

The Meaning of Knowing as a Network Tie*

Christopher McCarty

Introduction

In 1969, Mitchell focused our attention on the multiplexity of network ties. Each tie between two network alters, he said, might comprise several "strands." Two people might be connected by kinship, or debt, or affect, or any number of other links.

The importance of multiplexity was understood immediately by researchers in the field, as was the fact that some links were clearly stronger than others. In his study of scientists, for example, Friedkin (1980) differentiated between ties based on people saying that they had read one another's work and their actually citing one another's work. Erikson (1981) distinguished between college students being able to recognize one another (from pictures) and their claiming to like or interact with one another. Mitchell (1987) distinguished tie strength between homeless women based on things like whether they gave one another practical support. Wellman (1979) describes many different strands in the ties of East Yorkers.

In fact, recognizing the importance of multiplexity in network ties, researchers have developed different, quite specific, network generators. Burt (1984) asked people to focus on the alters with whom they "discuss important matters." McCallister and Fischer (1978) focused on ties defined by acts of social support ("Who could you leave your house with if you were going out of town for a few days?"). And so on.

* Christopher McCarty. "The Meaning of Knowing as a Network Tie." Connections 18:20-31 (1996).

The most exhaustive work on the multiplexity of network ties was done by Burt (1983). Through a secondary analysis of data collected by McCallister and Fischer, Burt uncovered five dimensions of social relationships: friendship, acquaintance, work, kinship and intimacy. He further suggested that these dimensions were not mutually exclusive, and were frequently associated with the concepts of “confusion between contents,” “substitutability of contents” and “content ambiguity.” These concepts categorized the types of errors which network analysts might be subject to in applying single word cues to elicit network ties among respondents.

“...it seems likely that there would be subcultures within which people understood relations in a similar way, but across which a single word or phrase could refer to quite different qualities of interaction.” (pg. 49)

In the same volume, McCallister and Fischer (ibid.) critiqued the use of general terms such as “best friend” and “close”, pointing to the potential for respondents to interpret them in varying ways. Still, most ego-centered research is based on asking people some form of the question "Who do you know ...?" where the blank is filled in with things like "who could lend you a hundred dollars" or (in studies of very old people) "who could take you shopping?"

In this article, I raise some questions about the use of the word "know." What exactly do we get when we generate ties in this way? Is knowing itself multiplex? And what is the measurable error which we can expect when knowing is used to define network ties?

In what follows, I summarize three studies that shed some light on this question. The first is an in-depth descriptive study that focused on defining "differences in knowing" of people. In the other two studies, these differences were tested with representative samples.

Study 1.- Method

I started with the notion that the concept of "knowing" was interpreted differently by different people. My goal was to elicit these differences with a minimum of experiment effect. To do this, I asked 47 respondents to free list 60 people they knew, given a loose definition of knowing. Respondents only had to "know" their listed alters by sight or name. Alters had to be alive and respondents had to be able to contact their alters. Alters had to know respondents by sight or by name (reciprocal knowing). After free-listing the 60 alters, respondents rated each alter, on a scale of 1 to 5, on how well they knew each, where a 1 was the lowest level of knowing and a 5 the most. In all, the 47 respondents selected and rated 2,820 alters.

Following the rating procedure, respondents were asked to think of examples of the types of people to whom they assigned the various levels of "knowing," and to describe in general terms what type of person corresponded to a level. My assumption was that, having used these rating levels 60 times, the descriptions would divulge understanding of knowing. I hoped to learn, for these 47 people, the criteria they used in assigning a higher or lower level of knowing to their network alters.

Study 1.- Results

Reading through the 235 responses (one description for each of 5 levels for all 47 respondents), I decided on 11 dimensions that were represented. These dimensions were defined subjectively, and to some extent, were influenced by the knowledge of the results from the studies listed above. Although another coder might come up with more or less dimensions, this is not critical to the purpose of this study. Assuming that another coder would perceive most of these dimensions, an assumption which is supported by earlier research, then these dimensions serve the purpose of showing how they vary by levels of “knowing,” the single word tie measure.

On average, respondents used 4.3 dimensions (SD 1.6) in their descriptions of the five levels, frequently expressing more than one dimension per level. All of the analyses are based on the presence or absence of dimensions for each respondent, and not on the number of times dimensions were used.

Table 1 shows the number of respondents who used each of the 11 dimensions and the distribution of the dimension over the five knowing levels¹. For example, the dimension of “friendship” was mentioned by 64 percent of the respondents over all knowing levels, and by 45 percent of the respondents on knowing level 5. This means that 45 percent of the respondents somehow thought of the idea of friendship when assigning knowing level 5 to alters². Friendship increases with knowing level, so that those whom we know at low knowing levels are frequently not

¹ Compare these 11 dimensions to those found by Burt. The fact that these dimensions were developed from respondent text independent of Burt’s dimensions lends support to their validity.

²Note that in all cases the frequency of dimension use counts the dimension whether it is used positively or negatively. Thus, saying level 1 is for friends, or level 1 is for people who are not really friends will both be counted as use of the “friend” dimension.

friends, and those at high knowing levels are. Unfortunately, friendship is conceptually nondescriptive³. If friendship had been coded into the categories of friend/not friend, there might be more differentiation, but in this study most of those listed were considered friends.

In contrast to friendship, the dimension of “acquaintance” has a negative association with knowing level. Although it provides little concrete information about the relation, other than a suggestion of little knowing, it is more informative than the dimension of friendship.

Frequency of contact is a concept which applies across all knowing levels except the highest. As might be expected, the higher the knowing level, the higher the percentage of respondents who used this dimension. This holds until level 5, where only six percent of the respondents used it. These data suggest that, for these respondents, defining knowing to be people whom you see the most could exclude alters whom the respondent knows the most.

An explanation for the fact that level 5 does not require frequency of contact emerges from examination of the family category. It tends to be associated with a high knowing level used by over half the respondents. Many people considered 5 to be reserved for family and close relationships, making frequency of contact less important in the evaluation of knowing strength. This may not be limited to family, however. Frequency of contact may be unimportant for most level 5 alters, regardless of the character of the network tie.

³In Burt’s examination of content ambiguity, he found the dimension of friendship exhibited the highest level of ambiguity of all five dimensions. This would explain the consistent use of this concept across all levels of knowing.

Another category of interest, "knowing personal data about the alter", also increased with knowing level. For some respondents, "the more you know about" an alter, or conversely "the more they know about you", the higher the level of knowing. Acquaintances were frequently characterized as those about whom respondents knew little (particularly personal information). Some respondents specifically mentioned knowing factual data about an alter, such as how many children alters have or where alters lived.

Although the dimension of "closeness" was mentioned by half of the respondents, analysis showed that respondents used the idea of closeness synonymously with knowing (suggestive of Burt's concept of the substitutability of content). When asked to rank on a scale of 1 to 5 how close they were to an alter, over 90 percent of the 2,820 rankings were within 1 point of the knowing rating. Further, there were no cases where knowing was high and closeness low, or vice versa.

Table 2 presents those dimensions that go together. Friendship as a criterion of knowing is mentioned along with closeness, family and acquaintance. None of these present any surprises, with the exception of a relatively weak negative association between "frequency of contact" and "knowing through someone". I call these ties "vias." When respondents mentioned knowing someone as a via, they tended not to mention frequency of contact as a dimension.

A principal components factor analysis on the frequencies of these dimensions yielded two factors that explained 37% of the variance in the 11 variables tested. The first factor loads the dimensions of family, closeness and friendship high, and the two knowing-data dimensions low.

Rotating the factors either orthogonally or obliquely causes the lower loadings, which were not demonstrably strong in the original solution, to vanish. This leaves the dimensions of family, closeness and friendship defining the first factor, which tends to be associated with high levels of knowing.

The second factor loads the via dimension against frequency of contact and duration of tie. This is sensible since the via relation is a consequence of another person. Thus, frequency of contact and the duration of the via tie also depend on the intermediary alter. The remaining factors accounted for little variance and loaded low.

Study 2.- Method

In the next study, I adapted the free list method to a telephone survey in which respondents were selected randomly from all Florida households that have a telephone. Interviewers read a randomized list of male first names to respondents. When a respondent knew a person who had one of the first names, he or she told the interviewer to stop. Respondents were then asked to say, on a scale of 1 to 3, how and how well they knew the choice. Interviewers then read a list of female first names following the same procedure. Interviewers alternated male and female lists until eight alters, four males and four females, had been selected.

My assumption here was that first names are more randomly distributed in a person's network than are names in a respondent's free list. Excluding names associated with particular ethnic groups biases against the inclusion of family members for respondents who belong to these ethnic groups. However, people named John are no more likely to be co-workers than people named Edward. My

assumption, then, was that the eight alters approximate a random selection of eight people from a respondent's network.

This method of alter selection, while arguably more random than other methods, is not perfect. There is a strong gender bias based on the fact that there is more variety to female names than to male names. Second, care must be taken to ensure that certain ethnic groups are represented, particularly Blacks and Hispanics. Finally, since it is more efficient to use popular names, the potential to use a name which applies to more than one alter is high. In such a case I assume that the respondent will use the first to come to mind, or the alter who is closest to them. Although this presents a potential bias towards closer alters, analysis of data generated by this method suggests there is very little "order effect" between alters which is more common with other methods (Brewer, 1995).

As in the experiment using 60 alters, the 233 respondents were asked to describe how they assigned knowing levels 1 to 3 after they had done so for the eight alters. Again, in the same study, the data were coded according to the presence or absence of certain dimensions.

Study 2.- Results

Telephone survey results are presented in Table 3. The big difference between Table 1 and Table 3 is the presence of "situation specific" and "emotional aspect" in the way respondents described their use of the three levels. Over a fourth (28%) of the respondents used a situation (such as "going to lunch" or "borrowing money"), and a fifth (22%) used what I categorized as emotional concepts (such as "casual relations" or "presence of feeling" or "likable"), to characterize relations.

In general the distribution of the dimensions across knowing levels in Table 3 resembles the distribution in Table 1. Frequency of contact is used across all levels and falls off in the highest knowing category. Family is generally a high knowing-level category. Closeness, however, is mentioned in all levels. For low knowing levels, closeness was mentioned in the negative; that is a 1 is used for people to whom respondents are not close. Acquaintance is consistently a low knowing-level dimension, while knowing data about alters increases with knowing level.

Summary of Studies 1 and 2

It is apparent that respondents have several ways in which they interpret knowing, levels of the three most dominant being frequency of contact, knowledge of personal data, and whether or not the alter is family. Remember that these dimensions arose from an already constrained definition of knowing--a definition far more limited than is typically used in network generators. Thus, knowing itself is multiplex. This means that respondents may interpret knowing differently from one another. And this introduces the likelihood of variability in the results due to the application of the construct, rather than to variables that researchers are investigating. How severe are the potential errors? To answer this question, a third study was done on a much larger sample of respondents.

Study 3.- Method

A list of fifty first names was selected from the 1993 University of Florida registrar's records. Twenty five uniquely male and twenty five uniquely female names were picked and an attempt was made to use names that are common among both black and white students and that do not demonstrate

a blatant age bias. Despite all precautions, only 37% of the 136 Black respondents were able to connect with 14 alters from the list of 50 first names, compared to 52% of the 1,313 White respondents. The final list was cross-checked with similar lists generated by Pluckett (1975).

Unlike Study 2, there was one list that alternated male and female names. This made it possible for a respondent to choose alters disproportionately by gender, although there was probably a bias against female alters⁴. On balance, I believe this list to be less biased than most network generators in eliciting a random grab of alters from an ego-centered network. Interviewers read the list of fifty names stopping when the respondent recognized a name. At this point a series of questions were asked about the alter and the respondent's relationship to the alter. A total of 14 alters were selected, and 747 respondents were able to make it through the list and generate all 14 alters. These data were summarized to provide characteristics of each network, such as the proportion of a network that were family members, or the average level of knowing.

Study 3. - Results

Studies 1 and 2 defined the dimensions that informants used when assigning a five-point knowing level to an alter. From the distribution of these dimensions across knowing levels it appears

⁴ As I and others have noted, there is more variety to female names than male names. From the registrar's list, 42% of all males are accounted for by the 21 most frequently occurring male names compared to 25% for the 21 most frequently occurring female names. Thus, a list with equal numbers of both is probably biased against selecting female alters in the proportion they actually exist in the network. The proportions such a list should have would vary by respondent and are unknown.

that, as knowing level increases, frequency of contact, duration of tie, the amount of information known about an alter and whether or not the alter is a family member become increasingly important.

However, while these studies demonstrate the multiplexity of knowing as a tie, they do not tell us about the direction of the relation between dimensions and knowing level. We assume that high knowing levels will be associated with frequent contact, lengthy relationships, high proportions of family alters and a high level of information known about the alter. It is this assumption, and the potential of cases that run counter to them, that make multiplexity important. After all, if these dimensions vary directly with level of knowing, then that level is a proxy for all of the dimensions. Studies that use this single concept to represent a network tie gain validity. Study 3 provides the opportunity to test whether these trends exist and whether anomalous cases are common enough to warrant caution in using knowing by itself to define a tie.

Turning first to frequency of contact, Figure 1 shows a clear tendency for average knowing level to decrease as frequency of contact decreases. The mean differences of knowing level between categories are significant, with the exception of those included between twice a week and once a month.

Cases running counter to this trend would be high knowing and low frequency, or low knowing and high frequency. On average, four percent of respondents' alters are those whom they see no more than every six months, but know at level five. Another four percent on average are those whom

respondents see at least twice a week but know at level one. So, a total of eight percent of the alters run counter to the assumed relationship between knowing and frequency of contact.

Figure 2 also shows the expected relation between duration of tie and knowing level; as the duration of the tie increases, so does the average knowing level. In contrast to frequency of contact, duration exhibits only one type of anomalous case. On average, nine percent of respondents' alters were those whom they knew for at least two years, but to whom they assigned a knowing level of one. There were virtually no cases of informants assigning an alter a knowing level five when they had known them for less than three months.

The amount of information known about an alter is correlated strongly with knowing level (Pearson $r=.92$). (It remains unclear whether this is because information known and knowing level are actually the same, or if this is an experiment effect due to the uniqueness of these two variables having a number assigned to them with no definition of the categories, that is, they are both constructs.)

Still, an average of 16 percent of the alters were knowing level one, but were assigned an information level of five. In contrast to the dimension of closeness, many alters are those whom respondents don't know well, but about whom they claim to know a lot about. The opposite case, claiming to know someone well (knowing level five) but knowing little about them (information level one), is very rare.

The final dimension is that of family. In Study 3 the distinction was made between blood relations and marital relations. Figure 3 presents the proportions of family of each type that account for

each of the five knowing levels. Thus, 46 percent of knowing level five assignments were blood relations, while 11 percent were marital relations. This figure shows that as knowing level increases, the proportion of family ties of both types increase. Interestingly, they are virtually identical for the first four levels, while blood relations are four times more prevalent in level five knowing assignments than are marital relations. The potential anomalous cases of low-knowing blood or marital relations combined represent about 1.5 percent of alters on average, or less than one percent each.

Discussion

From these three studies we see that the word "knowing" is understood by informants in different ways. We cannot assume that asking informants to assign tie strength based on this single word will assign values equally along all the dimensions informants think about when applying it to their alters.

The dimensions that appear to be important in making these assignments are friendship, closeness, frequency of contact, duration of tie, how much is known about an alter, whether the alter is family, and acquaintanceship. Friendship and acquaintanceship are not informative concepts as they tell us little about the nature of a relation. Closeness is by all accounts the same concept as knowing; there are virtually no cases of high knowing and low closeness or vice versa. Of the remaining four dimensions of knowing, all exhibit an overall trend with relation to knowing that is predictable. Information known, duration of tie, frequency of contact and proportions of family all increase with the level of knowing.

However, several cases run counter to these trends. These are high knowing/low frequency and low knowing/high frequency alters (4 percent of alters each, on average), low knowing/high duration (9 percent) and low knowing/high information known (16 percent). Consequently, those studies that measure ties and tie strength using the single word "knowing" may misinterpret the composition of the tie. Worse, assuming that defining a strong tie as high frequency, high duration or high information known by themselves could easily result in a low knowing tie.

How might such misinterpretations affect the results of research which uses a single word network generator? Consider the percentages listed above which represent cases where the level of a particular dimension does not have the expected relation with the level of the single word tie measure "knowing." Most of these discrepant cases do not pose a serious problem. Yet 16 percent of the time it would be erroneous to assume that an alter who was selected because the respondent "knew" them well was also someone the respondent knew a lot about.

One area of research that this might significantly effect is the use of information known by respondents about alters to estimate hard-to-count populations (see Killworth et al., 1995; Laumann et al., 1993). In these cases respondents are asked to list alters they "know" who belong to various subpopulations, such as diabetics, or victims of AIDS. The results of this study suggest that the ability of the respondent to know whether an alter is a member of any one subpopulation may be seriously compromised as much as 16 percent of the time.

These results lend support to the conclusions reached by Burt in 1983. Whereas Burt's data were based upon a limited set of five alters, these studies extend that to a free list of 60 alters, and a "random" draw of 14 alters from the networks of 747 respondents. Similar to Burt's findings, these relations are clearly multiplex. Further, respondents accept the concept of "knowing" to define their ties and interpret the tie in different ways.

In the end, we are probably safer to measure network ties using several questions with defined categories that individually address the various dimensions of ties. While the focus of this study was the word "knowing", there is no reason to believe other single word tie definitions, such as "closeness", are any better proxies for all of these dimensions. Given that all subsequent results are based on the way ties are measured, a standard set of questions should be developed that network researchers will use to define ties. This will make our results comparable and will remove the effect of different measures of ties and tie strength.

References Cited

Burt, Ron

1983 "Distinguishing Relational Contents" in Applied Network Analysis, Ron Burt and Michael Minor (eds.), Beverly Hills: Sage Publications.

1984 "Network Items and the General Social Survey" Social Networks 6:293-339.

Brewer, Devon

1995 "No associative biases in the first name cued recall procedure for eliciting personal networks." Presented at the meetings of the International Network for Social Network Analysis in London, England, July 1995.

Erickson, Bonnie H.

1981 "Network Sampling in Practice: Some Second Steps." Social Networks 3:127-136.

Fischer, Claude S.

1982 "What do we mean by 'friend'? An inductive study." *Social Networks* 3:287-306.

Friedkin, Noah

1980 "A Test of the Structural Features of Granovetter's Strength of Weak Ties Theory." *Social Networks* 2:411-422.

Killworth, Peter, Christopher McCarty, Eugene C. Johnsen, Gene Ann Shelley, H. Russell Bernard

1995 "A Social Network Approach to Estimating Seroprevalence in the United States." Submitted to the *Journal of the American Statistical Association*.

Laumann, Edward O., John H. Gagnon, Stuart Michaels, Robert T. Michael, L. Philip Schumm

1993 "Monitoring AIDS and Other Rare Population Events: A Network Approach." *Journal of Health and Social Behavior* 34:7-22.

McCallister, Lynne and Claude S. Fischer

1978 "A Procedure for Surveying Social Networks" *Sociological Methods and Research*, 7.

Mitchell, J. Clyde

1969 "The Concept and Use of Social Networks" in *Social Networks in Urban Situations*, J. Clyde Mitchell (ed.), Manchester, Manchester University Press.

1987 "The Components of Strong Ties Among Homeless Women." *Social Networks* 9:37-47.

Puckett, Newbell Niles

1975 *Black Names in America: Origins and Usage*. Boston, G.K. Hall.

Wellman, Barry

1979 "The Community Question: The Intimate Networks of East Yorkers." *American Journal of Sociology* 84:1201-1231.

Table 1. Distribution of knowing dimensions across respondents.

Dimension	Percent who used dimension					
	Overall	Level 1	Level 2	Level 3	Level 4	Level 5
Friendship	64	9	15	32	43	45
Frequency of contact	57	23	34	36	21	6
Know personal data about alter	55	9	11	13	36	57
Family/Relative	55	2	0	13	15	45
Closeness	51	0	11	9	21	34
Acquaintance	40	26	15	4	0	0
Duration of tie	30	11	4	15	11	17
Know factual data about alter	30	13	11	9	4	9
Work, Business, School	28	4	11	9	6	2
Via (Know through someone)	13	4	4	2	2	0
Depth of discussions	9	13	15	20	15	17

Table 2. Significant correlations (Pearson R) of categories.

Cooccurrence of	With	Correlation	Prob > R
Friendship	Closeness	0.5	0.01
Friendship	Family/Relative	0.48	0.01
Family/Relative	Closeness	0.4	0.01
Friendship	Acquaintance	0.34	0.02
Know Through Someone	Frequency of Contact	-0.32	0.03
Know Personal Data About Alter	Know Factual Data About Alter	0.31	0.04

Table 3. Percent of respondents who mentioned category, by three levels and over all levels.

Dimension	Percent who used dimension			
	Overall	Level 1	Level 2	Level 3
Frequency of Contact	38	14	29	16
Friendship	29	2	11	22
Closeness	28	10	8	9
Situation Specific	28	13	16	7
Family/Relative	24	4	5	16
Emotional Aspect	22	19	9	5
Acquaintance	21	19	3	0
Know Data About Alter	20	0	2	23

Figure 1. Average Knowing Level by Frequency of Contact

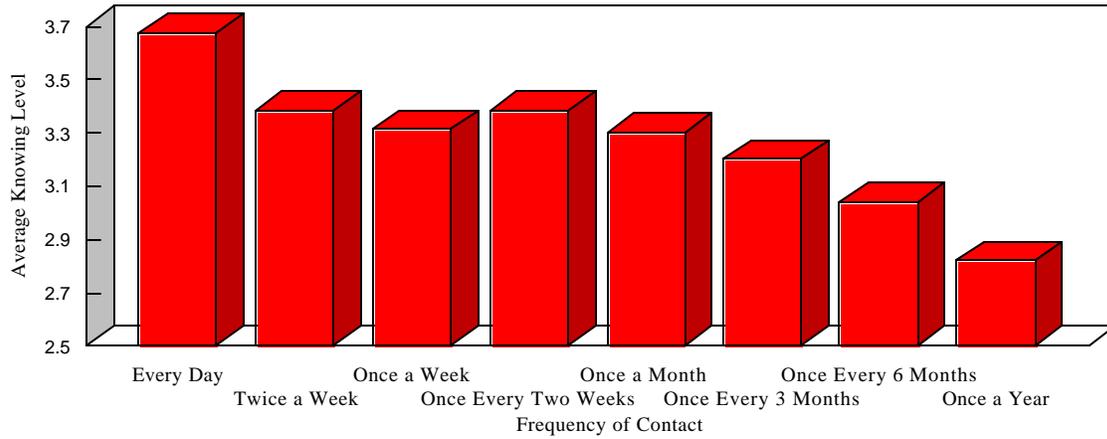


Figure 2. Average Knowing Level by Duration of Tie

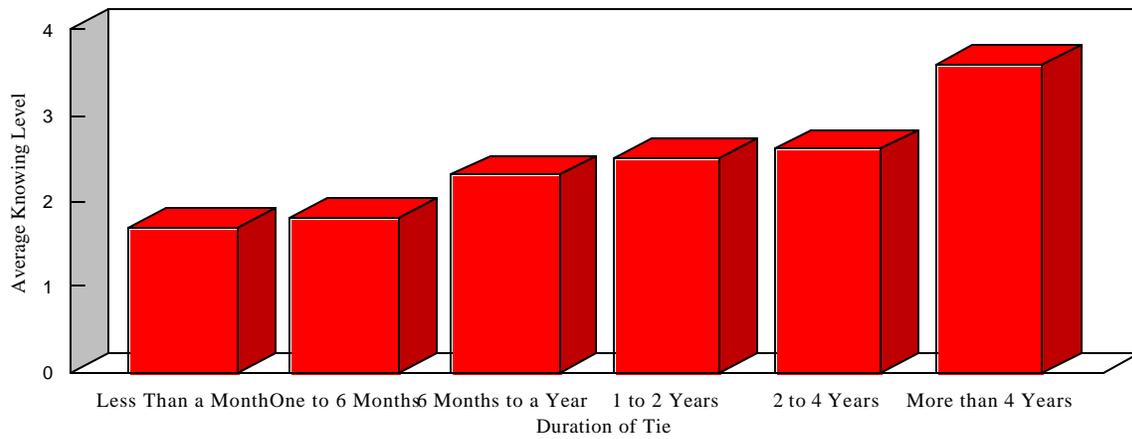


Figure 3. Proportion of Family by Knowing Level

