Reports

Sizes of Permanent Campsite Communities Reflect Constraints on Natural Human Communities

Tobias Kordsmeyer, Pádraig Mac Carron, and R. I. M. Dunbar

Department of Biological Personality Psychology, University of Goettingen, Gosslerstrasse 14, 37073 Göttingen, Germany (tobias.kordsmeyer@psych.uni-goettingen.de)/Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, United Kingdom (Kordsmeyer, Mac Carron, and Dunbar). This paper was submitted 17 XI 15, accepted 19 X 16, and electronically published 31 I 17.

Both small-scale human societies and personal social networks have a characteristic hierarchical structure with successively inclusive layers of 15, 50, 150, 500, and 1,500 individuals. It has been suggested that these values represent a set of natural social attractors, or "sweet spots," in organizational terms. We exploited the new phenomenon of permanent (i.e., residential) campsites to ask whether these values are present in the size distribution of the numbers of residents in these naturally small-scale communities. In two separate data sets of different grain, we find consistent evidence for sites with 50, 150, 500, and maybe 1,500 residents. We infer that these reflect numerical sizes at which communities may in some way be socially optimal. Our data do not allow us to say why this pattern emerges, but the consistency of the results and the fact that the predetermined sizes of permanent campsites adhere to this pattern suggest that it may arise from the limits on the number of relationships that make an effective community.

Hunter-gatherer communities typically occur in quite specific sizes that form a hierarchically scaled series of natural groupings of approximately 50, 150, 500, and 1,500 individuals, with a scaling ratio of approximately 3 (Hamilton et al. 2007; Lehmann, Lee, and Dunbar 2014; Zhou et al. 2005). These groupings correspond to communities that are conventionally labelled as bands (or overnight camp groups), communities (or clans), mega-bands, and tribes (Lehmann, Lee, and Dunbar 2014), respectively. Data from Facebook, Twitter, e-mail, and massive multiplayer online games suggest that this same grouping pattern and scaling ratio occurs even in online environments (Arnabaldi et al. 2015; Dunbar et al. 2015; Fuchs et al. 2014; Gonçalves, Perra, and Vespignani 2011; Haerter, Jamtveit, and Mathiesen 2012; Pollet, Roberts, and Dunbar 2011). These structural features of communities turn out to mirror the internal structure of personal social networks (Hill and Dunbar 2003; Sutcliffe et al. 2012; Zhou et al. 2005) and are similar to the layering pattern found in animal species that live in complex societies (Hill, Bentley, and Dunbar 2008).

Quite why social communities and networks should have the values and scaling ratio they do is, as yet, unclear. However, the core value of ∼150 fits with the predictions of the social brain hypothesis (Dunbar 1992, 1993). Since we know from a series of neuroimaging studies in humans that individual differences in the size of personal social networks are correlated with the volume of core brain regions associated with the mentalizing circuit, notably in the prefrontal cortex (Kanai et al. 2011; Lewis et al. 2011; Powell et al. 2012, 2014), it is likely that this reflects a real cognitive constraint of some kind. If so, it implies that these numbers are relatively fixed and are likely to reappear in many different social contexts.

In recent years, a new movement in German housing has emerged: elderly people, in particular, give up their regular housing (mostly rented apartments) and move to permanent camping sites, where little villages and communities emerge naturally, often equipped with amenities like pubs or nurseries (Soares 2013). In many such cases, the accommodation involves mobile homes and is probably similar to the trailer park phenomenon in the United States. Purportedly, the main reasons for this are the low cost of such locations, compared with the increasing price of regular housing, combined with the communality or sense of community that they provide. This unusual process of contemporary small-scale natural community formation provides us with a unique opportunity to test the hypothesis that there are natural grouping sizes for such communities.

We collated data on community sizes at permanent campsites in Germany and asked whether the same kind of patterns that have been found in hunter-gatherer societies are also evident in these. We considered two data sets: a small data set for which we established exact numbers of residents and a large data set for which only the number of "camping pitches" was available (which we used to estimate maximum community size). We ask two questions of the data. First, does the actual number of residents on these campsites exhibit peaks in frequency that correspond to the layers of natural human communities and personal social networks? In effect, we view these campsites as a form of joiner-leaver game in which individuals or families join or leave the community depending on whether they find its size socially congenial. Such decisions may, of course, be complex and involve many factors, but by focussing on the end product (the census size of camps), we observe the cumulative outcome of many such decisions over time. Sec-

 0011-3204/2017/5802-0009$10.00. DOI: 10.1086/690731

ond, using the larger data set, we ask whether campsite owners design their sites with these numbers in mind. The null hypothesis in this case is that, all else being equal, site sizes should be distributed with an arbitrary mean and variance reflecting the area and funding available to establish sites. At worst, there should be no particular pattern, and all sizes of sites might be equally represented. We assume that site owners would like to maximize the number of residents they have, but they might be intuitively aware that some community sizes are most attractive; hence, they might adjust the number of pitches they allow in light of this, but if they do, we imagine that they are likely to err on the high side to allow flexibility for more solo (as opposed to family) units.

As the baseline for comparison in both cases, we use the groupings identified for hunter-gatherer communities, as given by Lehmann, Lee, and Dunbar (2014), because these provide both means and variance statistics for all layers between 50 and 1,500. Lehmann, Lee, and Dunbar (2014) give mean ($\pm$ standard deviation [SD]) values of 42.8 $\pm$ 18.0, 127.3 $\pm$ 43.8, 566.6 $\pm$ 166.2, and 1,727.9 $\pm$ 620.6, respectively, for the four layers, which are numerically very similar to those obtained by Hamilton et al. (2007) using a different data set. The differences between these values and the nominal layers of 50, 150, 500, and 1,500 largely reflect ecological conditions; for example, band sizes average approximately 35 in ecologically stressed high-latitude hunter-gatherer populations but approximately 50 in low-latitude populations (Binford 2001).

It is important to be clear that the question we are asking is not whether the same kind of structured layering as has been found in human social organizations or personal social networks also occurs in these campsites, but rather whether the number of residents, and even the number of pitches, is dictated by this layering pattern across the range of camps. In other words, do these layers in some sense represent "sweet spots" in organization size that are preferred because they work better socially? To establish whether the same kinds of layerings occur within communities, we would need data of a very different kind (namely, data on interaction frequencies; e.g., Dunbar et al. 2015). Similarly, we cannot address the question of whether sites of particular size are socially optimal, because we do not have data on either community longevity or residents' satisfaction ratings. Our concern, rather, is whether the number of residents on permanent campsites favors particular values and, secondarily, whether site owners opt for particular sizes when deciding how many pitches to allow. We cannot say anything about why people choose to live in communities of particular sizes but simply ask whether they do. It is also important to emphasise that we are only concerned with individuals who live permanently on these sites and who regard the campsite as their main residence for legal purposes (as is now permitted in Germany). We are not concerned with temporary campers, who in any case invariably occupy a different part of the campsite from permanent campers.

Methods

Lists of camping sites in Germany that allow permanent camping were derived from online sources (http://www.mobilheim-forum.de, http://www.lebenaufdemcampingplatz.de, and http://www.camping.info). We used these sites to estimate camping community size in two separate ways. Most of these sites also offered camping places for temporary campers in addition to permanent residents. On most campsites, temporary and permanent camping areas are separate. We are concerned here only with the permanent camping places.

First, 53 camping sites that were referred to on the forums of the first two web pages or found via the search functionality of the third web page were contacted directly, and the number of residents on the site was obtained from the site office. Some could not provide us with the exact number of permanent campers but only with the number of permanent pitches that were occupied. In these cases, we asked the administrator to estimate the average number of campers per pitch on the site (this number consistently was either 2, 2.5, or 3 campers per pitch), and an average figure of 2.5 campers per pitch was then used to estimate the number of permanent campers at the site. The average for the number of people per pitch for those sites that provided this information was 2.39 (SD: 0.39; range: 1–3; $N = 25$). Second, we used a data set from a camping guide containing 1,216 camping sites, which gave the number of permanent camping pitches at each site (ADAC 2014). From these, 1,123 offered at least one permanent camping pitch. We used the average number of 2.5 campers per permanent camping pitch to estimate the maximum total number of permanent campers possible at the site.

We attempt to detect the clusters by two different methods: (i) we fit different distributions to the data and test whether a composite of distributions is a likely candidate, and (ii) we apply a clustering algorithm to the data and test whether the means and layers are similar to the layer sizes observed in other data sets.

To find a fit to the distribution, we use the method of maximum likelihood similar to Clauset, Shalizi, and Newman (2009). Because our data comprise positive integers, we treat distributions in a discrete manner. This involves normalizing the distribution by

$$\sum_{k=0}^{k_{\min}} \rho_k + \sum_{k=k_{\min}}^{\infty} p_k = 1, \tag{1}$$

where p_k is the distribution we are interested in for the variable k, and ρ_k is the distribution below some minimum value for that variable $k_{\min}$ in case the distribution p_k is only exhibited in the tail. We next numerically maximize the log of the likelihood to estimate the parameters for different distributions. The numerical maximization is implemented with the scipy (ver. 0.17.1) library in Python. We then compute the Akaike information criterion (AIC) for each of the candidate models (Akaike 1974) and identify the model with the lowest value of the AIC as being the most likely of the candidate models.

We tested the following models: power law, Gaussian, lognormal, geometric, compound Poisson, and compound negative binomial distributions (noting again that these are treated in a discrete manner). In the case of the last two, we treat the data as being made up of 1 to *n* distributions and calculate the AIC for each *n*, stopping when we reach a local minimum to give the optimal *n*. If the distribution is best fitted by one of these composite distributions, then the estimates for the means from the maximum likelihood parameters give the mean layer sizes.

We also apply a clustering algorithm to the data. To find the optimal number of clusters, we use the method of goodness of variance fit, also known as Jenks natural breaks optimization (Jenks 1967). This is an iterative process that moves values between clusters until the variance within each cluster is minimized. A goodness of fit value is calculated for different numbers of clusters. A goodness of fit of 1.0 can be attained only when there is zero within-class variation, which will typically be the case when the number of clusters is the same as the sample size. Here we follow Coulson (1987) and take a value of 0.85 as the threshold. The advantage of using the Jenks algorithm over other clustering techniques is that it is designed for one-dimensional data such as we have here. The data can be found at https://osf.io/v8jaf/.

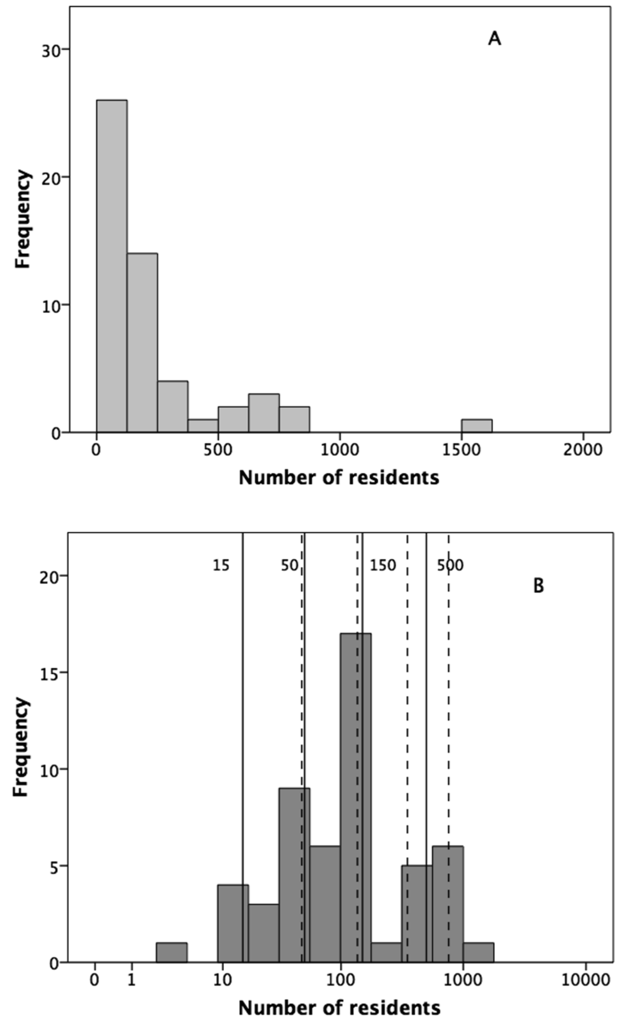

Figure 1. Frequency distribution of actual permanent residents at the 45 camping sites in Germany that provided numbers of actual permanent campers (*A*) and number of residents $\log_{10}$ transformed (*B*). The solid vertical lines demarcate values of 15, 50, 150, and 500; the dashed lines identify the cluster means identified by the compound negative binomial.

Results

Figure 1*A* plots the frequency distribution of the number of actual permanent residents at the 53 camping sites that provided this information. The distribution is highly skewed with a long right tail and a geometric mean of 97.5. A log-transformation yields a more normalized distribution (fig. 1*B*). For illustration, the values at 15, 50, 150, and 500 are superimposed as solid vertical lines on figure 1*B*. The binning in figure 1 is, of course, somewhat arbitrary (albeit determined by the SPSS software) and not of particular significance of itself, although the correspondence between the peaks and the theoretical values is striking nonetheless. The more important issue is whether the data themselves exhibit any kind of structure.

The AIC values using maximum likelihood estimates to the distributions described above are reported in table 1. The two most likely candidates are a compound of five Poisson distributions and a compound of four negative binomial distributions, the latter receiving the most support. The means for the compound Poisson distribution are 16.2, 56.4, 139.6, 350.0, and 677.2, whereas for the compound negative binomial, the means at each peak calculated from the parameter estimates are 47.3, 136.4, 349.8, and 759.5.

We also use the Jenks algorithm as described above to detect clusters in the data. This finds five clusters with means at 42.0 (23 cases), 139.6 (17 cases), 350 (5 cases), 623 (5 cases), and 1,075 (3 cases), with a scaling ratio of 2.33. The first three of these are remarkably close to the estimates from the compound negative binomial distribution and individually are not significantly different from the equivalent values of 42.8, 127.3, 566.6, and 1,727.9 given by the hunter-gatherer data (if we allow that the 350 and 623 are essentially picking up the same grouping level, with an average of 486.5; $Z = -0.04$, $Z = -0.28$, $Z = -0.34$, and $Z = -1.05$; $.968 \geq P \geq .297$).

We next apply the same method to the larger campsite data set. Although the values we have in this case are maximum possible community sizes and not actual number of residents, nonetheless we can ask whether they exhibit a similar patterning to the smaller data set. Figure 2*A* plots the distribution of estimated maximum residential capacity in the 1,123 campsites, and figure 2*B* plots this on a log-transformed scale. The geometric mean is 177.8. Once again, the dashed vertical lines indicate communities of 15, 50, 150, and 500 individuals. From

Table 1. Akaike information criteria for the candidate models applied to each data set

Model	Small data set (n = 53)	Large data set (n = 1,123)
Power law	749.4	17,390.0
Discrete normal	700.8	15,396.7
Lognormal	668.0	15,193.3
Geometric	677.5	15,145.3
Poisson	14,590.1	327,200.0
Negative binomial	678.4	15,157.8
Compound Poisson	**614.8**	46,579.9
Compound negative binomial	**558.4**	**12,809.4**

Note. The lower the Akaike information criteria value, the more likely the candidate model is. Values in bold are the most likely candidates for that data set.

the maximum-likelihood approach, we find that a compound negative binomial is also the most likely of the candidate models (table 1). This has means at 72.2, 233.6, 438.2, and 936.4. The Jenks algorithm also has an optimized value of four clusters which have means of 107.4 (628 cases), 369.2 (324 cases), 825.4 (142 cases), and 1,622.9 (29 cases), with a scaling ratio of 2.55. The values in the Jenks case are larger than those for the compound negative binomial, but in both cases the cluster means are larger in this data set than in the small data set.

Discussion

We examined the size distribution of communities in what was, in effect, a natural experiment created by the provision of permanent pitches at a large number of German campsites. The distributions are highly skewed in both of the data sets we had available, but their geometric means (~98 and ~178) straddle the value of 150 observed for natural human networks (and are within the natural range of variation for this value; Hill and Dunbar 2003). Partitioning the data sets into what seem to be natural subclusters suggests an even closer fit, with peaks at ~140 and ~107 for the small and large data sets, respectively. There is considerable evidence for the existence of a natural community size of approximately 150 in both ethnographic (Alberti 2014; Dunbar 2008) and online (Fuchs et al. 2014; Gonçalves, Perra, and Vespignani 2011; Haerter, Jamtveit, and Mathiesen 2012; Pollet, Roberts, and Dunbar 2011) environments, and the fact that the average camp community size in the present samples is in this same area adds support to the claim that this is an optimal or preferred community size. More generally, analysis of the substructuring patterns in both data sets suggests peaks at values that approximate the observed social layering values of 50, 150, and 500 observed in natural communities.

Not too surprisingly, the fit is rather better for the number of actual residents than it is for the maximum residential capacity: the latter will inevitably be driven by the site owners' economic interest in maximizing the number of residents, whereas the former is the outcome of a conventional joiner-leaver game in which individuals decide whether a particular community provides an appropriate social environment. Camp owners should always want to have more pitches available than they think they can fill, partly because they will want the flexibility to allow more single occupants, and in part because it will always be to their financial advantage to have a few extra people on their site (providing these do not exceed any optimal limit by too many and so disturb the community's equanimity). Nonetheless, the fact that both data sets are in broad agreement suggests that campsite owners must have, in their minds, a sense of what the ideal size is. It seems that these particular community sizes are in some way socially optimal and act as attractors when individuals, couples, or families decide to join or leave a community. It would be particularly illuminating to observe the conditional frequencies with which individuals joined and left

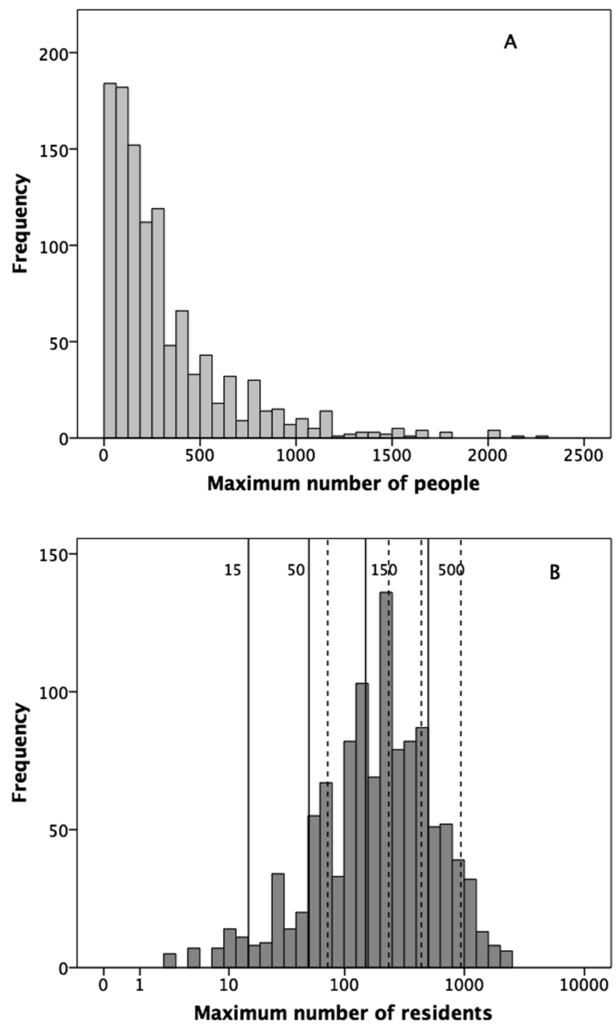

Figure 2. Frequency distribution of maximum number of residents for the 1,123 campsites from the camping guide (assuming an average number of 2.5 campers per pitch; A) and maximum number of residents $\log_{10}$ transformed (B). The solid vertical lines demarcate values of 15, 50, 150, and 500; the dashed lines identify the cluster means identified by the compound negative binomial.

communities so as to test the hypothesis that joining rates are higher on the downside of each number and leaving rates higher on the upside.

The hierarchically inclusive layered structuring of both natural communities and personal social networks is a consequence of a combination of cognitive limits on the number of relationships that can be maintained at a given emotional intensity and the time available to maintain such relationships (Miritello et al. 2013; Roberts and Dunbar 2011*a*, 2011*b*; Saramäki et al. 2014; Sutcliffe et al. 2012; Roberts et al. 2014), with very characteristic frequencies of interaction for each layer that are, in fact, mirrored even in the online world (Arnabaldi et al. 2015; Dunbar et al. 2015). We might expect the camp communities to be organized in the same way. Although outside the scope of this study, obvious predictions to test are (1) that average contact frequencies among camp members are higher in the smaller campsites than in the larger ones and (2) that large campsites are divided into subcommunities that interact within themselves especially frequently and between themselves only rarely. To test this, we would need data on interaction frequencies between individuals within the set of permanent residents. Such data would require detailed interviews that we are not able to perform, although, in principle, this could be done.

The existence of this apparently natural structuring to communities raises the question as to the functional significance of these layers. Whether the layers have functional properties or are simply an emergent property of how relationships are organized remains to be resolved, although some adaptive functions have been suggested for the various layers (Lehmann, Lee, and Dunbar 2014; Sutcliffe et al. 2012). Even so, it seems that, whereas specific functions can be ascribed to the different layers, there is some inflexibility in the number of individuals on whom one can draw for these functions, suggesting that there may be intrinsic constraints on layer size that require more detailed investigation. These "sweet spots" may arise because they allow natural communities to grow within them. Thus, 50 individuals may represent a natural social grouping (in the world of personal social networks, it is the set of individuals that provides the bulk of one's regular social contacts and all of one's emotional and economic support; Roberts et al. 2014; Sutcliffe et al. 2012), and functional communities must either be of this size or some multiple of this so as to allow several such self-contained communities to coexist.

Acknowledgments

We thank Paul Neupert (Humboldt University Berlin) for providing us with the data set of more than 1,200 camping sites and the number of available permanent camping pitches on these sites. P. Mac Carron and R. I. M. Dunbar are supported by a European Research Council advanced investigator grant (295663) to R. I. M. Dunbar.

References Cited

ADAC (Allgemeiner Deutscher Automobil-Club). 2014. *ADAC Campingführer 2014 Deutschland und Nordeuropa*. Munich: ADAC.

Akaike, H. 1974. A new look at the statistical model identification. *Automatic Control* 19:716–723.

Alberti, G. 2014. Modeling group size and scalar stress by logistic regression from an archaeological perspective. *PLoS One* 9:e91510.

Arnabaldi, V., A. Pessarella, M. Conti, and R. I. M. Dunbar. 2015. *Online social networks: human cognitive constraints in Facebook and Twitter personal graphs*. Amsterdam: Elsevier.

Binford, L. R. 2001. *Constructing frames of reference: an analytical method for archaeological theory building using ethnographic and environmental data sets*. Berkeley: University of California Press.

Clauset, A., C. R. Shalizi, and M. E. J. Newman. 2009. Power-law distributions in empirical data. *SIAM Review* 51:661–703.

Coulson, M. R. 1987. In the matter of class intervals for choropleth maps: with particular reference to the work of George F. Jenks. *Cartographica: The International Journal for Geographic Information and Geovisualization* 24:16–39.

Dunbar, R. I. M. 1992. Neocortex size as a constraint on group size in primates. *Journal of Human Evolution* 22:469–493.

———. 1993. Coevolution of neocortex size, group size and language in humans. *Behavioral and Brain Sciences* 16:681–735.

———. 2008. Mind the gap: or why humans aren't just great apes. *Proceedings of the British Academy* 154:403–423.

Dunbar, R. I. M., V. Arnaboldi, M. Conti, and A. Passarella. 2015. The structure of online social networks mirrors those in the offline world. *Social Networks* 43:39–47.

Fuchs, B., D. Sornette, and S. Thurner. 2014. Fractal multi-level organisation of human groups in a virtual world. *Scientific Reports* 4:6526.

Gonçalves, B., N. Perra, and A. Vespignani. 2011. Modeling users' activity on Twitter networks: validation of Dunbar's number. *PLoS One* 6(8):e22656.

Haerter, J. O., B. Jamtveit, and J. Mathiesen. 2012. Communication dynamics in finite capacity social networks. *Physics Review Letters* 109(16):168701.

Hamilton, M. J., B. T. Milne, R. S. Walker, O. Burger, and J. H. Brown. 2007. The complex structure of hunter-gatherer social networks. *Proceedings of the Royal Society B: Biological Sciences* 274(1622):2195–2202.

Hill, R. A., A. Bentley, and R. I. M. Dunbar. 2008. Network scaling reveals consistent fractal pattern in hierarchical mammalian societies. *Biology Letters* 4(6):748–751.

Hill, R. A., and R. I. M. Dunbar. 2003. Social network size in humans. *Human Nature* 14(1):53–72.

Jenks, G. F. 1967. The data model concept in statistical mapping. *International Yearbook of Cartography* 7:186–190.

Kanai, R., B. Bahrami, R. Roylance, and G. Rees. 2011. Online social network size is reflected in human brain structure. *Proceedings of the Royal Society B: Biological Sciences* 279(1732):1327–1334.

Lehmann, J., P. C. Lee, and R. I. M. Dunbar. 2014. Unravelling the evolutionary function of communities. In: *Lucy to language: the benchmark papers*. R. I. M. Dunbar, C. Gamble, and J. A. J. Gowlett, eds. Pp. 245–276. Oxford: Oxford University Press.

Lewis, P. A., R. Rezaie, R. Browne, N. Roberts, and R. I. M. Dunbar. 2011. Ventromedial prefrontal volume predicts understanding of others and social network size. *Neuroimage* 57(4):1624–1629.

Miritello, G., E. Moro, R. Lara, R. Martínez-López, J. Belchamber, S. B. G. Roberts, and R. I. M. Dunbar. 2013. Time as a limited resource: communication strategy in mobile phone networks. *Social Networks* 35(1):89–95.

Pollet, T., S. B. G. Roberts, and R. I. M. Dunbar. 2011. Use of social network sites and instant messaging does not lead to increased offline social network size, or to emotionally closer relationships with offline network members. *Cyberpsychology, Behavior and Social Networking* 14(4):253–258.

Powell, J. L., G. J. Kemp, R. I. M. Dunbar, N. Roberts, V. Sluming, and M. García-Fiñana. 2014. Different association between intentionality competence and prefrontal volume in left- and right-handers. *Cortex* 54:63–76.

Powell, J., P. A. Lewis, N. Roberts, M. García-Fiñana, and R. I. M. Dunbar. 2012. Orbital prefrontal cortex volume predicts social network size: an imaging study of individual differences in humans. *Proceedings of the Royal Society B: Biological Sciences* 279(1736):2157–2162.

Roberts, S. B. G., H. Arrow, J. Lehmann, and R. I. M. Dunbar. 2014. Close social relationships: an evolutionary perspective. In *Lucy to language: the Benchmark Papers*. R. I. M. Dunbar, C. Gamble, and J. A. J. Gowlett, eds. Pp. 151–180. Oxford: Oxford University Press.

Roberts, S. B. G., and R. I. M. Dunbar. 2011*a*. The costs of family and friends: an 18-month longitudinal study of relationship maintenance and decay. *Evolution and Human Behavior* 32(3):186–197.

———. 2011*b*. Communication in social networks: effects of kinship, network size and emotional closeness. *Personal Relationships* 18(3):439–452.

Saramäki, J., E. Leicht, E. López, S. B. G. Roberts, F. Reed-Tsochas, and R. I. M. Dunbar. 2014. The persistence of social signatures in human communication. *Proceedings of the National Academy of Sciences of the USA* 111(3):942–947.

Soares, P. A. d. S. 2013. Wohnen auf Rädern. *Der Spiegel*. http://www.spiegel.de/spiegel/print/d-122579481.html. Accessed January 13, 2014.

Sutcliffe, A. J., R. I. M. Dunbar, J. Binder, and H. Arrow. 2012. Relationships and the social brain: integrating psychological and evolutionary perspectives. *British Journal of Psychology* 103(2):149–168.

Zhou, W-X., D. Sornette, R. A. Hill, and R. I. M. Dunbar. 2005. Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B: Biological Sciences* 272(1572):439–444.