Social Networks 8 (1986) 387–396
North-Holland

A NOTE ON SCALING THE GENERAL SOCIAL SURVEY NETWORK ITEM RESPONSE CATEGORIES *

Ronald S. BURT and Miguel G. GUILARTE **
Columbia University

The idea of structural balance is used to suggest quantitative intervals between relationship strength response categories in the GSS network data. In contrast to an assumption of equal intervals between the categories of relationship strength, the intervals appear quite unequal. Relations with discussion partners "less close" to their respondent than other cited discussion partners are about 0.17 the strength of relations with "especially close" discussion partners. The middle category of relations between discussion partners appear to be little more than acquaintance relations; about 0.2 of the distance from people who are "total strangers" to people who are "especially close".

1. Introduction

In the near future, a great many people interested in interpersonal relations and network theory are likely to study the 1985 General Social Survey (GSS) network data. These unusual data describe the interpersonal environment surrounding each of the 1531 respondents in a national probability sample of adult Americans during the winter of 1985. The network data describe relations with and among up to five of each respondent's important discussion partners and the structure of these relations can be studied independently or used to predict the data on respondent opinion and behavior routinely collected in the survey. [1]

* This technical note is a by-product of support from the National Science Foundation Sociology Program (SES-8208203) and the Measurement Methods and Data Improvement Program (SES-8513327) and has been produced as part of the Research Program in Structural Analysis housed at Columbia University's Center for the Social Sciences. Portions of this note were presented at the 1985 annual meetings of the American Sociological Association.

** Department of Sociology, Columbia University, New York, NY 10027, U.S.A.

[1] Burt (1984) provides a detailed discussion of the data and various issues taken into account by the GSS Board of Overseers in their deliberations over the network items.

In order to construct indices of network structure from the data however, quantitative scores will have to be assigned to the response categories. A discussion partner can be especially close to the respondent, as close as the other discussion partners, or less close. Any two discussion partners can be especially close, total strangers, or somewhere between strangers and especially close. Compared to an especially close relationship, how much weaker is a less close relationship? Compared to the relationship between total strangers, how much stronger is a relationship that is not viewed as especially close? These and related questions have to be answered in order to construct often used network indices such as the mean strength of relations in a respondent's network (density), or the connectedness of certain kinds of people in her network (centrality), or the diversity of kinds of people represented in her network (range). Assumptions have to be made about the relationship strength indicated by each response category. In this brief note, we use the idea of balance in social relations and data on the strength of relations to suggest some scale values for the GSS data.

2. Equal interval assumptions

Some reasonable scaling assumptions were presumed – based on question wording and face validity – in the network item proposal to the GSS Board of Overseers. For simplicity, we will sometimes refer to a respondent's cited discussion partners as her alters. It seems reasonable to consider "total strangers" a minimal relationship and so set such inter-alter relations to a value of 0. At the other extreme, it seems reasonable to set "especially close" relationships equal to a maximum strength relation at 1. Alters who are are neither "total strangers" nor "especially close" have a relationship of strength somewhere between 0 and 1. In the absence of further information these intermediate strength relations could be set to 0.5 assuming equal intervals to the two extremes of relationship strength. Turning to respondent–alter relations, each could be set to 1 as a sociometric citation. This certainly seems reasonable when the respondent feels "equally close" to her discussion partners. When some alters are"especially close" making the others "less close", however, the quantitative meaning of "less close" is unclear. In the absence of further information, "less close" could be set

to 0.5 making "less close" relations count half as much as "close" relations in network measures.

3. Order and balance in the data

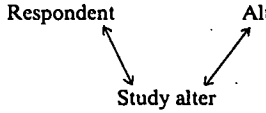
Point one, we know that there is a citation order effect in the GSS network data. The relationship between respondent and the first named discussion partner, alter one, is very strong. In fact, relationship strength indicated by closeness and contact frequency has a steep, linear decline across the first three discussion partners and a slower, but continuing, decline across the fourth and fifth named discussion partners. [2]

Point two, the idea of balance in social relations stated in its most basic form implies that two people strongly tied to one another will have similar relations to any third person. [3] *Ceteris paribus*, the strong relationship between respondent and alter one therefore implies that relations from respondent and alter one to any third person should have the same strength. A person especially close to the respondent should be especially close to alter one. A person especially close to the respondent should not be a stranger to alter one.

The above two points together mean that the response categories used to measure respondent relations with the persons named second through fifth (study alters in Figure 1) can be scaled in terms of the first alter's relations with the same people, and vice versa. An example tabulation of respondent–alter relations by alter–alter relations is presented in Figure 1. Frequencies are the number of respondent–alter dyads involving second, third, fourth or fifth cited discussion partners. For example, 154 of the people named second or later were "especially close" to the respondent and "especially close" to alter one while 31 were "especially close" to the respondent and "total strangers" to alter one.

[2] The analysis supporting this conclusion is presented in Burt (1986). The order effect exists before and after network size is held constant and continues after the contents in a relationship (kinship, homophily, co-worker, friend, etc.) are held constant.

[3] Leik and Meeker (1971: 54–73) and Burt (1982: 55–60) review the network models developing Heider's idea of cognitive balance into the sophisticated network transitivity models studied in the 1970s. Many of the key articles in this development are reprinted in Leinhardt (1977).



Tie between study alter and the respondent	Tie between study alter and alter one		
	Especially close	Acquainted	Stranger
Especially close	154 (1.32)	118 (0.99)	31 (0.76)
Equally close	798 (1.26)	515 (0.80)	222 (1.00)
Less close	201 (0.60)	431 (1.27)	153 (1.31)

Figure 1. Using alter one as a scaling criterion (multiplicative loglinear effects are presented in parentheses and frequencies exclude the 388 "less close" first alters).

4. Relations between discussion partners

Relations between discussion partners can be scaled using fixed points on respondent–alter relations as a criterion. After naming discussion partners, respondents were asked whether they felt equally close to all of the people named or closer to some than others. Those feeling closer to some than others were asked to indicate the people to whom they felt especially close. Table 1 shows how the proportion of respondent–alter relations distinguished to be "especially close" rather than "less close" shifts across levels of relationship between alters under different conditions. The invocation of balance here implies that the proportion of especially close relations with the respondent increases with the strength of the relation between study alter and alter one.

The story is told by the results in the first row of Table 1, taken from the frequencies in Figure 1. There are 355 discussion relations with people cited second or later (Figure 1 study alters) in which the discussion partner has an especially close relation with alter one and either an especially close or less close relation with the respondent (154 and 201 dyads, respectively, in the first column of the table). Of these, 43.4 percent have an especially close relation with the respondent. The percentage drops to 21.5 percent (of 549 dyads in the second column) if the discussion partner is neither especially close nor a stranger to alter one. It drops to 16.9 percent (of 184 dyads in column three) if the discussion partner is a stranger to alter one. Summarizing these results,

Table 1
Scaling relations between alters

	Relationship between alter one and study alter		
	Especially close	Acquainted	Strangers
Alter one close to respondent (1088)	0.434	0.215	0.169
All dyads (1476)	0.440	0.295	0.265
Kinship constant (1088)	0.373	0.239	0.215
Size and order constant (1088)	0.418	0.223	0.175
Suggested scaling	1.0	0.2	0.0

Note: Study alter is a discussion partner cited second or later in response to the GSS sociometric name generator. All results except the "all dyads" results are based on the exclusion of less close alter ones, as illustrated in Figure 1. Results give the probability of a study alter and respondent being "especially close" rather than "less close" at each level of relationship between the study alter and alter one as discussed in the text. The number of dyads on which results are based is given in parentheses. A study alter is kin to the respondent if he is a spouse, parent, sibling, child, or member of the extended family. Sociometric order distinguishes study alters cited second, third, fourth, and fifth. Network size refers to the number of sociometric citations a respondent made (2, 3, 4, 5, 6 or more).

the middle category of interalter relations lies about 0.2 of the distance from total strangers to people being especially close. [4] In contrast to the equal interval assumption, in other words, alter pairs who are

[4] More specifically, 0.174 equals the difference (0.215 − 0.169) divided by the difference (0.434 − 0.169). These effect-proportional scalings are conveniently expressed as a regression model (e.g. see Lyons 1971). Let Z be the dichotomous criterion variable distinguishing an especially close relation between respondent and study alter (coded 1) from a less close relation (coded 0). Let E be a dummy variable equal to 1 if the study alter has an especially close relation with alter one and 0 otherwise. Let A be a dummy variable equal to 1 if the study alter is acquainted with alter one (neither especially close nor a stranger) and 0 otherwise. Given estimates of the regression coefficients in the following equation (where R is a residual term):

$$Z = b + b_e E + b_a A + R,$$

b is the probability of an especially close relation with the respondent for a study alter who is a stranger to alter one, $b + b_a$ is the probability of such a relation for study alters merely acquainted with alter one, and $b + b_e$ is the probability of such a relation for study alters especially close to alter one. This is the model used to estimate the first two rows of Table 1. A control for kinship is added to get the estimates in the third row and controls for network size and sociometric order are added to get the estimates in the fourth row. The relative position of acquaintance between stranger and especially close has been computed with the ratio b_a/b_e.

neither especially close nor strangers are much more like strangers that especially close. They are acquainted, but not close.

The remaining rows in Table 1 show that the row one results hold across various conditions. In the second row of Table 1, acquaintance is 0.17 of the difference between especially close and stranger for all discussion partners. In other words, weakening the relationship between respondent and alter one (and so departing from ideal conditions for balance) by including dyads in which alter one is less close to the respondent yields the same results. Returning again to the discussion partners cited by respondents close to their first cited alters, acquaintance in the third row of Table 1 is 0.15 of the difference between especially close and stranger when kinship between respondent and study alter is held constant. [5] In the fourth row, acquaintance is 0.20 of the difference between especially close and stranger when network size and sociometric citation order are held constant. [6] Rounded to the

[5] The regression model in the preceding footnote controls for the tendency for strong relations with kin. A more thorough test was conducted with loglinear models of the data. In a three-way tabulation of (a) the trichotomous relation between study alter and alter one by (b) the dichotomous relation between study alter and respondent by (c) whether or not the study alter and respondent were kin, the results in Table 1 concern interactions between categories of the first two variables. Ignoring the less close alter ones as in Figure 1, the data are adequately described by the hypothesis that these interactions are independent of whether or not the respondent is related to the study alter (4.13 likelihood ratio χ^2 statistic with 2 degrees of freedom, $p \sim 0.13$). Similarly, interactions between the relation variables in Table 2 are independent of kinship between respondent and study alter (4.70 χ^2 statistic with 2 degrees of freedom, $p \sim 0.10$).

[6] The regression model in footnote 4 measures continuous size and order effects. A more thorough test was conducted with loglinear models of the data. In a three-way tabulation of (a) the trichotomous relation between study alter and alter one by (b) the dichotomous relation between study alter and respondent by (c) the order in which the study alter was cited (second, third, fourth, or fifth), the results in Table 1 concern interactions between categories of the first two variables. Ignoring the less close alter ones as in Figure 1, the data are adequately described by the hypothesis that these interactions are independent across levels of sociometric order – 7.63 χ^2 statistic with 6 degrees of freedom ($p \sim 0.27$). All χ^2 statistics reported here are likelihood ratio statistics. The hypothesis is an even more acceptable description of the data in Table 2 (χ^2 statistic of 2.76 with 6 degrees of freedom, $p \sim 0.84$). The data in a similar table where network size is the third variable (size categories of 2, 3, 4, 5, 6 or more), are also well described by the hypothesis that the interactions reflected in the Table 1 probabilities are independent of network size (4.25 χ^2 statistic with 8 degrees of freedom, $p \sim 0.83$). Similarly, interactions between the relation variables in Table 2 are independent of network size (χ^2 of 7.61 with 8 degrees of freedom, $p \sim 0.47$). Note that these results do not imply that the relation variables are independent of network size or sociometric order. In fact, there are strong order effects in the data as documented elsewhere (Burt 1986) and evident here. The hypothesis that the two relation variables in Table 1 are independent of sociometric order, for example, provides an unacceptable description of the data (38.12 χ^2 statistic with 15 degrees of freedom, $p < 0.001$). The χ^2 statistics reported above merely show that any order and size effects on levels of relationship with respondent and alter one are independent of the interactions between the levels of relationship in Figure 1.

nearest decimal, these results suggest the scalings indicated at the bottom of the table – 1 for especially close discussion partners, 0.2 for acquaintance, and 0 for total strangers.

5. Relations with discussion partners

Relations between respondent and discussion partners can be scaled using fixed points on alter–alter relations as a criterion. Focusing on the relatively clear extremes in relations between alters to define a scaling criterion, especially close versus strangers, Table 2 shows how the proportion of especially close relations with alter one shifts across levels of relationship with the respondent. The invocation of balance here implies that the proportion of especially close relations with alter one increases with the strength of relationship between respondent and discussion partner.

Table 2
Scaling relations between respondent and alter

	Respondent to study alter relationship		
	Especially close	Equally close	Less close
Alter one close to respondent (1559)	0.832	0.782	0.568
All dyads (1746)	0.722	0.782	0.544
Kinship constant (1559)	0.798	0.769	0.624
Size and order constant (1559)	0.831	0.780	0.577
Suggested scaling	1.0	1.0	0.7

Note: Study alter is a discussion partner cited second or later in response to the GSS sociometric name generator. All results except the "all dyads" results are based on the exclusion of less close alter ones, as illustrated in Figure 1. Results give the probability of alter one and a study alter being "especially close" rather than "strangers" at each level of relationship between the respondent and study as discussed in the text. The number of dyads on which results are based is given in parentheses. A study alter is kin to the respondent if he is spouse, parent, sibling, child, or member of the extended family. Sociometric order distinguishes study alters cited second, third, fourth, and fifth. Network size refers to the number of sociometric citations a respondent made (2, 3, 4, 5, 6 or more).

The basic results are given in the first row of the table, taken from the frequencies in Figure 1. There are 185 discussion relations with people cited second or later (Figure 1 study alters) in which the discussion partner has an especially close relation with the respondent and either an especially close or stranger relation with alter one (154 and 31 dyads respectively in the first row of the table in Figure 1). Of these, 83.2 percent have an especially close relation with alter one. The percentage drops to 78.2 percent (of 1020 dyads in the second row) if the discussion partner and all other cited discussion partners are equally close to the respondent. It drops to 56.8 percent (of 354 dyads in the third row) if the discussion partner is less close to the respondent. Summarizing these results, less close relations with the respondent are about 0.7 of the strength of especially close relations. [7]

As in Table 1, the remaining rows in Table 2 merely show that this result holds across various conditions. In the second row, less close relations are 0.75 the strength of especially close relations for all discussion partners. In the third row, less close is 0.78 of especially close when kinship is held constant. [8] In the fourth row, less close is 0.69 of especially close when network size and sociometric order are held constant. [9] Rounding to the nearest decimal, these results suggest

[7] More specifically, 0.682 equals 0.568 divided by 0.832. As with Table 1 (see footnote 4), these effect-proportional scalings are conveniently expressed as a regression model. Let Z be the dichotomous criterion variable distinguishing study alters with an especially close relation to alter one (coded 1) from those who are strangers to alter one (coded 0). Let E be a dummy variable equal to 1 if the study alter has an especially close relation with the respondent and 0 otherwise. Let S be a dummy variable equal to 1 if the study alter has the same strength relation with the respondent as all other cited discussion partners and 0 otherwise. Given estimates of the regression coefficients in the following equation (where R is a residual term):

$$Z = b + b_e E + b_s S + R,$$

b is the probability of an especially close relation with alter one if the discussion partner is less close to the respondent, $b + b_s$ is the probability of an especially close relation with alter one if the discussion partner is as close to the respondent as all other alters, and $b + b_e$ is the probability of an especially close relation with alter one if the discussion partner is especially close to the respondent. This is the model used to estimate the first two rows in Table 2. A control for kinship is added to get the estimates for the third row and controls for network size and sociometric order are added to get the estimates for the fourth row. The relative magnitude of "less close" to "especially close" relation has been computed with the ratio $b/(b + b_e)$.

[8] See footnote 5 for further details on the stability of the scaling with kinship held constant.

[9] See footnote 6 for further details on the stability of the scaling with network size and sociometric order held constant.

the scalings indicated at the bottom of the table – 1 for especially close and equally close discussion relations and 0.7 for less close relations.

6. Conclusion

In looking at the internal consistency of the GSS network data on relationship strength, we find evidence of unequal intervals between the response categories. Pairs of discussion partners perceived as neither strangers nor especially close are acquainted but far from especially close. Their relation has a strength that is about 0.2 of the distance from total strangers to people who are especially close. Discussion partners distinguished as less close than others to the respondent are about 0.7 as close to the respondent as the people with whom she feels especially close.

To illustrate the variable impressions one can obtain from the data under quite reasonable scaling assumptions, summary data on network

Table 3
Density under alternative scaling assumptions

	Mean	Standard deviation	95 percent confidence interval
Density of strangers between alters	0.189	0.279	0.17–0.21
Density of especially close relations between alters	0.413	0.378	0.40–0.43
Density of scaled relations between alters	0.492	0.331	0.47–0.51
Density of equal interval relations between alters	0.612	0.281	0.60–0.63
Density of scaled relations including respondent	0.721	0.196	0.71–0.73

Note: Density is computed as the average strength of relations in a network. Results are based on the 1161 respondents citing two or more discussion partners. The density of stranger and especially close relations measures the proportion of alter pairs who are respectively strangers or especially close. Scale values for the sociometric response categories are given in Tables 1 and 2. Equal interval relations are measured 1 if two alters are "especially close", 0 if they are "strangers", and 0.5 if their relationship lies somewhere between stranger and especially close.

density – the average strength of relations in a respondent's network are presented in Table 3. The density of structural holes (proportion stranger relations) and the density of especially close relations are based only on the extreme categories of relations between discussion partners and so are unaffected by the alternative scaling assumptions we have discussed. The average network contains less than half of the "especially close" relations possible (41%). This is comparable to the 44 percent density of binary "knowing well" relations among core alters that Fischer (1982: 145) reports in his Northern California Communities Study. At the other extreme, about a fifth of the possible relations among discussion partners are missing (19% density of stranger relations). It is the network measures that take advantage of data on all three categories of relationship that are affected by the discussed scaling assumptions. In the remaining rows of Table 3, the average strength of relationship within the average network ranges from 49 percent of the maximum strength possible with a 33 percent standard deviation up to an average strength 72 percent of the maximum possible and a much smaller, 20 percent standard deviation. Any of the rows in Table 3 is a legitimate report of density in American discussion networks. The moral is that quantitative values for the GSS network response categories should be selected carefully when operationalizing network concepts. The scale values in Tables 1 and 2 seem to better represent the formal data than the equal interval scaling initially proposed to the GSS Board of Overseers.

References

Burt, Ronald S.
1982 *Toward a Structural Theory of Action: Network Models of Social Structure, Perception and Action*. New York: Academic Press.
1984 "Network items and the General Social Survey". *Social Networks 6*: 293–339.
1986 "A note on sociometric order in the General Social Survey network data". *Social Networks 8*: 149–174..
Fischer, Claude S.
1982 *To Dwell Among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.
Leik, Robert K. and Barbara F. Meeker
1971 *Mathematical Sociology*. Englewood Cliffs, NJ: Prentice-Hall.
Leinhardt, Samuel (ed.)
1977 *Social Networks: A Developing Paradigm*. New York: Academic Press.
Lyons, Morgan
1971 "Techniques for using ordinal measures in regression and path analysis". In H.L. Costner (ed.) *Sociological Methodology 1971*. San Francisco: Jossey-Bass.