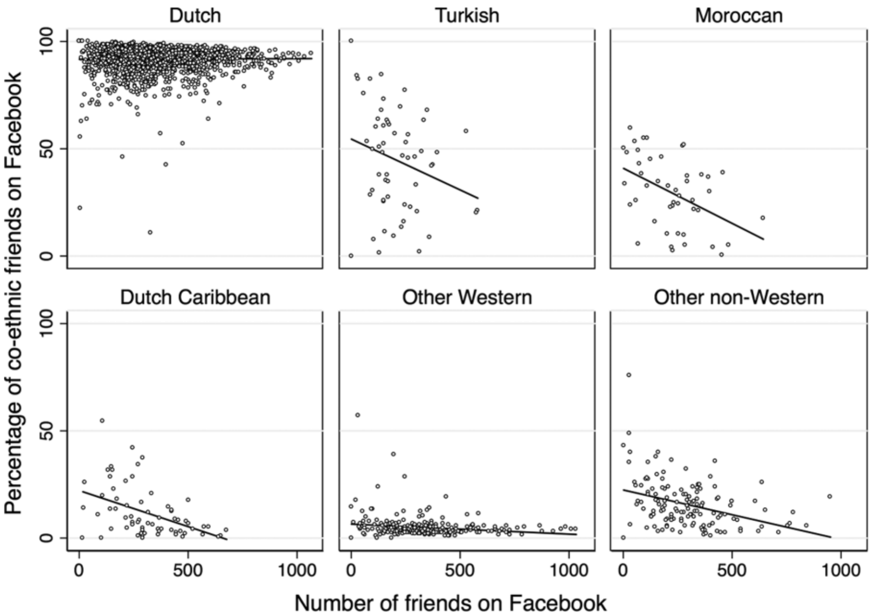

Figure 2. Ethnic Homogeneity of Large Online Networks by Number of Friends, Broken Down by Ethnicity and Including a Fitted Regression Slope

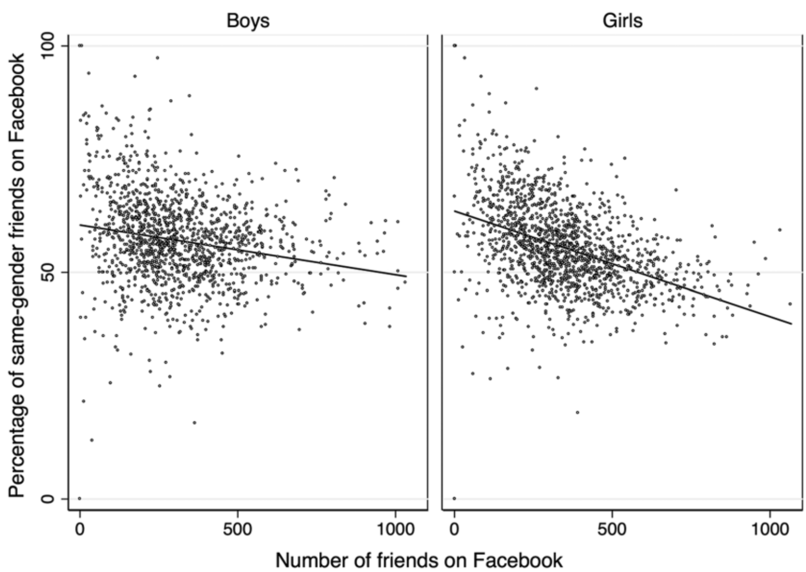

Figure 3. Gender Homogeneity of Large Online Networks by Number of Friends, Broken Down by Gender and Including a Fitted Regression Slope

with all other variables kept constant. The designation of gender homogeneity in classes and school as covariates does not explain away the difference. This finding suggests that, compared to larger online networks, smaller core networks tend to be more gender homogeneous (consistent with Hypothesis 3d).

Confounding Factors

We examined these differences while adjusting for a variety of factors, as shown in Tables 6 and 7. Although the percentage of co-ethnic online network friends does not differ significantly between girls and boys, girls have slightly fewer same-gender online ties than do boys. Turkish minority members also have significantly more same-gender online network friendships than do Dutch majority members. Furthermore, ethnic minority adolescents who are more highly educated have fewer co-ethnic online network friendships than do their less-educated counterparts, and students in a senior general educational track have more same-gender ties online than do students in the lowest educational track. The percentage of co-ethnic friends in general and same-gender friends in a class is also positively associated with ethnic and gender homogeneity in online networks. Dutch majority members who have more positive ethnic out-group attitudes have fewer co-ethnic friends online (implying that they are more likely to connect to minority members), and adolescents who hold more progressive gender role attitudes have fewer same-gender friends online. Finally, adolescents who have more friends in their online networks with the same surnames have a higher percentage of co-ethnic friends online (consistent with the idea that kin ties increase ethnic homogeneity).

CONCLUSIONS AND DISCUSSION

We aimed to answer three key questions that are unresolved in the literature on segregation in social networks: How high are segregation levels in large online networks? Under what conditions does this segregation occur? And how can we explain disparities in segregation between core and larger networks? We show that digital footprint data from online social networks, specifically Facebook, can be used to obtain novel and robust tests of predictions derived from seminal theories of the determinants of segregation in personal networks.

The answer to our first question is that we find high levels of ethnic segregation in online networks. Averaged over all respondents, we find that approximately three-quarters of respondents' Facebook friends are of a similar ethnic background. This ratio is on par with ethnic homogeneity in core networks. However, if we split these estimates by ethnic group, only majority members' core and online networks are equally ethnically homogeneous, whereas minority members have lower levels of ethnic homogeneity in their online than in their core networks. Slightly more than half of online networks friends are the same gender as respondents, whereas in the core networks, the ratio was well above 80 percent.

Second, under what conditions do these patterns of segregation occur in online networks? In the tradition of Blau (1977a, 1977b), Feld (1981), and others who have studied the role of meeting opportunities in the genesis of core ties (e.g., Kalmijn and Flap 2001; Mouw and Entwisle 2006; Wimmer and Lewis 2010), we found that relative group size and social foci are strongly associated with segregation in larger personal networks. Specifically, large networks tend to mirror the structural features of the population and foci. The gender distribution in a population is often 50/50, whereas the distribution of ethnicities is much more unequal. Therefore, we hypothesized and confirmed that gender homogeneity is lower than ethnic homogeneity in online networks. Because ethnic majority members have more opportunities to meet similar others, we expected, and found, that ethnic majority members, compared to ethnic minorities, have higher levels of ethnic homogeneity in their large personal networks. Groups in society segregate over foci and the ties that emerge within them (Feld 1981). Therefore, personal networks resemble the levels of segregation of foci. We thus hypothesized, and found, that homogeneity in foci is positively related to homogeneity among friends online.

Third, we hypothesized and corroborated that as network size increases, larger online networks are characterized by lower gender homogeneity, and that among ethnic minority groups, as online network size increases ethnic homogeneity decreases.

Our results are in line with the propositions that core ties are more segregated than weaker ties (e.g., Blackwell and Lichter 2004; Granovetter 1983; Son and Lin 2012), that dyadic similarity fosters tie strength because returns on investments are more likely (Leszczensky and Pink 2015; Windzio and Bicer 2013), and triadic closure is more pronounced among homogeneous triads. Personal networks initially mirror the features of meeting opportunities, but over time, similar dyads are more likely to become stronger bonds, whereas weak ties will continue to reflect features of meeting opportunities. Ethnic majority members have limited opportunities to befriend dissimilar others, as reflected in core and larger networks that are equally ethnically homogeneous.

Limitations of this Study

Four shortcomings of this study merit acknowledgment. First, the data we used might not be perfectly representative of the overall Dutch adolescent population, due to attrition rates between waves (unit non-response) and selectivity in which respondents are more likely to maintain a public Facebook friend list (item non-response). Future studies might utilize representative samples to generalize our findings to entire adolescent populations, to other age groups, and to other nations. However, when we estimate statistical models that account for (at least some of) this selectivity, we do not find qualitatively different results than those presented by the main analyses, nor do other model specifications (e.g., fixed effects for classes or schools) lead to different results.

Second, predicting the ethnic background of online network friends by their names with our machine-learning algorithm may be an imperfect method, with the potential to misclassify individuals' ethnic backgrounds. More precise measurements of ethnicity among online network friends may be needed to establish robust evidence. One way to address this issue is to collect the birthplace of friends in the data and infer their ethnic background from these birthplaces. Nevertheless, there is a strong correlation between ethnic background and names (Mateos, Longley, and O'Sullivan 2011), and we find evidence for this in our data, especially for respondents of Dutch, Turkish, and Moroccan backgrounds. Limiting our analyses to these groups did not alter our results.

Third, one might argue that Facebook networks do not capture respondents' *complete* networks. There may be selectivity in the Facebook friends of the analyzed respondents, and we may have missed something specific to these friends. For instance, individuals may add to their network contacts on Facebook only others with whom they most closely relate, which could potentially bias the results toward segregation. However, at a mean network size of 336, we have, at the very least, provided insight into *a large portion* of the complete personal network, and most certainly a larger portion of networks than has previously been investigated, as further confirmed by studies on the overall size of social networks (see note 2).

Fourth, our measurement of kin ties likely has measurement error. Kinship tags in online networks can potentially lead to false positive kin matches, as adolescents may tag non-kin as kin. We partly solved this issue by considering realistic tags among adolescents—that is, by not considering granddaughters or grandsons. Nevertheless, some kin tags may seem plausible (e.g., sibling tags) but are not kin. Despite this limitation, we believe that the number of kin tags on Facebook is correlated to the number of real family members present among kin tags—that is, the number of accurate kin tags exceeds the number of fictive kin tags. Some indication for this is the negative correlation between the number of kin tags and gender homogeneity online. However, future research may consider the strategic use of fictive kinship tags in online networks. Additionally, by determining kin based on shared surnames among online friends, we may miss kin who have a different surname than the respondent. However, the number of friends sharing a respondent's surname is positively correlated to ethnic and gender homogeneity on Facebook. This provides some evidence that this variable is a good proxy for the relative amount of kin in online networks. Future research may consider both parents' surnames, as we potentially miss kin ties to mothers' families since children more often share fathers' surnames. Despite these limitations, with our two measurement approaches to kin, we innovatively corrected for kinship ties in estimating segregation in large online networks.

Implications and Future Research

Our study elaborates the work of DiPrete and colleagues (2011), one of the few studies to investigate segregation in larger networks. DiPrete and colleagues used U.S. survey data, whereas we used Facebook data from Dutch adolescents. Despite these differences, some of the conclusions and intuitions related to DiPrete and colleagues' findings are upheld in this study; we found similar ethnic-racial segregation in core networks and larger networks (at least when the estimates by ethnic-racial groups are not split). DiPrete and colleagues (2011:1271) speculated about its causes, stating that meeting opportunities "do not play a strongly integrative role in contemporary . . . society." We aimed to use "imaginative strategies . . . to determine the individual and structural factors that can explain heterogeneity in segregation across individuals" (DiPrete et al. 2011:1273). Based on expectations and considering the relative group sizes of ethnic groups and both genders (Blau 1977a, 1977b), and the segregated nature of foci (Feld 1981), our study confirms that meeting opportunities partially drive segregation among hundreds of contacts on Facebook.

Our findings of different levels of segregation between core networks and larger networks (e.g., Blackwell and Lichter 2004; Granovetter 1973, 1983; Putnam 2000; Son and Lin 2012) seem to contrast with DiPrete and colleagues (2011). However, when we take into account the full range of tie-formation mechanisms, the contrast is attenuated. DiPrete and colleagues (2011) studied segregation along religious, political, socioeconomic, and racial lines—characteristics along which social settings segregate—and consequently found that both core networks and larger networks are segregated. However, we found similarity in ethnic segregation only among core and weaker ties of ethnic majority members. We specifically found that ethnic minority members have far higher levels of homogeneity among their core networks than among their larger networks, and gender homogeneity is significantly higher among core networks than among larger networks. This finding confirms the speculation (e.g., Granovetter 1973, 1983) that weaker ties are less segregated than core ties, but it also explains the findings of DiPrete and colleagues (2011). Moreover, disparities in segregation between core ties and weaker ties occur only under specific circumstances as a consequence of the interplay among meeting opportunities, homophily, and balance.

Because our study uses data on adolescents, we should be careful in generalizing our results to a broader (adult) population. However, the tie-formation mechanisms we consider are not unique to the adolescent population, and many studies show opportunity effects on segregation in other target populations (e.g., Kalmijn and Flap 2001; Mollenhorst et al. 2008). Therefore, we conjecture that our results generalize to different target populations, such as employees, as they do for networks measured from name generators among employees (e.g., Feld 1982; Ibarra 1995). Nevertheless, it may be difficult to empirically observe the patterns we found, given that adolescents' social contexts are well defined (see, e.g., Mollenhorst et al. 2008:62) and that schools are a major focus of tie formation (Coleman 1961; McPherson et al. 2001).

We acknowledge that cross-sectional analyses cannot be used to establish causal direction. Nevertheless, given that the relationship between homogeneity in online networks and opportunity is robust (while controlling for many confounders), we tentatively assume that similar results will be replicated using longitudinal data. We recommend that future research considers segregation in large personal networks over time and settings as a next step to obtain potential causal estimates.

Another area for future research is feedback effects: What are the effects of large personal networks online vis-à-vis core networks offline? Will weak ties characterized by dyadic similarity turn into strong ties, and what online behaviors (e.g., Facebook wall posts) will lead people out of the mode of segregated large online networks?

Finally, future research may examine whether the implications of core network segregation, such as out-group attitude formation, result from segregation in online networks. Contact theory predicts that having more contact with out-group members reduces prejudice toward them (Allport 1954; Pettigrew and Tropp 2006). Does this apply to the relationship between ethnic segregation in large online networks (on Facebook) and ethnic prejudice?

APPENDIX

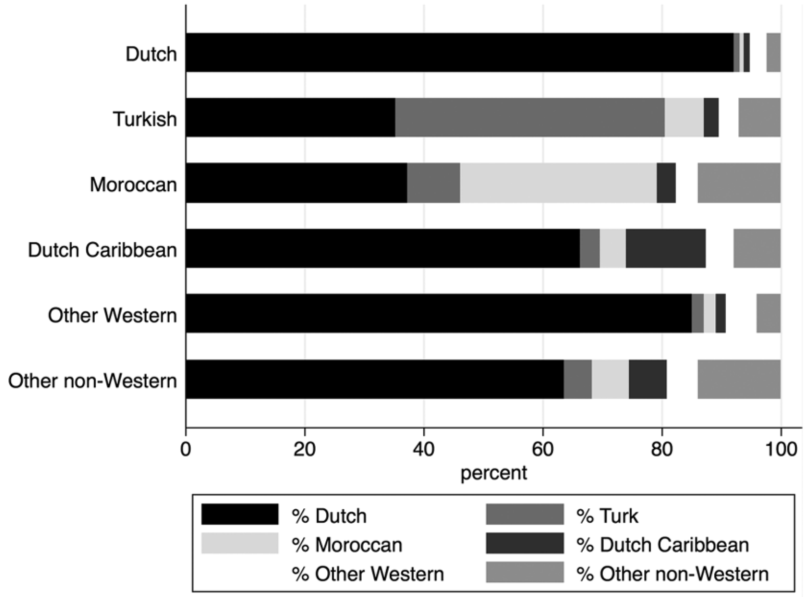

Figure A1. Ethnic Segregation of Social Networks Online, Broken Down by Ethnicity

Table A1. Multilevel Model Estimating the Difference between the Percentage of Co-ethnic and Same-Gender Friends in Online Networks; Test for Hypothesis 1a

	Intercept Only (Hypothesis 1a)		
	Coef.	SE	*p*
Fixed part			
Intercept	16.851	(1.515)	***
Random part			
$\sigma^2_{s_{0k}}$ (school level)	185.033	(44.113)	
$\sigma^2_{c_{0jk}}$ (class level)	3.629	(12.4899)	
$\sigma^2_{p_{0ijk}}$ (pupil level)	1016.715	(45.640)	
Number of schools	114		
Number of classes	315		
Number of pupils	2,690		
Log likelihood		–13216.774	

Note: Robust standard errors, adjusted for the school-identifier.
*p < .05; **p < .01; ***p < .001 (one-tailed tests).

Table A2. Multilevel Model Estimating the Difference between Ethnic Segregation in Core and Online Networks; Tests for Hypothesis 3c

	$\text{Co-ethnic}_{\text{FRIENDS IN GENERAL}} - \text{Co-ethnic}_{\text{FACEBOOK}}$		
	Coef.	SE	*p*
Fixed part			
Intercept	–13.164	(6.136)	*
Core-network			
$\text{Co-ethnic}_{\text{FRIENDS IN GENERAL}}$			
$\text{Co-ethnic}_{\text{FRIENDS IN CLASS}}$			
Opportunity			
$\text{Co-ethnic}_{\text{IN CLASS}}$	.004	(.038)	
$\text{Co-ethnic}_{\text{IN SCHOOL}}$	.155	(.051)	**
Ethnicity (ref.: Dutch)			
Turkish	26.375	(5.171)	***
Moroccan	27.692	(5.561)	***
Dutch Caribbean	32.342	(4.668)	***
Other Western	25.326	(3.785)	***
Other non-Western	29.889	(4.607)	***
Number of Facebook friends	.003	(.003)	
Facebook membership (ref.: 2013)			
2012	1.843	(4.865)	
2011	1.210	(4.699)	
2010	–.107	(4.722)	
2009	.642	(4.814)	
2008	–1.757	(4.672)	
2007	–.693	(5.337)	
2006	–3.755	(6.975)	
Girls (ref.: boys)	.551	(.804)	
Educational track (ref.: lower voc.)			
Senior general	2.189	(.928)	**
University preparatory	.647	(1.127)	*
Indegree popularity	–.071	(.033)	*
Ethnic out-group attitudes	–.761	(.260)	**
% Kinship ties declared	–.323	(.337)	
% Similar surname on Facebook	–.170	(.175)	
Random part			
$\sigma^2_{s_{0k}}$ (school level)	.000	(.000)	
$\sigma^2_{c_{0jk}}$ (class level)	2.643	(4.490)	
$\sigma^2_{p_{0ijk}}$ (pupil level)	387.398	(28.982)	
Number of schools	112		
Number of classes	309		
Number of pupils	2,549		
Log likelihood		–11220.572	

Note: Robust standard errors, adjusted for the school-identifier.
*p < .05; **p < .01; ***p < .001 (one-tailed tests).

Table A3. Multilevel Model Estimating the Difference between Gender Segregation in Core and Online Networks; Tests of Hypothesis 3d

	Same-gender$_{\text{FRIENDS IN CLASS}}$ – Same-gender$_{\text{FACEBOOK}}$		
	Coefficient	SE	p
Fixed part			
Intercept	–29.991	(7.390)	***
Core-network			
Same-gender$_{\text{FRIENDS IN CLASS}}$			
Opportunity			
Same-gender$_{\text{IN CLASS}}$	.027	(.039)	
Same-gender$_{\text{IN SCHOOL}}$	.047	(.046)	
Ethnicity (ref.: Dutch)			
Turkish	19.791	(5.266)	***
Moroccan	16.665	(6.377)	**
Dutch Caribbean	3.218	(4.045)	
Other Western	1.271	(2.214)	
Other non-Western	4.921	(3.830)	
Number of Facebook friends	–.010	(.004)	**
Facebook membership (ref.: 2013)			
2012	–3.543	(6.967)	
2011	–2.803	(6.326)	
2010	–1.098	(6.356)	
2009	–2.031	(6.527)	
2008	1.233	(6.746)	
2007	10.760	(10.483)	
2006	3.549	(10.683)	
Girls (ref.: boys)	–1.670	(1.528)	
Educational track (ref.: lower voc.)			
Senior general	5.468	(1.604)	**
University preparatory	–1.001	(2.014)	
Indegree popularity	.082	(.050)	*
Gender role attitudes	.809	(.485)	*
% Kinship ties declared	–.420	(.348)	
% Similar surname on Facebook	.455	(.366)	
Random part			
$\sigma^2_{s_{0k}}$ (school level)	.000	(.000)	
$\sigma^2_{c_{0jk}}$ (class level)	50.935	(14.677)	
$\sigma^2_{p_{0ijk}}$ (pupil level)	856.391	(38.042)	
Number of schools	108		
Number of classes	301		
Number of pupils	2,596		
Log likelihood	–12508.024		

Note: Robust standard errors, adjusted for the school-identifier.
*p < .05; **p < .01; ***p < .001 (one-tailed tests).

Acknowledgments

We thank Manja Coopmans, Maaike van der Vleuten, Jesper Rözer, Joseph B. Bayer, Niek de Schipper, Müge Simsek, Bernie Hogan, René Veenstra, and Jeroen Weesie for helpful advice on earlier drafts of this article. This article benefited from discussions at the "Migration and Social Stratification seminar" in Utrecht, the "CILS4EU conference" in Stockholm, the "XXXVI SUNBELT" in Newport Beach, the "CITP/MILAB workshop" in Vienna, the "INAS 2016" in Utrecht, Nuffield College in Oxford, the "Social Sciences and the Internet Conference" in Eindhoven, the "Dag van de Sociologie" in Tilburg, and the "Dutch Demography Day 2016" in Utrecht. Finally, we acknowledge the detailed feedback of five anonymous reviewers as well as the valuable insights of the *ASR* editors to improve this article.

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 [grant number: 406-12-004], NWO middelgroot [grant number: 480-11-013], and NWO veranderingsstudies [grant number: 481-11-004].

Notes

1. There has been a debate on the increase in social isolation (i.e., having zero alters to discuss "important matters" with) of Americans between 1985 and 2004, as reported by McPherson and colleagues (2006). A 2008 erratum (McPherson, Smith-Lovin, and Brashears 2008) corrected a coding error in the original data release. In 2009 (Fischer 2009; McPherson, Smith-Lovin, and Brashears 2009), a discussion emerged on whether the trend was a data artifact that resulted from respondent fatigue and training. Paik and Sanchagrin (2013) showed that the increase in social isolation may be attributed to substantial interviewer effects.
2. McCarty and colleagues (2001) found a mean total network size of 290; Hill and Dunbar (2003) found a size of 150; Zheng and colleagues (2006) found a size of 750; McCormick, Salganik, and Zheng (2010) found a size of 611; and DiPrete and colleagues (2011) found a median network size of 550.
3. Balance is restored in this network configuration (when *A* is friends with *B*, and *A* with *C*, then *B* and *C* are likely to connect) under the assumption that these ties are positively signed, undirected, and of the same tie strength. For instance, a closed triad of three mutual foes is an unbalanced triad.
4. One can apply for data access to waves 1, 2, and 3 of the CILS4EU via the following link: https://dbk.gesis.org/dbksearch/sdesc2.asp?no=5353&db=e&doi=10.4232/cils4eu.5353.2.1.0.
5. In wave 1, 600 respondents who were not part of the random sampling frame were sampled because some schools wanted to participate in the survey with more than two classrooms. Therefore, a *random* sample of 4,363 pupils was drawn in wave 1. Because of the attrition rates between waves 1 and 2, our sample is not necessarily representative. We included as many respondents as possible in the sample for analyses, including newcomers (nonrandom) and the nonrandom sample of wave 1, to ensure a large sample size.
6. An anonymized version of the DFS will be available in October 2017 (https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:62379).
7. We performed robustness analyses using Heckman selection models (Heckman 1979). Therefore, we corrected for selectivity in modeling an outcome only when a second selection equation determined that this outcome was non-missing. The errors of both equations are allowed to correlate. We correct in the selection equation for ethnicity, gender, popularity, and educational-track level. We cluster-corrected standard errors for the class cluster and school cluster, because multilevel Heckman models were computationally infeasible. These analyses did not provide different results than those we present here. We present the analyses that consider the clustered data. Tables are available upon request.
8. The combined total of 1,158,227 friends is a raw count of all respondents' friendships. Respondents likely have similar friends in their online networks. Counting the unique set of friends would most likely result in a lower number.
9. The collection and use of these data for scientific purposes were internally approved by an ethical review board for the social and behavioral sciences.
10. We also assigned ethnicities based on name carriers' last names, following the procedure outlined in Part A of the online supplement. We obtained correlations similar to those based on first names. We reperformed all analyses pertaining to ethnicity, and the results are robust when we consider last names.
11. We also performed all of our descriptive analyses for ethnicity with an index of qualitative variation (IQV)—the inverse of network *diversity* (Agresti and Agresti 1977). The IQV for pupil *i* is formally defined as follows:

$$\mathrm{IQV}_i = 1 - [\left(\frac{k}{k-1}\right)(1 - \sum_{b=1}^{k} p_b^2)], \tag{3}$$

where k is the number of ethnic categories and p_b is the fraction of Facebook friends in the bth category ($b = 1, \ldots, k$). IQV has been used in various studies to measure (ethnic) diversity in networks (e.g., Lewis et al. 2008; Marsden 1987; McPherson et al. 2006). In none of the analyses did the results differ from those presented in the article.
12. In removing kin from the Facebook homogeneity estimates, we assume that all kin are of a similar

ethnic background as the respondent. We reduce the number of co-ethnic friends and the number of Facebook friends by the number of identified kin ties and calculate the percentage of co-ethnic friends on Facebook. We assume that half the kin ties are of similar gender as the respondent. We reduce the number of same-gender friends by *half* the number of identified kin and subtract the *total* kin ties that are identified from the number of Facebook friends.

13. *Indegree popularity* can be formally defined as follows:

$$(\sum_i B_{ji} / N-1)\times 100, \tag{4}$$

where i is the actor, B_{ji} indicates whether pupil j nominates pupil i as popular, and N is the total number of pupils in a classroom.

14. We also controlled for dummy variables that indicate to which stratum in the sampling frame the respondent belongs; thus, we account for some of the selectivity in the sampling strategy. In none of the analyses does this control variable lead to qualitatively different results. To keep the results parsimonious, we present the results without these variables.

15. We furthermore controlled for dummy variables that indicate respondents' generational immigration status. Categories are, for instance, Dutch majority adolescents or adolescents who have only one foreign-born grandparent. Thus, we account for differences in immigration background. In none of the analyses does this control variable lead to qualitatively different results. To keep the results parsimonious, we present the results without these variables. For more information about generational status in the CILS4EU data, see Dollmann, Jacob, and Kalter (2014) and the CILS4EU Wave 2 Codebook (2016:273).

16. Pupils from the same class may look more alike than pupils from different classes, and the proportion of variance explained at the class and school levels represents the expected correlation between two randomly selected pupils within the same class. This is defined as follows:

$$\rho_{s+c} = \frac{\sigma^2_{s_{0k}} + \sigma^2_{c_{0jk}}}{\sigma^2_{s_{0k}} + \sigma^2_{c_{0jk}} + \sigma^2_{p_{0ijk}}}. \tag{5}$$

17. Pupils from the same school may resemble each other more than pupils from different schools, and the expected correlation between two randomly selected pupils from the same school is defined as follows:

$$\rho_{s} = \frac{\sigma^2_{c_{0jk}}}{\sigma^2_{s_{0k}} + \sigma^2_{c_{0jk}} + \sigma^2_{p_{0ijk}}}. \tag{6}$$

18. We estimated similar models with the *number* of co-ethnic and same-gender friendships on Facebook and offline instead of percentages as dependent and independent variables (we controlled for network sizes). We estimated random-effect models with either a random intercept at the class or school level. We estimated fixed-effect models with dummies for schools and classes. Finally, we estimated a model with only Dutch, Turkish, and Moroccans because we could best predict these ethnicities. In none of these analyses did the results qualitatively differ from the results presented here. Full tables are available upon request.

19. We performed additional robustness analyses to investigate whether our main results are driven by socially isolated or highly connected adolescents on Facebook. We obtained similar results when we selected respondents with more than 50 and fewer than 750 friends on Facebook in our statistical models.

20. For Dutch majority members, we performed analyses considering the percentage of co-ethnic neighborhood residents based on supplementary data from Statistics Netherlands. More co-ethnic neighborhood residents positively relate to the percentage of co-ethnic friends on Facebook. We could not perform similar analyses for ethnic minority members, because these data do not distinguish between the presence of various ethnic minority groups in a neighborhood.

21. Correlations between degree and individual properties may happen by chance. By design, there may be a negative correlation between degree and individual-level clustering in social networks. The friends of high-degree individuals are less likely to be linked than are friends of low-degree individuals (Jackson 2008). We tested whether the correlation between homogeneity and degree originates from design. We randomly rewired the ties of Facebook networks while keeping individual degree constant; we found no correlations between degree and homogeneity in this random mixing model. We therefore assume that H_0 is r(degree, homogeneity) = 0 (for more details, see Part B of the online supplement).

References

Allport, Gordon W. 1954. *The Nature of Prejudice*. Reading, MA: Addison-Wesley.

Agresti, Alan, and Barbara F. Agresti. 1977. "Statistical Analysis of Qualitative Variation." Pp. 204–237 in *Sociological Methodology*, edited by K. F. Schuessler. San Francisco: Jossey-Bass.

Anderson, Ashton, Sharad Goel, Gregory Huber, Neil Malhotra, and Duncan J. Watts. 2014. "Political Ideology and Racial Preferences in Online Dating." *Sociological Science* 1:28–40; doi: 10.15195/v1.a3.

Baerveldt, Chris, Marijtje A. J. Van Duijn, Lotte Vermeij, and Dianne A. Van Hemert. 2004. "Ethnic Boundaries and Personal Choice: Assessing the Influence of

Individual Inclinations to Choose Intra-Ethnic Relationships on Pupils' Networks." *Social Networks* 26(1):55–64.

Blackwell, Debra L., and Daniel T. Lichter. 2004. "Homogamy among Dating, Cohabiting and Married Couples." *Sociological Quarterly* 45(4):719–37.

Blau, Peter M. 1977a. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.

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

Bloothooft, Gerrit, and Marijn Schraagen. 2011. "Name Fashion Dynamics and Social Class." Working paper. Retrieved August 8, 2006 (http://www.let.uu.nl/~Gerrit.Bloothooft/personal/Publications/ICOS2011_Name%20fashion%20dynamics_Bloothooft_Schraagen.pdf).

Brashears, Matthew E., Emily Hoagland, and Eric Quintane. 2016. "Sex and Network Recall Accuracy." *Social Networks* 44:74–84; doi: 10.1016/j.socnet.2015.06.002.

Byrne, Donn E. 1971. *The Attraction Paradigm*. New York: Academic Press.

Castles, Stephen, Hein de Haas, and Mark J. Miller. 2013. *The Age of Migration: International Population Movements in the Modern World*, 5th ed. New York: The Guilford Press.

Centola, Damon. 2015. "The Social Origins of Network and Diffusion." *American Journal of Sociology* 120(5):1295–1338.

CILS4EU. 2016. *Children of Immigrants Longitudinal Survey in Four European Countries. Codebook. Wave 2 – 2011/2012, v2.3.0*. Mannheim University. (http://www.cils4.eu/images/wave2_material/codebook/za5353_cod_wave2.pdf).

Coleman, James S. 1961. *The Adolescent Society*. New York: Free Press.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 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):4857–61.

Davis, Shannon N., and Theodore N. Greenstein. 2009. "Gender Ideology: Components, Predictors, and Consequences." *Annual Review of Sociology* 35:87–105.

DiPrete, Thomas A., Andrew Gelman, Tyler McCormick, Julien Teitler, and Tian Zheng. 2011. "Segregation in Social Networks Based on Acquaintanceship and Trust." *American Journal of Sociology* 116(4):1234–83.

Dollmann, Jörg, Konstanze Jacob, and Frank Kalter. 2014. "Examining the Diversity of Youth in Europe: A Classification of Generations and Ethnic Origins Using CILS4EU Data (Technical Report)." Retrieved May 6, 2016 (http://www.mzes.uni-mannheim.de/publications/wp/wp-156.pdf).

Duggan, Maeve, Nicole B. Ellison, Cliff Lampe, Amanda Lenhard, and Mary Madden. 2015. "Social Media Update 2014." Pew Research Center. Retrieved July 5, 2016 (http://www.pewinternet.org/2015/01/09/social-media-update-2014/).

Dunbar, Robin 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; doi:10.1098/rsos.150292.

Dunbar, Robin 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; doi: 10.1016/j.socnet.2015.04.005.

Ellison, Nicole B., and danah m. boyd. 2013. "Sociality through Social Network Sites." Pp. 150–72 in *Oxford Handbook of Internet Studies*, edited by W. H. Dutton. Oxford, UK: Oxford University Press.

Ellison, Nicole B., Charles Steinfield, and Cliff Lampe. 2007. "The Benefits of Facebook 'Friends': Social Capital and College Students' Use of Online Social Network Sites." *Journal of Computer-Mediated Communication* 12(4):1143–68.

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

Facebook. 2015. "Facebook Company Info." Retrieved July 29, 2015 (http://newsroom.fb.com/company-info/).

Feld, Scott L. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86(5):1015–35.

Feld, Scott L. 1982. "Social Structural Determinants of Similarity among Associates." *American Sociological Review* 47(6):797–801.

Feld, Scott L. 1984. "The Structured Use of Personal Associates." *Social Forces* 62(2):640–52.

Feld, Scott L. 1997. "Structural Embeddedness and Stability of Interpersonal Relationships." *Social Networks* 19:94–95; doi: 10.1016/S0378-8733(96)00293-6.

Feld, Scott L., and William C. Carter. 1999. "Foci of Activity as Changing Contexts for Friendship." Pp. 136–52 in *Placing Friendships in Context*, edited by R. G. Adams and G. Allan. Cambridge, UK: Cambridge University Press.

Feliciano, Cynthia, Belinda Robnett, and Golnaz Komaie. 2009. "Gendered Racial Exclusion among White Internet Daters." *Social Science Research* 38(1):39–54.

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

Fukuyama, Francis. 1995. *Trust*. New York: Free Press.

Gambetta, Diego. 1988. *Trust: Making and Breaking Cooperative Relations*. New York: Blackwell.

Golder, Scott A., and Michael W. Macy. 2014. "Digital Footprints: Challenges and Opportunities for Online Social Research." *Annual Review of Sociology* 40:129–52.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78(6):1360–80.

Granovetter, Mark S. 1983. "The Strength of Weak Ties: A Network Theory Revisited." *Sociological Theory* 1:201–233; doi: 10.2307/202051.

Grosetti, Michel. 2005. "Where do Social Relations Come From? A Study of Personal Networks in the Toulouse Area of France." *Social Networks* 27(4):289–300.

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.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153–61.

Heider, Fritz. 1946. "Attitudes and Cognitive Organization." *Journal of Psychology* 21(1):107–112.

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

Hofstra, Bas, Rense Corten, and Frank van Tubergen. 2015. *Dutch Facebook Survey: Wave 1* [dataset and codebook]. Utrecht University, Department of Sociology/ICS (https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:62379). doi: 10.17026/dans-274-azju.

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):2340–58.

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* 60:611–21.

Ibarra, Herminia. 1995. "Race, Opportunity, and Diversity of Social Circles in Managerial Networks." *Academy of Management Journal* 38(3):673–703.

Jackson, Matthew O. 2008. *Social and Economic Networks*. Oxfordshire, UK: Princeton University Press.

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

Kalmijn, Matthijs, and Henk Flap. 2001. "Assortative Meeting and Mating: Unintended Consequences of Organized Settings for Partner Choices." *Social Forces* 79(4):1289–1312.

Kalter, Frank, Anthony F. Heath, Miles Hewstone, Jan O. Jonsson, Matthijs Kalmijn, Irena Kogan, and Frank van Tubergen. 2013. *Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU)*. ZA5353, Data file version 1.0.0. GESIS Data Archive, Cologne [distributor].

Kalter, Frank, Anthony F. 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. GESIS Data Archive, Cologne [distributor].

Kossinets, Gueorgi, and Duncan J. Watts. 2009. "Origins of Homophily in an Evolving Social Network." *American Journal of Sociology* 115(2):405–450.

Krackhardt, David, and Mark S. Handcock. 2007. "Heider vs Simmel: Emergent Features in Dynamic Structures." Pp. 14–27 in *Statistical Network Analysis: Models, Issues and New Directions*, Vol. 4503, *Lecture Notes in Computer Science*, edited by E. Airoldi, D. M. Blei, S. E. Fienberg, A. Goldenberg, E. P. Xing, and A. X. Zheng. Berlin: Springer Berlin Heidelberg.

Lancee, Bram, and Oriane Sarrasin. 2015. "Educated Preferences or Selection Effects? A Longitudinal Analysis of the Impact of Educational Attainment on Attitudes towards Immigrants." *European Sociological Review* 31(4):490–501.

Lazarsfeld, Paul F., and Robert K. Merton. 1954. "Friendships as a Social Process: A Substantive and Methodological Analysis." Pp. 18–66 in *Freedom and Control in Modern Society*, edited by M. Berger. New York: Van Nostrand.

Leszczensky, Lars, and Sebastian Pink. 2015. "Ethnic Segregation of Friendship Networks in School: Testing a Rational-Choice Argument of Differences in Ethnic Homophily between Classroom- and Grade-Level Networks." *Social Networks* 42:18–26; doi: 10.1016/j.socnet.2015.02.002.

Lewis, Kevin, Marco Gonzalez, and Jason Kaufman. 2012. "Social Selection and Peer Influence in an Online Social Network." *Proceedings of the National Academy of Sciences* 109(1):68–72.

Lewis, Kevin. 2013. "The Limits of Racial Prejudice." *Proceedings of the National Academy of Sciences* 110(47):18814–19.

Lewis, Kevin. 2015. "How Networks Form: Homophily, Opportunities, and Balance." Pp. 1–14 in *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*, edited by R. A. Scott and S. M. Kosslyn. Hoboken, NJ: John Wiley & Sons.

Lewis, Kevin, Jason Kaufman, Marco Gonzalez, Andreas Wimmer, and Nicholas Christakis. 2008. "Tastes, Ties, and Time: A New Social Network Dataset Using Facebook.com." *Social Networks* 30(4):330–42.

Lichter, Daniel T., Domenico Parisi, and Michael C. Taquino. 2015. "Toward a New Macro-Segregation? Decomposing Segregation within and between Metropolitan Cities and Suburbs." *American Sociological Review* 80(4):843–73.

Lin, Nan. 1999. "Social Capital and Status Attainment." *Annual Review of Sociology* 25:467–87.

Lubbers, Miranda J. 2003. "Group Composition and Network Structure in School Classes: A Multilevel Application of the P* Model." *Social Networks* 25(4):309–332.

Marin, Alexandra. 2004. "Are Respondents More Likely to List Alters with Certain Characteristics? Implications for Name Generator Data." *Social Networks* 26(4):289–307.

Marsden, Peter V. 1987. "Core Discussion Networks of Americans." *American Sociological Review* 52(1):122–31.

Marsden, Peter V. 1988. "Homogeneity in Confiding Relations." *Social Networks* 10(1):57–76.

Mäs, Michael, and Andreas Flache. 2013. "Differentiation without Distancing: Explaining Bi-Polarization of Opinions without Negative Influence." *PloS One* 8(11):e74516.

Mateos, Pablo, Paul A. Longley, and David O'Sullivan. 2011. "Ethnicity and Population Structure in Personal Naming Networks." *PLoS One* 6(9):e22943.

Mayer, Adalbert, and Steven L. Puller. 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, H. Russell Bernard, Eugene C. Johnsen, and Gene A. 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, and Lynn Smith-Lovin. 1987. "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups." *American Sociological Review* 52(3):370–79.

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):353–75.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2008. "ERRATA: Social Isolation in America: Changes in Core Discussion Networks over Two Decades." *American Sociological Review* 73(6):1022.

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):670–81.

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.

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.

Mollenhorst, Gerald, Beate Volker, and Henk Flap. 2011. "Shared Contexts and Triadic Closure in Core Discussion Networks." *Social Networks* 33(4):292–302.

Mollenhorst, Gerald, Beate Volker, and Henk Flap. 2014. "Changes in Personal Relationships: How Social Contexts Affect the Emergence and Discontinuation of Relationships." *Social Networks* 37:65-80; doi: 10.1016/j.socnet.2013.12.003.

Moody, James. 2001. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology* 107(3):679–716.

Mouw, Ted, and Barbara Entwisle. 2006. "Residential Segregation and Interracial Friendship in Schools." *American Journal of Sociology* 112(1):394–441.

Paik, Anthony, and Kenneth Sanchagrin. 2013. "Social Isolation: An Artifact." *American Sociological Review* 78(3):339–60.

Pettigrew, Thomas F., and Linda R. Tropp. 2006. "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality Social Psychology* 90(5):751–83.

Potârcă, Gina, and Melinda Mills. 2015. "Racial Preferences in Online Dating across European Countries." *European Sociological Review* 31(3):326–41.

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.

Rainie, Lee, and Barry Wellman. 2012. *Networked: The New Social Operating System*. Cambridge, MA: The MIT Press.

Reich, Stephanie M., Kaveri Subrahmanyam, and Guadalupe Espinoza. 2008. "Friending, IMing, and Hanging Out Face-to-Face: Overlap in Adolescents' Online and Offline Social Networks." *Developmental Psychology* 48(2):356–68.

Roberts, Sam G. B., Robin I. M. Dunbar, Thomas V. Pollet, and Toon Kuppens. 2009. "Exploring Variation in Active Network Size: Constraints and Ego Characteristics." *Social Networks* 31:138–46.

Robinson, Laura, Shelia R. Cotten, Hiroshi Ono, Anabel Quan-Haase, Gustavo Mesch, Wenhong Chen, Jeremy Schulz, Timothy M. Hale, and Michael Stern. 2015. "Digital Inequalities and Why They Matter." *Information, Communication & Society* 18(5):569–82.

Semyonov, Moshe, and Anya Glikman. 2009. "Ethnic Residential Segregation, Social Contacts, and Anti-Minority Attitudes in European Societies." *European Sociological Review* 25(6):693–708.

Shrum, Wesley, Neil H. Cheek, and Saundra MacD. Hunter. 1988. "Friendship in School: Gender and Racial Homophily." *Sociology of Education* 61(4):227–39.

Smith, Andrea, and Barry H. Schneider. 2000. "The Interethnic Friendships of Adolescent Students: A Canadian Study." *International Journal of Intercultural Relations* 24(2):247–58.

Smith, Jeffrey A., Miller McPherson, and Lynn Smith-Lovin. 2014. "Social Distance in the United States: Sex, Race, Religion, Age, and Education Homophily among Confidants, 1985 to 2004." *American Sociological Review* 79(3):432–56.

Smith, Sanne, Ineke Maas, and Frank van Tubergen. 2014. "Ethnic Ingroup Friendships in Schools: Testing the By-Product Hypothesis in England, Germany, the Netherlands and Sweden." *Social Networks* 39:33–45; doi: 10.1016/j.socnet.2014.04.003.

Snijders, Tom A. B., and Roel J. Bosker. 2012. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London, UK: Sage Publishers.

Son, Joonmo, and Nan Lin. 2012. "Network Diversity, Contact Diversity, and Status Attainment." *Social Networks* 34(4):601–613.

Stark, Tobias H., and Andreas Flache. 2012. "The Double Edge of Common Interest." *Sociology of Education* 85(2):179–99.

Statistics Netherlands. 2012. "Migranten, Vreemdelingen en Vluchtelingen: Begrippen op het Terrein van Asiel en Buitenlandse Migratie." Retrieved June 16, 2015 (http://www.cbs.nl/nl-NL/menu/themas/bevolking/publicaties/bevolkingstrends/archief/2012/2012-10-bt-btmve-migratie.htm).

Statistics Netherlands. 2015. "CBS Statline: Bevolking; kerncijfers." Retrieved March 23, 2015 (http://statline.cbs.nl).

Subrahmanyam, Kaveri, Stephanie M. Reich, Natalia Waechter, and Guidalupe 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 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.

Vermeij, Lotte, Martije A. J. van Duijn, and Chris Baerveldt. 2009. "Ethnic Segregation in Context: Social Discrimination among Native Dutch Pupils and Their Ethnic Minority Classmates." *Social Networks* 31(4):230–39.

Wimmer, Andreas, 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.

Windzio, Michael, and Enis Bicer. 2013. "Are We Just Friends? Immigrant Integration into High- and Low-Cost Social Networks." *Rationality and Society* 25(2):123–45.

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

Zhou, Wei-Xing, Didier Sornette, Russel A. Hill, and Robin I. M. Dunbar. 2005. "Discrete Hierarchical Organization of Social Group Sizes." *Proceedings of the Royal Society B* 272(1561):439–44.

Bas Hofstra is a PhD candidate at the Department of Sociology at Utrecht University and affiliated with the Interuniversity Center for Social Science Theory and Methodology (ICS). This study is part of his PhD thesis on the interplay between off- and online social networks. Specifically, he studies (early) adoption of social media platforms, privacy, and the (consequences of) size, structure, and segregation in networks off- and online. His research combines traditional survey data and (big) data obtained from online platforms, such as presented in this article.

Rense Corten is Assistant Professor at the Department of Sociology at Utrecht University and member of the ICS. His research revolves around the themes of cooperation, trust, and social networks, with empirical applications including adolescents' networks, social media, the sharing economy, online criminal networks, and laboratory experiments.

Frank van Tubergen is Professor of Theoretical and Empirical Sociology at the Department of Sociology, Utrecht University, and Distinguished Adjunct Professor at the King Abdulaziz University in Jeddah, Saudi Arabia. He is a member of the ICS and the European Academy of Sociology. His research topics are social networks, religion, and immigration.

Nicole B. Ellison is a Professor in the School of Information at the University of Michigan. Her research addresses issues of social capital, relationship development, self-presentation, and identity in online environments such as Facebook and online dating sites. A list of her publications is available here: http://scholar.google.com/citations?user=P0XTYBEAAAAJ.