The structure of negative social relationships

Ignacio Tamarit[1,2*], **María Pereda**[1,2,3], **José A. Cuesta**[1,2,4,5], **Robin I.M. Dunbar**[6,7], **and Angel Sánchez**[1,2,4,5]

[1]Grupo Interdisciplinar de Sistemas Complejos, Departamento de Matemáticas, Universidad Carlos III de Madrid, 28911, Leganés, Madrid, Spain.
[2]Unidad Mixta Interdisciplinar de Comportamiento y Unidad Social (UMICSS) UC3M-UV-UZ, Universidad Carlos III de Madrid, 28911 Leganés, Spain.
[3]Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros Industriales, Departamento de Ingeniería de Organización, Administración de Empresas y Estadística, Madrid, 28006, Spain.
[4]Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, 50018, Zaragoza, Spain.
[5]Institute UC3M-BS of Financial Big Data, Universidad Carlos III de Madrid, 28903, Getafe, Spain.
[6]Department of Experimental Psychology, University of Oxford, UK. New Richards Building, Old Road Campus, Oxford OX3 7LG.
[7]Department of Computer Science, Aalto University, Finland. FI-00076 AALTO.
[*]ignacio.tamarit@uc3m.es

ABSTRACT

The composition of human personal social networks exhibits robust properties. A number of studies have shown that humans typically organize their social relationships following consistent patterns in terms of emotional closeness and communication dynamics. These regularities have been recently connected to a mathematical model describing these phenomena as a resource allocation problem. However, not all human relationships are positive, and the precise composition of negative personal networks is not yet well understood. Here we analyse data collected from a school (n=322 students) to find that individuals tend to treat positive and negative relationships as separate entities, each of which is internally organised with its own independent resource allocation pattern. We show that, although the internal organisation of negative relationship networks is independent of the equivalent positive relationship networks, it mirrors its structure. We argue that the composition of all kinds of personal networks (online, offline, workplace, positive, negative, etc.) seems to be governed by a sort of psychological *social template* that humans adapt to different social contexts, with different cognitive demands, resulting in a characteristic scaling.

Introduction

The ability to cooperate and coordinate with large groups has led to the extraordinary level of human sociality. This reflects our brain size, which had to increase, among other things, to be able to process the information of an increasingly complex society[1]. Nonetheless, this ability is finite and limits the amount of information that can be stored and processed[2]. From the perspective of a particular individual (whom we will call ego) this limitation translates into a maximum number of relationships (about 150) that can actively managed[3–7]. This personal network is the interface through by which the ego interacts with society as a whole.

In addition to a limitation to their size, a number of studies have shown that personal networks are typically organised with a hierarchical pattern consisting of a series of inclusive circles of increasing size (5, 15, 50, and 150) but decreasing intensity[3,8–10]. Although there is a continuous flux of individuals in and out of the personal networks the number of relationships is stable over time[11]. Interestingly, even though relationships are volatile, the ego still displays a distinctive and robust way to distribute her interactions that also remains stable over time[11–13]. Similar regularities have been found in the way we visit familiar locations[14]. In this case, both the number of locations (~ 25) and the individuals' time allocation for categories of places are conserved at any point—despite the fact that the total number of places visited grows sublinearly. Furthermore, this study[14] also found a significant correlation between individuals' location capacity and the size of their personal network.

All these results, taken together, suggest that the way we organise our relationships is deeply nested in our psychology[9], and seems to act as a sort of *social template* with which we handle our social interactions. The structure of such a template has been recently connected to a simple model of resource allocation[15]. The model is based on two strong and robust empirical observations. Firstly, the number of alters in the ego networks, $\mathcal{L}$, tends to be stable over time[11]. Secondly, there are different *costs* to maintaining different types of relationships[16–18], and the total cognitive resources we apply to them, $\mathcal{S}$, are limited. The maximum entropy principle[19] is then used to incorporate that information into a multinomial prior, and the result is a posterior distribution that measures the likelihood of different allocations of resources to relationships characterised with different costs

Table 1. Descriptive statistics of participants

Grade	Average age (std. err.)	Group	Participants (males)
5	10.29 (0.46)	A	22 (14)
		B	23 (15)
6	11.36 (0.47)	A	22 (11)
		B	23 (14)
7	12.47 (0.62)	A	23 (12)
		B	24 (16)
8	13.61 (0.63)	A	23 (13)
		B	28 (21)
9	14.60 (0.69)	A	27 (15)
		B	25 (18)
10	15.60 (0.64)	A	25 (13)
		B	23 (16)
11	16.58 (0.87)	A	8 (1)
		B	13 (7)
		C	12 (9)
			321 (195)

(see Ref. 15 for details)

$$P(\boldsymbol{\ell}|\mathcal{S},\mathcal{L},N) = \mathscr{B}(L,\mathcal{L}/N,N)\binom{L}{\boldsymbol{\ell}}\frac{e^{-\mu\sum_k s_k \ell_k}}{\left(\sum_k e^{-\mu s_k}\right)^L}, \quad \mathscr{B}(L,p,N) = \binom{N}{L} p^L (1-p)^{N-L} \tag{1}$$

with N the total population, ℓ_k the number of individuals in relationship $k = 1,2,\ldots,r$, s_k the cost of that relationship, and $L = \sum_k \ell_k$ the actual size of the personal network.

The only free parameter of the model, μ, arises as the Lagrange multiplier associated with the constraint on the total resources $\mathcal{S}$, and it is a function of the ratio $\mathcal{S}/\mathcal{L}$. The value of μ decreases monotonically with $\mathcal{S}/\mathcal{L}$, so high values of the ratio $\mathcal{S}/\mathcal{L}$ lead to negative values of μ and lower values lead to positive ones[15]. Notice that if $\mu = 0$ then Eq. 1 reduces to a combination of binomial (choosing L individuals from the population N) and multinomial (placing these L at random in layers) coefficients, so it is just what we would expect at random. The situation is rather different when $\mu \neq 0$. On one hand, if $\mu > 0$ then there appears a tendency towards a structure of hierarchically inclusive layers with a characteristic scaling ratio[15], in agreement with the available empirical observations[3,8–10]. In this case, the more intense (costly) a relationship the fewer individuals one can have in that layer. Additionally, if we define circle k as including all links from layers $1,2,\ldots,k$, we get that the number of individuals in consecutive circles grows with a constant factor given (approximately) by e^{μ}. On the other hand, if $\mu < 0$, the scaling is reversed and the more intense a relationship is the more alters there are—a phenomenon that had not been reported before but which, while rare, turned out to occur in real world networks under the appropriate conditions[15].

Therefore, according to the model, this characteristic *template* is just a consequence of having to allocate a finite cognitive (or time) capacity to relationships (or places) that require different amounts of resource, or effort. It is just the most likely outcome according the maximum entropy principle[19]. Notice that, although the model is presented in association with personal networks, it is a fully general description of the problem of allocating any set of objects (balls) into predefined categories (urns) with different costs.

All these regularities have been studied in considerable detail in the case of of networks of positive (i.e. affiliative) relationships. In contrast, far less attention has been given to negative relationships[20] and the precise composition of negative personal networks is not yet well understood. Nonetheless, from a socio-centric viewpoint, the interplay between positive and negative relationships forms the basis for a number of sociological theories such as Structural Balance[21,22] or Status Theory[23]. Therefore, understanding the organisation of negative networks from the ego viewpoint may help us understand its influence in the overall social structure. In what follows, we provide empirical evidence showing that networks of negative relationships mirror what we find in positive (affiliative) networks, albeit with a very different behaviour from the positive ones, and that, despite this, the above model of resource allocation also applies to them.

Table 2. Summary of network measures

	Average degree		Reciprocity	Assortativity			Connected components		Number of nodes
	in (std. err.)	out (std. err.)	Avg. (std. err.)	Gender (std. err.)	Grade (std. err.)	Class (std. err.)	Weakly (WCC)	Strongly (SCC)	N
Collected networks									
SP (+)	6.23 (5.39)	6.23 (3.24)	0.45 (0.24)	0.46 (0.02)	0.84 (0.01)	0.61 (0.01)	1	17	307
DL (+)	10.96 (9.48)	10.96 (4.69)	0.45 (0.21)	0.23 (0.02)	0.87 (0.01)	0.60 (0.01)	1	29	308
LT (+)	8.17 (5.48)	8.17 (4.04)	0.47 (0.21)	0.33 (0.02)	0.88 (0.01)	0.64 (0.01)	1	15	308
WW (+)	6.55 (5.52)	6.55 (3.90)	0.36 (0.22)	0.28 (0.02)	0.91 (0.01)	0.67 (0.01)	2	27	307
DW (-)	6.83 (8.19)	6.83 (4.84)	0.18 (0.16)	-0.09 (0.02)	0.93 (0.01)	0.55 (0.01)	6	110	308
DN (-)	5.97 (7.81)	5.97 (4.67)	0.19 (0.18)	-0.08 (0.02)	0.84 (0.01)	0.46 (0.01)	6	124	305
Constructed networks									
UPN (+)	14.92 (9.09)	14.92 (5.97)	0.54 (0.16)	0.21 (0.01)	0.85 (0.01)	0.58 (0.01)	1	1	308
UNN (-)	9.04 (9.88)	9.04 (6.04)	0.23 (0.18)	-0.07 (0.02)	0.86 (0.01)	0.49 (0.01)	1	63	308

Results

We present results from a sociometric study conducted in a school (Madrid, May 2017) with students from grades 5 to 11—the highest grade offered in the school. The total number of students within this range was 322, of whom 321 (195 males) agreed to participate and 308 ended up participating (see Table 1 for a summary of the characteristics of the participants).

In order to collect the social information, we asked the participants six questions about both their positive and negative relationships within the school. To that end, we developed a web tool, *ConectAula*, so that the students could easily find others using an organised name roster (see Methods for details). The questions we used were in Spanish so we provide here a translated version:

- If you had a serious personal problem, which other schoolmates would you be willing to share it with? (SP)
- If there are schoolmates that you wouldn't want (by any means) they had to leave the school, please mark them. (DL)
- If you could choose with whom to seat at the lunch table (independently of the actual size of the tables), who would you choose? (LT)
- If there are schoolmates with whom you'd rather not do any kind of activity, please mark them. (DN)
- If you had to complete an assignment in the school, who would you like to work with? (Mark 'None' if you'd rather work on your own) (WW)
- Mark the schoolmates with whom you'd rather not work in a school assignment (regardless of your personal relationship with them). (DW)

The answers to these questions allow us to recreate the social structure of the school. For each question, we create a directed network in which there is an edge from node i to node j if student i nominated student j in that particular question. We name these networks: SP (*share problem*), DL (*don't leave*), LT (*lunch table*), WW (*work with*), DW (*don't work*), and DN (*do nothing*). For example, if student i nominated student j as someone with whom she would like to seat at the lunch table, then there will be a directed edge from i to j in the network LT.

Some of the features of these networks are summarised in Table 2. Notice that the number of nodes is slightly smaller in three of them. The reason is that some of the 308 questionnaires gathered had missing information, resulting in a different number of nodes in each of the networks. A number of approaches can be adopted to deal with this missing data in order to minimize its impact in the estimation of network measures[24–26]. In our worst case scenario, we have complete information of 305 out the 308 who actually answered from the 322 potential participants. Since the fraction of nodes for which we don't have information is negligible ($\sim 5\%$), and they are missing at random, we simply eliminate these nodes.

Degree correlations

The information needed to characterize the composition of an ego network from the ego point of view is encoded in her out degree distribution—or *embeddedness*. Indeed, ego's answers are the best proxy we can have for her internal representation of her social world. In Table 3 we can see how the out degree distributions are positively correlated if they belong to relationships of the same type (both positive or both negative) and, interestingly, not correlated otherwise. Therefore, there is no apparent connection between the number of reported positive relationships and the number of reported negative ones.

From a socio-centric perspective, however, the best way to characterize an ego is to look at her in-degree (or *popularity*), since it reflects how the ego is seen by others. What we observe (see Table 3) is that the distributions are (highly) positively

Table 3. Degree correlations. The upper triangular part of the table shows the Pearson correlation coefficient for each pair of variables. The lower triangular shows the 99% confidence interval for each coefficient, computed using BCa Bootstrap[27,28] with 10^6 samples in each case. In bold, the values for which the confidence interval does not cross zero.

		SP		DL		LT		WW		DW		DN	
		out	in	out	in	out	in	out	in	out	in	out	in
SP	out	——	0.08	**0.36**	0.06	**0.51**	0.07	**0.40**	0.03	0.00	−0.07	−0.02	−0.05
	in	(−0.09, 0.26)	——	0.07	**0.68**	0.05	**0.66**	0.00	**0.69**	−0.02	−**0.43**	−0.04	−**0.38**
DL	out	**(0.18, 0.53)**	(−0.08, 0.22)	——	0.14	**0.43**	**0.18**	**0.20**	0.15	0.04	−0.12	0.02	−0.10
	in	(−0.09, 0.20)	**(0.58, 0.76)**	(−0.01, 0.31)	——	0.05	**0.78**	−0.04	**0.72**	−0.01	−**0.45**	−0.01	−**0.44**
LT	out	**(0.31, 0.68)**	(−0.10, 0.20)	**(0.26, 0.59)**	(−0.09, 0.19)	——	0.08	**0.42**	0.05	0.01	−0.10	0.10	−0.05
	in	(−0.09, 0.21)	**(0.56, 0.74)**	**(0.01, 0.37)**	**(0.71, 0.83)**	(−0.08, 0.22)	——	−0.03	**0.70**	−0.03	−**0.43**	−0.03	−**0.38**
WW	out	**(0.20, 0.60)**	(−0.15, 0.14)	**(0.04, 0.38)**	(−0.18, 0.12)	**(0.19, 0.61)**	(−0.16, 0.10)	——	0.03	0.02	−0.02	0.13	−0.01
	in	(−0.13, 0.19)	**(0.59, 0.76)**	(−0.01, 0.32)	**(0.63, 0.79)**	(−0.09, 0.19)	**(0.60, 0.77)**	(−0.11, 0.17)	——	0.07	−**0.56**	−0.02	−**0.47**
DW	out	(−0.12, 0.17)	(−0.14, 0.11)	(−0.09, 0.17)	(−0.14, 0.12)	(−0.12, 0.17)	(−0.16, 0.11)	(−0.10, 0.12)	(−0.06, 0.20)	——	−0.07	**0.62**	−0.07
	in	(−0.19, 0.08)	**(−0.53, −0.32)**	(−0.24, 0.03)	(−**0.56**, −**0.32**)	(−0.22, 0.03)	(−**0.54**, −**0.29**)	(−0.15, 0.12)	(−**0.63**, −**0.47**)	(−0.19, 0.06)	——	0.06	**0.85**
DN	out	(−0.15, 0.12)	(−0.17, 0.10)	(−0.12, 0.17)	(−0.13, 0.13)	(−0.06, 0.29)	(−0.16, 0.12)	(−0.00, 0.27)	(−0.14, 0.12)	**(0.28, 0.77)**	(−0.09, 0.25)	——	0.10
	in	(−0.18, 0.09)	(−**0.49**, −**0.26**)	(−0.23, 0.04)	(−**0.54**, −**0.31**)	(−0.17, 0.07)	(−**0.49**, −**0.26**)	(−0.13, 0.13)	(−**0.56**, −**0.36**)	(−0.19, 0.07)	**(0.81, 0.89)**	(−0.06, 0.36)	——

correlated if they belong to the same type and, unlike the out-degree distributions, negative otherwise. Unsurprisingly, this result shows that individuals who are popular in a network of positive relationships tend to be so in all of them, and not 'popular' in negative ones—and vice versa for individuals who are listed frequently in a network of negative relationships. In summary, from the ego's viewpoint, the positive and negative networks seem to be handled as independent entities: from a socio-centric perspective, being popular in one type of network is closely related to not being popular in the other.

Overlapping of networks and social circles

If we focus on the positive networks, we can appreciate how the answers (*out edges*) overlap to a great extent (see Fig. 1), meaning that the participants typically nominated the same alter in more than one network. In fact, if we consider the network constructed as the union of all the positive networks (UPN), we see that the average degree becomes 14.92 (see Table 2) as opposed to 31.91, which would be the expected value if the networks were absolutely disjoint. This number (~ 15) is precisely the typical size of the *sympathy group*, defined as '*all the people whose death tomorrow would cause great distress*'[18,29], and corresponds with the second of a series of circles with sizes: 5, 15, 50, 150, in which an ego normally organizes her social relationships[3] —see Introduction. We include a summary of network measures for UPN in Table 2. Remarkably, this network is the only one with a single strongly connected component (SCC) —that is, for every pair of nodes i, j there exists a path from i to j and also a path (possibly involving different nodes) from j to i.

Since the network UPN includes the *sympathy groups* of the students, we explore whether there is any evidence for the first of the circles, the *support network*[18], in our data. This network represents the closest relationships of an ego, her primary source of emotional support, advice, and assistance in time of need, and has a typical size of[3,30] ~ 5. Among all the questions we asked, the one with the closest meaning is clearly the question about sharing a serious personal problem —SP. The average degree in this case is 6.23 (see Table 2), which would be in good agreement with the commonly observed value of ~ 5.

Let us focus now on the negative relationships. In fig 1 we can appreciate how DN and DW have also a strong level of overlap between them and very little ($Jaccard = 0.05$) overlap with the positive relationships—which was not unexpected. Additionally, the union of the negative networks (UNN) has degree 9.04 and the *most intense* one (DN) has an average degree of 5.97. The similarities between the positive and negative networks make us wonder whether more information would have led us to find an exact (negative) parallel of the sympathy group with ~ 15 antagonistic relationships. A meta study on antipathetic relationships[20] showed that unlimited choice procedures and the use of less intense items assessing dislike (for example, *least like* instead of *enemy*) lead to a higher prevalence of antipathetic relationships. Therefore, it is reasonable to expect that we would have obtained a higher number of antipathetic relationships if more questions designed to elicit *weak*, negative relationships had been included. Although we do not have enough information to answer this question, both the independence of the degree distributions (see previous section) and the (apparently) parallel overlapping behaviour, suggest that this could indeed be the case.

Scaling properties

As we described in the introduction, the model presented in Ref. 15 predicts that the distribution of the intensities (i.e. in terms of emotional closeness[18]) of human relationships may exhibit two distinct patterns, namely, hierarchical inclusive layers with a constant scaling ($\mu > 0$) or a tendency to accumulate very intense relationships leaving aside less significant ones ($\mu < 0$). This is a consequence of the different *costs* associated with the different types of relationships[16,31] —the closer the relationship, the more costly. In order to fit this model to our data we first need to assign a value to each of the relationships of the ego that

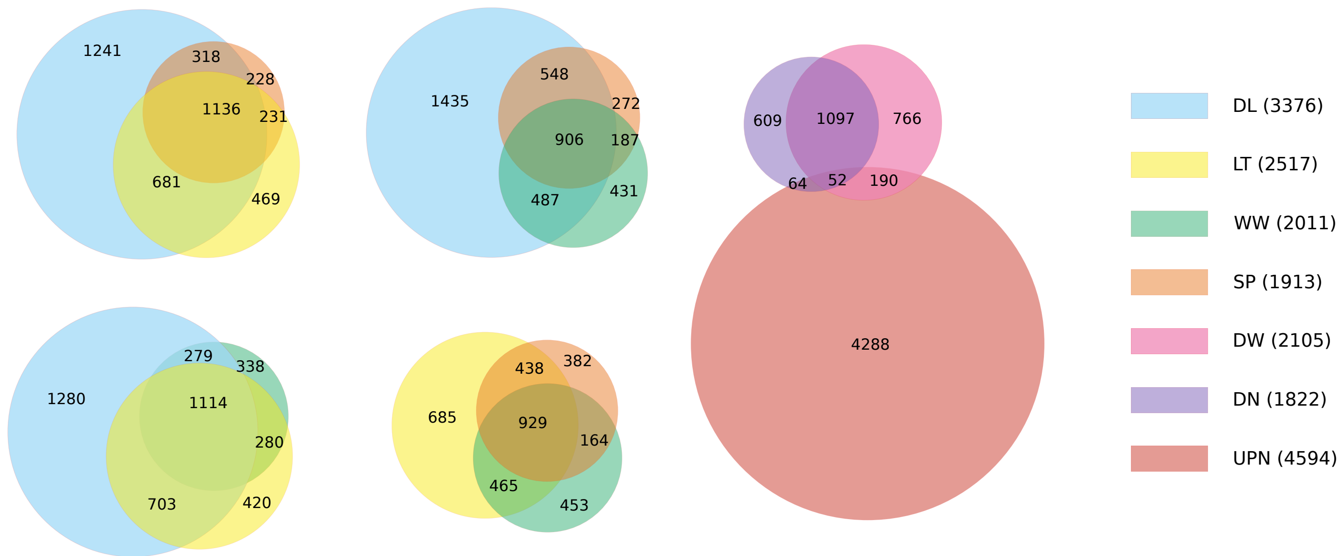

Figure 1. Overlapping of networks. The numbers in each separate region show the number of edges common to the corresponding overlapping networks. The areas are depicted for illustrative purposes and do not reflect the exact sizes of the networks.

reflects its intensity. Although we do not have any direct measure of the intensity of the reported relationships, we can use proxies to assess it.

The concept of emotional closeness is very much related to that of *strength* of a tie as discussed by Granovetter[32] and, in fact, it is considered as one of its best indicators[33]. Another indicator of the strength of a tie is clearly whether it is reciprocated or not. In his classic paper[32], Granovetter classified reciprocal links as 'strong ties' and non-reciprocal links as 'weak ties', and further studies[34] have shown that reciprocal links were associated with stronger ties than non-reciprocal (or asymmetric) ones —and those in turn revealed to be stronger than relationships not elicited by any link at all. Yet another indicator (also noted by Granovetter[32]) is the extent of multiplexity within a tie, that is, the multiple contents present in a relationship. In our setting, for example, this can be inferred by the number of nominations from an ego to a given alter. Since we have information on both the reciprocity and the multiplexity of the relationships, we build on these two ideas to create a measure of the intensity of the links.

When there are multiple networks (layers) reciprocity can occur, at least, at three levels; a link can be reciprocated in the same layer (S), in a different layer (O), or not reciprocated at all (N). From a social perspective these three types of behaviours are rather different. A reciprocal link from the same network is signalling a relationship that is not only mutual, but also agreed. Therefore, we consider that links reciprocated at the same layer (S) have a higher value (intensity, strength) than those reciprocated in any other layer (O). Similarly, we also consider that the latter (O) have a higher value than those not reciprocated at all (N).

The previous ordering of S, O and N can be extended to multiplex settings[35], in which relationships may be formed out of more than one of these types of links. All we have to do is to agree on how to compare any two pairs of these combinations. For example, we all would agree that a relationship based on four (as in our positive networks) S links has a higher value than one based on a single N, non-reciprocated link. To extend this idea we choose a very simple decision algorithm: *Take any two sets with elements* $\in \{S, O, N\}$. *Then, the one with more Ss is assigned the highest value. If a draw happens then the one with more Os wins. If a draw still happens then the one with more Ns wins —in case of a draw in all categories assign them the same value*.

If we apply the former rule to a single network the results are trivial. In this case, only two types of links can happen, S or N, and S is considered stronger. Adding an extra layer induces six different possibilities, and the above described algorithm results in the following (unique) ordering: $[S,S] > [S,N] > [S] > [O] > [N,N] > [N]$. Since we have precisely two negative networks, we can use this ordering as a proxy for the intensity of the (negative) relationship between any two individuals. The number of alters with whom an ego has a relationship of strength $[S,S]$ (assumed to have the highest cost and intensity) is assigned to ℓ_1, that corresponding to $[S,N]$ is assigned to ℓ_2, and so on until the number of relationships with strength $[N]$ (assumed to have the lowest cost) is assigned to ℓ_6. The *costs* s_k are simply assumed to decrease linearly from $[S,S]$ to $[N]$ —see Methods. The same reasoning applied to four layers results in a (unique) sequence of twenty ordered strengths: $[S,S,S,S] > [S,S,S,N] > \cdots > [N,N] > [N]$.

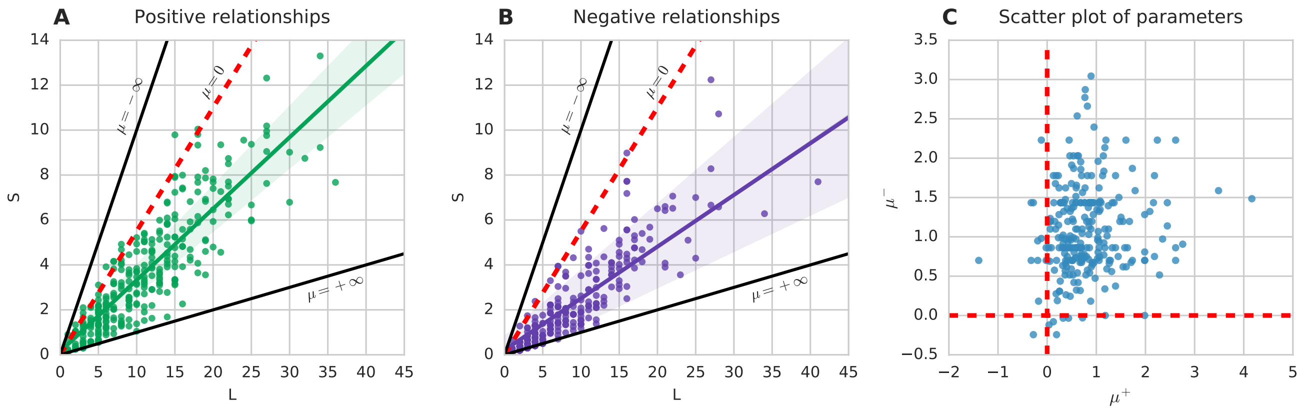

Figure 2. Parameter estimates for positive and negative relationships. (A) Linear regression $S \sim L$ for the positive relationships ($n = 300$, the individuals with no relationships have been excluded). The fitted line ($p < 10^{-3}, R^2 = 0.77$) has a slope of 0.32 ($[0.28, 0.35]_{99\%}$), resulting in an estimated value of $\mu^+ = 0.70$ ($[0.58, 0.85]_{99\%}$) —see Methods. The green, shadowed region corresponds to the 99% confidence interval for the slope and the intercept. An outlier with $L = 51, S = 13.5$ has been excluded in the plot but taken into consideration in the fitting. **(B)** Linear regression $S \sim L$ for the negative relationships ($n = 284$). The fitted line ($p < 10^{-3}, R^2 = 0.74$) has a slope of 0.23 ($[0.16, 0.29]_{99\%}$), which results in an estimated value of $\mu^- = 1.13$ ($[0.79, 1.76]_{99\%}$) —see Methods. The purple,shadowed region corresponds to the 99% confidence interval for the slope and the intercept. An outlier with $L = 85, S = 12.6$ has been excluded in the plot but taken into consideration in the fitting. **(C)** Scatter plot of pairs of parameters. The figure shows the pairs (μ^+, μ^-) for the 268 cases for which we can estimate both μ^+ and μ^- for the same individual —see Methods . The Pearson correlation coefficient is small ($r = 0.136$) and, although $p = 0.026$, the 99% confidence interval ($[-0.006, 0.277]_{99\%}$) crosses zero.

With this information we have a setting for which the model described in the introduction applies. For each individual we can now estimate his *cognitive* capacity and the size of his network as $S = \sum_{k=1}^{r} s_k \ell_k$ and $L = \sum_{k=1}^{r} \ell_k$ respectively. In Figs. 2A and 2B we show the scatter plots of the pairs (L, S) for the positive and the negative relationships. As we can see, in both cases the values are grouped around straight lines —and the 99% confidence intervals of the intercepts include the zero, as stated by the model. The slopes of these lines can be used to estimate the value of the parameter μ that characterises the population —see Methods. The positive relationships turn out to be centred around a value of $\mu^+ = 0.70$ ($[0.58, 0.85]_{99\%}$) and the negative ones around $\mu^- = 1.13$ ($[0.79, 1.76]_{99\%}$). Notice that the confidence intervals of the estimates for both types of relationships overlap, so the hypothesis that both values are indeed the same can not be rejected. In addition, these values are in good agreement with the typically observed scaling of circles[8,18] ~ 3, which would correspond to $\mu \approx 1.01$. Thus, in both cases there is a preferred, positive scaling, which translates into personal networks typically organised with disproportionally more weak relationships than strong ones[3,10,15]. Lastly, let us note that this tendency towards a preferred, positive scaling is consistent across different ages, and does not exhibit any sort of gender-wise differences.

The former analysis reveals the existence of a particular value of the parameter μ that characterises the population for both positive and negative relationships. However, there exists variability across different individuals, each of them having a different value of μ. In Fig. 2C we show the scatter plots for the cases in which we can estimate both μ^+ and μ^- for the same individual —see Methods. Although there exists a weak correlation ($r = 0.136, p = 0.026$), the 99% confidence interval ($[-0.006, 0.277]_{99\%}$) includes the zero, so the the hypothesis that there exists no correlation can not be rejected —and in any case it is weak. In agreement with the fact that the out-degree distributions are also not correlated (see Table 3), this result further suggests that the internal organisation of positive and negative personal networks follow independent (but similar) patterns.

Parameter estimates and centrality measures.

Lastly, since the parameters μ^+ and μ^- characterise the structure of the individuals' personal networks (positive and negative) with a single real number, we can think of them as attributes of the nodes and analyse their connection with other properties of the same nodes in the network. In particular, we focus on four centrality measures: degree, closeness, *PageRank*, and betweenness. We compute the value of each of these measures for every node in UPN and UNN, and look for correlations with the estimated parameters —see Methods. In the case of UPN we find some weak but significant correlations. Firstly, there is a negative correlation between μ^+ and *PageRank* ($r = -0.24, [-0.37, -0.09]_{99\%}$). Recall that *PageRank* is a measure of the relative importance of a node in a network, and that lower values of the parameter μ correspond to higher investments per relationship. Thus, this result suggests that individuals investing more in close, strong relationships tend to benefit from powerful positions in the socio-centric network. Secondly, there is a significant correlation between μ^+ and *in-degree* ($r = -0.19, [-0.33, -0.04]_{99\%}$), so that people with higher ratios of investments per link tend to be popular. Lastly, there exists also a positive correlation between μ^- and *in-degree* ($r = 0.17, [0.02, 0.31]_{99\%}$) in the network UPN. Thus, interestingly, individuals who tend to allocate few resources per negative relationship (i.e. have a higher μ^-) tend also to be popular in the network of positive relationships. In the case of the network UNN, however, we do not find any significant correlation between the analysed centrality measures and the parameter estimates.

Discussion

In this paper we have presented results from a sociometric study conducted in a school ($n = 322$). We have focused on analysing the structure of personal networks of positive and negative relationships, and found that the similarities between them are considerable. In particular, we have shown that a model of resource allocation[15] that explains the organisation of personal networks of positive relationships also applies to antagonistic ones. There are however some limitations to this study. Firstly, the intensity of the links had to be inferred from indirect measures such as *multiplexity* and reciprocity. Although these measures are reasonable proxies for the intensity of relationships, further experiments in which the intensity is measured in a more direct manner should be performed. Secondly, the number of questions about positive and negative relationships was not balanced. This is due to the fact that the experiment had other additional research questions for which these questions were specially selected. To overcome this issue, and assess the robustness of the results we present in this paper, we are currently carrying out a longitudinal study in which we measure the intensity of positive and negative relationships through specific,well-defined questions.

Our results show (see Fig. 2) that the average number of positive relationships converged to ~ 15, the typical size of *sympathy groups*, and that the number of most intimate ones (SP) was compatible with the typical size of the *support networks*[18] ~ 5. It is remarkable that these numbers arise in a setting in which the universe was restricted to schoolmates, without considering family members and other possible relationships from outside the school. However, ours is not the only setting in which a restrictive social environment exhibits numbers in agreement with these social circles. A good example is provided by on-line social networks (OSN) such as Facebook or Twitter, in which this precise structure has also been found[36]. In this case, it is likewise unlikely that all social relationships of the individuals occur within the boundaries of such platforms.

Both the school and the OSN can be regarded as closed, social environments in which an individual establishes relationships in a very concentrated social environment. The fact that the structure of personal networks in such settings resembles that of an entire, open social environment suggests that this pattern of relationships, the circles and their characteristic sizes (or their scaling), act as a sort of "template" that is reproduced at different scales, in different environments[9, 11]. Interestingly, our results further reveal that this tendency towards a positive scaling is also present when the intensity of the relationships is measured using the number of nominations (multiplexity) and its reciprocity as a proxy.

More surprising is the finding that personal networks of negative relationships seem to be organised in a similar manner. The analysis of the negative networks showed similar intersection patterns (see Fig. 1), and the network DN (that can be argued to be the most intense) exhibit sizes very similar to those of positive support networks, a finding also reported by other studies reporting negative relationships[37]. Moreover, in a meta study on antipathetic relationships[20] it was found that antipathetic relationships were also decreasingly common with more intense items assessing dislike: that is, the more intense a negative relationship, the fewer of them one has. Our results further suggest that, from the ego's viewpoint, positive and negative networks are intrinsically different, working as separate, but self-consistent, social environments. As we have seen, their degree distributions were not correlated, and both exhibited different, but self-consistent scaling patterns. Thus, from the individuals' perspective, the different types of relationships seem to be idiosyncratically different. In other words, an equivalent pattern to *Dunbar's circles* also exists for negative relationships.

In the light of the model presented in the introduction[15], these results have a precise interpretation. If we assume that individuals devote a certain amount of resource to each of the different social environments, then, the internal organization of such networks would still be governed by Eq. 1. The reason is that, in each of these contexts, a given amount of resource has to be allocated among the different relationships as if that particular context was the only one available. A good example of this would be precisely the scenario on which this paper focuses: a school. Although students may certainly have relationships outside the centre, the school's social ecosystem functions as an independent entity in which they spend a given amount of time and are *forced* to interact socially with a more or less stable set of individuals. These are the individuals with whom they share most common interests (the homophily effect[38]) and hence are the individuals they feel most emotionally engaged with.

Nevertheless, if we aggregate all possible social contexts, and their associated costs, we would recover an identical situation where the cost of a relationship is the combination of the ones in the different contexts, and the total limitation of resource is given by the overall capacity of the individual. Indeed, the resulting personal networks would be organised similarly.

Furthermore, if the actual equilibrium between resource ($\mathcal{S}$) and size of the personal network ($\mathcal{L}$) occurs around a value of $\mathcal{L} = 150$, with estimated parameter $\mu \sim 1$, then the expected numbers in an hypothetical distribution of alters into four layers would be very similar to the so called Dunbar's circles[8]: 5, 15, 50, 150. Therefore, the model gives a plausible explanation for the ubiquity of this type of structure and connects it to an overall limitation on humans' relationships.

The reason why there appears to be a discretisation in the structure of egocentric networks of this kind is uncertain. Indeed, the model allows for any number of layers (or even a continuum) and there is no a priori reason to choose one over the other. The dual interpretation of personal networks, that is, whether they have a precise discrete structure or they are a continuum (or, perhaps, both), is an interesting question to be addressed. To shed light on this issue, any closed, social environment in which one can measure the intensity of the relationships of the individuals can be used to test the hypothesis that we indeed use a sort of social template to manage the complexities of social interaction.

Lastly, our results show that there is a connection between the (internal) organisation of personal networks and socio-centric measures such as node centrality. As we have seen, the internal organisation is well explained as a consequence of a constraint in the capacity of the individuals to establish different types of relationships, and there is not any a priori reason for it to be connected with socio-centric aspects. Besides, we have shown that there exist precise values of the parameter μ that characterise the populations (see Fig. 2), which was also not required by the model specifications —notice that the parameter emerges as a characteristic of the individuals, not the populations. It is therefore plausible that the structure of our personal networks has evolved towards a particular type of structure (characterised by a central value of μ) influenced by social processes.

Methods

Experimental set-up and software.

To perform the experiment we divided the participants into groups (about 25 participants per group) and scheduled a time when their teachers had to bring them to the computer room of the school. Once there, after a brief presentation of the activity by the researcher, they were instructed to enter an online, especially designed platform (*ConectAula*) and answer the survey. The whole activity took about 30-45 minutes to be completed. The online platform allowed the participants to easily find and select any other participant (while keeping their identities anonymous to the researchers). It was possible to nominate as many alters as they wanted in each question. To facilitate the process, each of the network questions was displayed together with drop-down buttons organising the possible answers (i.e. all the participants with the exception of the ego herself) with the same structure of grades and groups as in the school. This information had been previously uploaded to the software by an employee of the school, and the data was stored codified in the database so that researchers never had access to the real identity of the participants. At any given moment, at least one researcher and one teacher were present in the room, and occasionally assisted the students who had troubles understanding a question. This procedure was repeated during three consecutive days and enabled us to collect 308 surveys —that is, 96 % of the total population. However, for various reasons, 5 of these surveys missed information of at least one of the questions about social relationships. A complete description of the number of respondents for each question can be seen in Table 2.

Estimate of the parameter.

The value of μ is determined by the value of $\mathcal{S}/\mathcal{L}$ according to (see Ref. 15 for details)

$$\frac{s_1 - \mathcal{S}/\mathcal{L}}{s_1 - s_r} = f(\mu) \equiv e^{\mu} \frac{(r-1)e^{r\mu} - re^{(r-1)\mu} + 1}{(r-1)(e^{r\mu} - 1)(e^{\mu} - 1)},$$

were s_1 is the cost associated with ℓ_1 (the most expensive), s_r the cost associated with ℓ_r (the least expensive), and r is the number of considered layers. The former expression leads to a linear relationship between $\mathcal{S}$ and $\mathcal{L}$ given by $\mathcal{S} = (s_1 - f(\mu)(s_1 - s_r))\mathcal{L} \equiv g(\mu)\mathcal{L}$. For the analysis shown in Figs. 2A and 2B we choose, without loss of generality, $s_1 = 1$, $s_r = 0.1$, and $r = 4$, and estimate $\mathcal{S}$ and $\mathcal{L}$ for each individual as $\mathcal{S} \equiv S = \sum_{k=1}^{r} s_k \ell_k$ and $\mathcal{L} \equiv L = \sum_{k=1}^{r} \ell_k$ respectively —which are indeed the maximum likelihood estimators. The slope of the linear regression, let's say x, is used to estimate μ by numerically solving the equation $g(\mu) = x$. The 99% confidence intervals for the slope and the intercept are computed using Bca Bootstrap[27,28] ($n_{samples} = 10^6$). These values are then used to compute the 99% confidence interval for the estimated μ. In the case of Fig. 2C, the individuals' parameters are estimated as explained before but using simply the corresponding value S/L of the individual —and not the estimated slope of the regression. The values excluded are those for which either $L = 0$ or $|\mu| > 10$ —which are divergences. The confidence interval for the Pearson correlation coefficient is also computed used Bca Bootstrap ($n_{samples} = 10^6$).

Notice that the actual number of layers in which we divided the positive ($r = 20$) and the negative ($r = 6$) relationships is not considered here. The reason is that the value of μ depends somewhat arbitrarily on the number of categories (layers) in which we split the relationships. This is easy to see if we take a look to the so-called *Dunbar's circles*, which consist of the hierarchical inclusion of $r = 4$ different layers and have sizes: 5, 15, 50, and 150. In this case, the ratio between consecutive circles is ~ 3. However, if the layers had been grouped differently (for example into two circles of sizes 20 and 130) the resulting scaling would be different (6.5 in our example). Therefore, different values of r simply reflect arbitrary decisions about the considered number of layers and result in different estimations of μ (and the constant scaling e^{μ}) accordingly —in fact, the model can be cast in a continuous manner with no discretisation of the layers at all (see Supplementary Information in Ref. 15 for details). It is to facilitate the comparison with the empirical results observed for *Dunbar's circles* that we choose $r = 4$ to estimate the parameter. Note, however, that the observed value of $r = 4$ for the empirical data is not arbitrary: it is the optimized value for large datasets[39].

Parameter estimates and centrality measures.
To compute the results presented in this section, we compute the centrality measures (degree, closeness, *PageRank*, and betweenness) for each of the nodes in the networks UPN and UNN. Then, using the 268 cases for which we also have both μ^{+} and μ^{-}, we compute the Pearson's correlation coefficient (r) between these estimates and the above mentioned centrality measures. The 99% confidence intervals are computed using Bca Bootstrapping ($n_{samples} = 10^{6}$). Results are considered significant in this interval does not cross the zero value.

References

1. Dunbar, R. The social brain hypothesis. *Evol. Anthropol.* **6**, 178–190 (1998).
2. Omodei, E., Brashears, M. E. & Arenas, A. A mechanistic model of human recall of social network structure and relationship affect. *Sci. reports* **7**, 17133 (2017).
3. Hill, R. A. & Dunbar, R. I. Social network size in humans. *Hum. nature* **14**, 53–72 (2003).
4. Roberts, S. G., Dunbar, R. I., Pollet, T. V. & Kuppens, T. Exploring variation in active network size: Constraints and ego characteristics. *Soc. Networks* **31**, 138–146 (2009).
5. Haerter, J. O., Jamtveit, B. & Mathiesen, J. Communication dynamics in finite capacity social networks. *Phys. review letters* **109**, 168701 (2012).
6. Gonçalves, B., Perra, N. & Vespignani, A. Modeling users' activity on twitter networks: Validation of dunbar's number. *PloS one* **6**, e22656 (2011).
7. Dunbar, R. I. The social brain psychological underpinnings and implications for the structure of organizations. *Curr. Dir. Psychol. Sci.* **23**, 109–114 (2014).
8. Zhou, W.-X., Sornette, D., Hill, R. A. & Dunbar, R. I. Discrete hierarchical organization of social group sizes. *Proc. Royal Soc. Lond. B: Biol. Sci.* **272**, 439–444 (2005).
9. Fuchs, B., Sornette, D. & Thurner, S. Fractal multi-level organisation of human groups in a virtual world. *Sci. reports* **4**, 6526 (2014).
10. Dunbar, R. I. Do online social media cut through the constraints that limit the size of offline social networks? *Open Sci.* **3**, 150292 (2016).
11. Saramäki, J. *et al.* Persistence of social signatures in human communication. *Proc. Natl. Acad. Sci.* **111**, 942–947 (2014).
12. Heydari, S., Roberts, S. G., Dunbar, R. I. & Saramäki, J. Multichannel social signatures and persistent features of ego networks. *Appl. Netw. Sci.* **3**, 8 (2018).
13. Liu, J.-G., Guo, Q., Zhang, Y.-C. *et al.* Social signature identification of dynamical social networks. *Phys. A: Stat. Mech. its Appl.* **508**, 213–222 (2018).
14. Alessandretti, L., Sapiezynski, P., Sekara, V., Lehmann, S. & Baronchelli, A. Evidence for a conserved quantity in human mobility. *Nat. Hum. Behav.* 1 (2018).
15. Tamarit, I., Cuesta, J. A., Dunbar, R. I. & Sánchez, A. Cognitive resource allocation determines the organization of personal networks. *Proc. Natl. Acad. Sci.* 201719233 (2018).
16. Oswald, D. L., Clark, E. M. & Kelly, C. M. Friendship maintenance: An analysis of individual and dyad behaviors. *J. Soc. clinical psychology* **23**, 413–441 (2004).
17. Miritello, G. *et al.* Time as a limited resource: Communication strategy in mobile phone networks. *Soc. Networks* **35**, 89–95 (2013).
18. Dunbar, R. The anatomy of friendship. *Trends cognitive sciences* **22**, 32–51 (2018).
19. Jaynes, E. T. *Probability theory: The logic of science* (Cambridge university press, 2003).

20. Card, N. A. Antipathetic relationships in child and adolescent development: A meta-analytic review and recommendations for an emerging area of study. *Dev. Psychol.* **46**, 516 (2010).

21. Heider, F. Attitudes and cognitive organization. *The J. psychology* **21**, 107–112 (1946).

22. Cartwright, D. & Harary, F. Structural balance: a generalization of heider's theory. *Psychol. review* **63**, 277 (1956).

23. Leskovec, J., Huttenlocher, D. & Kleinberg, J. Signed networks in social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 1361–1370 (ACM, 2010).

24. Robins, G., Pattison, P. & Woolcock, J. Missing data in networks: exponential random graph (p) models for networks with non-respondents. *Soc. networks* **26**, 257–283 (2004).

25. Kossinets, G. Effects of missing data in social networks. *Soc. networks* **28**, 247–268 (2006).

26. Huisman, M. Imputation of missing network data: some simple procedures. *J. Soc. Struct.* **10**, 1–29 (2009).

27. Bishara, A. J. & Hittner, J. B. Testing the significance of a correlation with nonnormal data: comparison of pearson, spearman, transformation, and resampling approaches. *Psychol. methods* **17**, 399 (2012).

28. DiCiccio, T. J. & Efron, B. Bootstrap confidence intervals. *Stat. science* 189–212 (1996).

29. Buys, C. J. & Larson, K. L. Human sympathy groups. *Psychol. Reports* **45**, 547–553 (1979).

30. Dunbar, R. I. & Spoors, M. Social networks, support cliques, and kinship. *Hum. Nat.* **6**, 273–290 (1995).

31. Takano, M. Two types of social grooming methods depending on the trade-off between the number and strength of social relationships. *Royal Soc. Open Sci.* **5**, 180148 (2018).

32. Granovetter, M. S. The strength of weak ties. In *Social networks*, 347–367 (Elsevier, 1977).

33. Marsden, P. V. & Campbell, K. E. Measuring tie strength. *Soc. forces* **63**, 482–501 (1984).

34. Friedkin, N. A test of structural features of granovetter's strength of weak ties theory. *Soc. networks* **2**, 411–422 (1980).

35. Kivelä, M. *et al.* Multilayer networks. *J. complex networks* **2**, 203–271 (2014).

36. Dunbar, R. I., Arnaboldi, V., Conti, M. & Passarella, A. The structure of online social networks mirrors those in the offline world. *Soc. Networks* **43**, 39–47 (2015).

37. Berger, C. & Dijkstra, J. K. Competition, envy, or snobbism? how popularity and friendships shape antipathy networks of adolescents. *J. Res. on Adolesc.* **23**, 586–595 (2013).

38. McPherson, M., Smith-Lovin, L. & Cook, J. M. Birds of a feather: Homophily in social networks. *Annu. review sociology* **27**, 415–444 (2001).

39. Mac Carron, P., Kaski, K. & Dunbar, R. Calling dunbar's numbers. *Soc. Networks* **47**, 151–155 (2016).

Acknowledgements

We thank the school administration, teachers, families, and students for their participation in this study. We also thank Juan A. Zamora and Nieves Maestro for developing the software *ConectAula*. This study was supported in part by Fundación BBVA through grant DUNDIG (I.T., M.P., and A.S.), by Ministerio de Economía, Innovación y Competitividad (Spain) through grants FIS2015-64349-P VARIANCE (MINECO/FEDER, UE) (I.T., M.P., J.A.C. and A.S.), and by the European Commission through FET Open RIA 662725 IBSEN (I.T., M.P., R.I.M.D., J.A.C., and A.S.) and FET Proactive RIA 640772 DOLFINS (A.S.). R.I.M.D. was supported by the European Research Council Advanced Investigator through grant [295663].

Author contributions statement

All authors conceived the experiment, I.T., M.P., and J.A.C. conducted the experiment, I.T. analysed the results, I.T. wrote the manuscript. All authors reviewed the manuscript.

Additional information

The experiment was approved by the Ethics Committee of the *Universidad Carlos III de Madrid* and the school gathered parental consent agreements for all the participants.

The authors declare no conflict of interest.