**An Experimental Study of the Small World Problem**

Jeffrey Travers, Stanley Milgram

*Sociometry*, Volume 32, Issue 4 (Dec., 1969), 425-443.

---



---

# An Experimental Study of the Small World Problem*

JEFFREY TRAVERS

Harvard University

AND

STANLEY MILGRAM

The City University of New York

*Arbitrarily selected individuals (N=296) in Nebraska and Boston are asked to generate acquaintance chains to a target person in Massachusetts, employing "the small world method" (Milgram, 1967). Sixty-four chains reach the target person. Within this group the mean number of intermediaries between starters and targets is 5.2. Boston starting chains reach the target person with fewer intermediaries than those starting in Nebraska; subpopulations in the Nebraska group do not differ among themselves. The funneling of chains through sociometric "stars" is noted, with 48 per cent of the chains passing through three persons before reaching the target. Applications of the method to studies of large scale social structure are discussed.*

The simplest way of formulating the small world problem is "what is the probability that any two people, selected arbitrarily from a large population, such as that of the United States, will know each other?" A more interesting formulation, however, takes account of the fact that, while persons *a* and *z* may not know each other directly, they may share one or more mutual acquaintances; that is, there may exist a set of individuals, *B,* (consisting of individuals $b_1$, $b_2$ . . . $b_n$) who know both *a* and *z* and thus link them to one another. More generally, *a* and *z* may be connected not by any single common acquaintance, but by a series of such intermediaries, *a-b-c-* . . . *-y-z*; i.e., *a* knows *b* (and no one else in the chain); *b* knows *a* and in addition knows *c, c* in turn knows *d,* etc.

To elaborate the problem somewhat further, let us represent the popula-

* The study was carried out while both authors were at Harvard University, and was financed by grants from the Milton Fund and from the Harvard Laboratory of Social Relations. Mr. Joseph Gerver provided invaluable assistance in summarizing and criticizing the mathematical work discussed in this paper.

tion of the United States by a partially connected set of points. Let each point represent a person, and let a line connecting two points signify that the two individuals know each other. (Knowing is here assumed to be symmetric: if *a* knows *b* then *b* knows *a*. Substantively, "knowing" is used to denote a mutual relationship; other senses of the verb, e.g. knowing about a famous person, are excluded.) The structure takes the form of a cluster of roughly 200 million points with a complex web of connections among them. The acquaintance chains described above appear as pathways along connected line segments. Unless some portion of the population is totally isolated from the rest, such that no one in that subgroup knows anyone outside it, there must be at least one chain connecting any two people in the population. In general there will be many such pathways, of various lengths, between any two individuals.

In view of such a structure, one way of refining our statement of the small world problem is the following: given two individuals selected randomly from the population, what is the probability that the minimum number of intermediaries required to link them is 0, 1, 2, . . . *k*? (Alternatively, one might ask not about the minimum chains between pairs of people, but mean chain lengths, median chain lengths, etc.)

Perhaps the most direct way of attacking the small world problem is to trace a number of real acquaintance chains in a large population. This is the technique of the study reported in this paper. The phrase "small world" suggests that social networks are in some sense tightly woven, full of unexpected strands linking individuals seemingly far removed from one another in physical or social space. The principal question of the present investigation was whether such interconnectedness could be demonstrated experimentally.

The only example of mathematical treatment dealing directly with the small world problem is the model provided by Ithiel Pool and Manfred Kochen (unpublished manuscript). Pool and Kochen assume a population of $N$ individuals, each of whom knows, on the average, $n$ others in the population. They attempt to calculate $P_k$, the probability that two persons chosen randomly from the group can be linked by a chain of $k$ intermediaries. Their basic model takes the form of a "tree" or geometric progression. Using an estimate of average acquaintance volume provided by Gurevitch (1961), they deduce that two intermediaries will be required to link typical pairs of individuals in a population of 200 million. Their model does not take account of social structure. Instead of allowing acquaintance nets to define the boundaries of functioning social groups, Pool and Kochen must, for the purposes of their model, conceive of society as being partitioned into a number of hypothetical groups, each with identical populations. They are then able

to devise a way to predict chain lengths within and between such hypothesized groups.

In an empirical study related to the small world problem Rapoport and Horvath (1961) examined sociometric nets in a junior high school of 861 students. The authors asked students to name in order their eight best friends within the school. They then traced the acquaintance chains created by the students' choices. Rapoport was interested in connectivity, i.e. the fraction of the total population that would be contacted by tracing friendship choices from an arbitrary starting population of nine individuals. Rapoport and his associates (Rapoport and Horvath, 1961; Foster et al., 1963; Rapoport, 1953; 1963) have developed a mathematical model to describe this tracing procedure. The model takes as a point of departure random nets constructed in the following manner: a small number of points is chosen from a larger population and a fixed number of "axones" is extended from each of these points to a set of target points chosen at random from the population. The same fixed number of axones is then extended from each of the target points to a set of second generation target points, and the process is repeated indefinitely. A target point is said to be of the $t$th remove if it is of the $t$th generation and no lower generation. Rapoport then suggests a formula for calculating the fraction, $P_t$, of the population points which are targets of the $t$th remove. He is also able to extend the formula to nonrandom nets, such as those created in the Rapoport and Horvath empirical study, by introducing a number of "biases" into the random net model. Rapoport shows that two parameters, obtainable from the data, are sufficient to produce a close fit between the predictions of the model and the empirical outcome of the trace procedure.[1]

Rapoport's model was designed to describe a trace procedure quite different from the one employed in the present study; however, it has some relation to the small world problem. If we set the number of axones traced from a given individual equal to the total number of acquaintances of an average person, the Rapoport model predicts the total fraction of the population potentially traceable at each remove from the start, serving precisely the aims of the model of Pool and Kochen. (It should, however, be noted that Rapoport's model deals with asymmetric nets, and it would be difficult to modify the model to deal with general symmetric nets, which characterize the small world phenomenon.)

Despite the goodness of fit between Rapoport's model and the data from

[1] There is additional empirical evidence (Fararo and Sunshine, 1964) and theoretical support (Abelson, 1967) for the assumption that two parameters are sufficient to describe the Rapoport tracing procedure, i.e. that more complex biases have minimal effects on connectivity in friendship nets.

two large sociograms, there are unsolved problems in the model, as Rapoport himself and others (Fararo and Sunshine, 1964) have pointed out. The Pool-Kochen model involves assumptions difficult for an empirically oriented social scientist to accept, such as the assumption that society may be partitioned into a set of groups alike in size and in internal and external connectedness. In the absence of empirical data, it is difficult to know which simplifying assumptions are likely to be fruitful. On the other hand, with regard to the empirical study of Rapoport and Horvath, the fact that the total population employed was small, well-defined, and homogeneous leaves open many questions about the nature of acquaintance nets in the larger society.[2] An empirical study of American society as a whole may well uncover phenomena of interest both in their own right and as constraints on the nature of any correct mathematical model of the structure of large-scale acquaintanceship nets.

## PROCEDURE

This paper follows the procedure for tracing acquaintance chains devised and first tested by Milgram (1967). The present paper introduces an experimental variation in this procedure, by varying "starting populations"; it also constitutes a first technical report on the small world method.

The procedure may be summarized as follows: an arbitrary "target person" and a group of "starting persons" were selected, and an attempt was made to generate an acquaintance chain from each starter to the target. Each starter was provided with a document and asked to begin moving it by mail toward the target The document described the study, named the target, and asked the recipient to become a participant by sending the document on. It was stipulated that the document could be sent only to a first-name acquaintance of the sender. The sender was urged to choose the recipient in such a way as to advance the progress of the document toward the target; several items of information about the target were provided to guide each new sender in his choice of recipient. Thus, each document made its way along an acquaintance chain of indefinite length, a chain which would end only when it reached the target or when someone along the way declined to participate. Certain basic information, such as age, sex and occupation, was collected for each participant.

[2] In addition to the Pool-Kochen and Rapoport work, there are numerous other studies of social network phenomena tangentially related to the small-world problem. Two well-known examples are Bailey's *The Mathematical Theory of Epidemics* and Coleman, Katz and Menzel's *Medical Innovation.* Bailey's work deals with diffusion from a structured source, rather than with convergence on a target from a set of scattered sources, as in the present study. The Coleman, Katz and Menzel study deals with an important substantive correlate of acquaintance nets, namely information diffusion.

We were interested in discovering some of the internal structural features of chains and in making comparisons across chains as well. Among the questions we hoped to answer were the following: How many of the starters —if any—would be able to establish contact with the target through a chain of acquaintances? How many intermediaries would be required to link the ends of the chains? What form would the distribution of chain lengths take? What degree of homogeneity in age, sex, occupation, and other characteristics of participants would be observed within chains? How would complete chains differ from incomplete on these and other dimensions?

An additional comparison was set up by using three distinct starting subpopulations. The target person was a Boston stockbroker; two of the starting populations were geographically removed from him, selected from the state of Nebraska. A third population was selected from the Boston area. One of the Nebraska groups consisted of bluechip stockholders, while the second Nebraska group and the Boston group were "randomly" selected and had no special access to the investment business. By comparisons across these groups we hoped to assess the relative effects of geographical distance and of contact with the target's occupational group. Moreover we hoped to establish a strategy for future experimental extensions of the procedure, in which the sociological characteristics of the starting and target populations would be systematically varied in order to expose features of social structure.

The primary research questions, then, involved a test of the feasibility and fruitfulness of the method as well as an attempt to discover some elementary features of real social nets. Several experimental extensions of the procedure are already underway. A more detailed description of the current method is given in the following sections.

PARTICIPANTS. *Starting Population.* The starting population for the study was comprised of 296 volunteers. Of these, 196 were residents of the state of Nebraska, solicited by mail. Within this group, 100 were systematically chosen owners of blue-chip stocks; these will be designated "Nebraska stockholders" throughout this paper. The rest were chosen from the population at large; these will be termed the "Nebraska random" group. In addition to the two Nebraska groups, 100 volunteers were solicited through an advertisement in a Boston newspaper (the "Boston random" group). Each member of the starting population became the first link in a chain of acquaintances directed at the target person.

*Intermediaries.* The remaining participants in the study, who numbered 453 in all, were in effect solicited by other participants; they were acquaintances selected by previous participants as people likely to extend the chain toward the target. Participation was voluntary. Participants were not paid, nor was money or other reward offered as incentive for completion of chains.

THE DOCUMENT. The 296 initial volunteers were sent a document which was the principal tool of the investigation.[3] The document contained:

a. a description of the study, a request that the recipient become a participant, and a set of rules for participation;
b. the name of the target person and selected information concerning him;
c. a roster, to which each participant was asked to affix his name;
d. a stack of fifteen business reply cards asking information about each participant.

*Rules for Participation.* The document contained the following specific instructions to participants:

a. Add your name to the roster so that the next person who receives this folder will know whom it came from.
b. Detach one postcard from the bottom of this folder. Fill it out and return it to Harvard University. No stamp is needed. The postcard is very important. It allows us to keep track of the progress of the folder as it moves toward the target person.
c. If you know the target person on a personal basis, mail this folder directly to him (her). Do this only if you have previously met the target person and know each other on a first name basis.
d. If you do not know the target person on a personal basis, do not try to contact him directly. Instead, mail this folder to a personal acquaintance who is more likely than you to know the target person. You may send the booklet on to a friend, relative, or acquaintance, but it must be someone you know personally.

*Target Person.* The target person was a stockholder who lives in Sharon, Massachusetts, a suburb of Boston, and who works in Boston proper. In addition to his name, address, occupation and place of employment, participants were told his college and year of graduation, his military service dates, and his wife's maiden name and hometown. One question under investigation was the type of information which people would use in reaching the target.

*Roster.* The primary function of the roster was to prevent "looping," i.e., to prevent people from sending the document to someone who had already received it and sent it on. An additional function of the roster was to motivate people to continue the chains. It was hoped that a list of prior participants, including a personal acquaintance who had sent the document to

[3] A photographic reproduction of this experimental document appears in Milgram, 1969: 110–11.

the recipient, would create willingness on the part of those who received the document to send it on.

*Tracer Cards.* Each participant was asked to return to us a business reply card giving certain information about himself and about the person to whom he sent the document. The name, address, age sex and occupation of the sender and sender's spouse were requested, as were the name, address, sex and age of the recipient. In addition, the nature of the relationship between sender and recipient—whether they were friends, relatives, business associates, etc.—was asked. Finally, participants were asked why they had selected the particular recipient of the folder.

The business reply cards enabled us to keep running track of the progress of each chain. Moreover, they assured us of getting information even from chains which were not completed, allowing us to make comparisons between complete and incomplete chains.

## RESULTS

COMPLETIONS. 217 of the 296 starting persons actually sent the document on to friends. Any one of the documents could reach the target person only if the following conditions were met: 1) recipients were sufficiently motivated to send the document on to the next link in the chain; 2) participants were able to adopt some strategy for moving the documents closer to the target (this condition further required that the given information allow them to select the next recipient in a manner that increased the probability of contacting the target); 3) relatively short paths were in fact required to link starters and target (otherwise few chains would remain active long enough to reach completion). Given these contingencies, there was serious doubt in the mind of the investigators whether any of the documents, particularly those starting in an area remote from the target person, could move through interlocking acquaintance networks and converge on him. The actual outcome was that 64 of the folders, or 29 per cent of those sent out by starting persons, eventually reached the target.

DISTRIBUTION OF CHAIN LENGTHS. *Complete Chains.* Figure 1 shows the frequency distribution of lengths of the completed chains. "Chain length" is here defined as the number of intermediaries required to link starters and target. The mean of the distribution is 5.2 links.

It was unclear on first inspection whether the apparent drop in frequency at the median length of five links was a statistical accident, or whether the distribution was actually bimodal. Further investigation revealed that the summary relation graphed in Figure 1 concealed two underlying distributions: when the completed chains were divided into those which approached the target through his hometown and those which approached him via

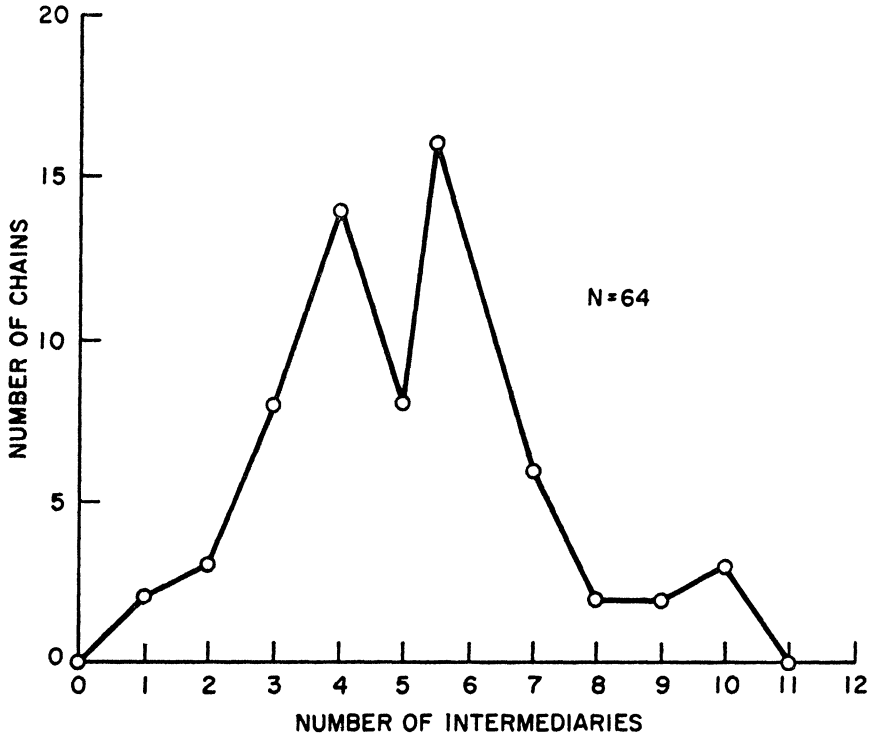


FIGURE 1

*Lengths of Completed Chains*

Boston business contracts, two distinguishable distributions emerged. The mean of the Sharon distribution is 6.1 links, and that of the Boston distribution is 4.6. The difference is significant at a level better than .0005, as assessed by the distribution-free Mann-Whitney *U* test. (Note that more powerful statistical tests of the significance of differences between means cannot be applied to these data, since those tests assume normality of underlying distributions. The shape of the true or theoretical distribution of lengths of acquaintance chains is precisely what we do not know.)

Qualitatively, what seems to occur is this. Chains which converge on the target principally by using geographic information reach his hometown or the surrounding areas readily, but once there often circulate before entering the target's circle of acquaintances. There is no available information to narrow the field of potential contacts which an individual might have within the town. Such additional information as a list of local organizations

of which the target is a member might have provided a natural funnel, facilitating the progress of the document from town to target person. By contrast, those chains which approach the target through occupational channels can take advantage of just such a funnel, zeroing in on him first through the brokerage business, then through his firm.

*Incomplete Chains.* Chains terminate either through completion or dropout: each dropout results in an incomplete chain. Figure 2 shows the number of chains which dropped out at each "remove" from the starting population. The "0th remove" represents the starting population itself: the "first remove" designates the set of people who received the document directly from members of the starting population. The "second remove" received the document from the starters via one intermediary, the third through two intermediaries, etc. The length of an incomplete chain may be defined as the number of removes from the start at which dropout occurs, or, equivalently, as the number of transmissions of the folder which precede dropout. By this definition, Figure 2 represents a frequency distribution of the lengths of incomplete chains. The mean of the distribution is 2.6 links.

The proportion of chains which drop out at each remove declines as

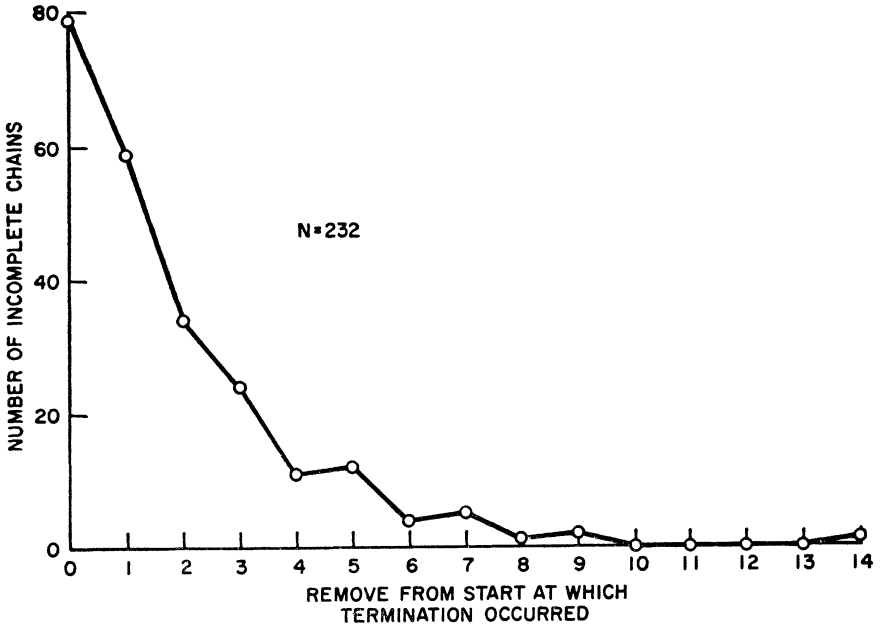


FIGURE 2

*Lengths of Incomplete Chains*

chains grow in length, if that proportion is based on all chains active at each remove (those destined for completion as well as incompletion). About 27 per cent of the 296 folders sent to the starting population are not sent on. Similarly, 27 per cent of the 217 chains actually initiated by the starters die at the first remove. The percentage of dropouts then appears to fall. It also begins to fluctuate, as the total number of chains in circulation grows small, and an increasing proportion of completions further complicates the picture.

It was argued earlier that, in theory, any two people can be linked by at least one acquaintance chain of finite length, barring the existence of totally isolated cliques within the population under study. Yet, incomplete chains are found in our empirical tracing procedure because a certain proportion of those who receive the document do not send it on. It is likely that this occurs for one of two major reasons: 1) individuals are not motivated to participate in the study; 2) they do not know to whom to send the document in order to advance it toward the target.

For purposes of gauging the significance of our numerical results, it would be useful to know whether the dropouts are random or systematic, i.e., whether or not they are related to a chain's prognosis for rapid completion. It seems possible, for example, that dropouts are precisely those people who are least likely to be able to advance the document toward the target. If so, the distribution of actual lengths of completed chains would understate the true social distance between starters and target by an unknown amount. (Even if dropouts are random, the observed distribution understates the true distribution, but by a potentially calculable amount.) We can offer some evidence, however, that this effect is not powerful.

First, it should be clear that, though people may drop out because they see little possibility that any of their acquaintances can advance the folder toward the target, their subjective estimates are irrelevant to the question just raised. Such subjective estimates may account for individual decisions not to participate; they do not tell us whether chains that die in fact would have been longer than others had they gone to completion. People have poor intuitions concerning the lengths of acquaintance chains. Moreover, people can rarely see beyond their own acquaintances; it is hard to guess the circles in which friends of friends—not to mention people even more remotely connected to oneself—may move.

More direct evidence that dropouts may be treated as "random" can be gleaned from the tracer cards. It will be recalled that each participant was asked for information not only about himself but also about the person to whom he sent the document. Thus some data were available even for dropouts, namely age, sex, the nature of their relationship to the people

TABLE 1

*Activity of Chains at Each Remove*

| All Chains | | | | | Incomplete Chains Only | | | |
|---|---|---|---|---|---|---|---|---|
| Remove | Chains Reaching this Remove | Completions at this Remove | Dropouts at this Remove | Per cent Dropouts | Remove | Chains Reaching this Remove | Dropouts at this Remove | Per cent Dropouts |
| 0 | 296 | 0 | 79 | 27 | 0 | 232 | 79 | 34 |
| 1 | 217 | 0 | 59 | 27 | 1 | 153 | 59 | 39 |
| 2 | 158 | 2 | 34 | 22 | 2 | 94 | 34 | 36 |
| 3 | 122 | 3 | 24 | 20 | 3 | 60 | 24 | 40 |
| 4 | 95 | 8 | 11 | 12 | 4 | 36 | 11 | 31 |
| 5 | 76 | 14 | 12 | 16 | 5 | 25 | 12 | 48 |
| 6 | 50 | 8 | 4 | 8 | 6 | 13 | 4 | 31 |
| 7 | 38 | 16 | 5 | 13 | 7 | 9 | 5 | 55 |
| 8 | 17 | 6 | 1 | 6 | 8 | 4 | 1 | 25 |
| 9 | 10 | 2 | 2 | 20 | 9 | 3 | 2 | 67 |
| 10 | 6 | 2 | 0 | 0 | 10 | 1 | 0 | 0 |
| 11 | 4 | 3 | 0 | 0 | 11 | 1 | 0 | 0 |
| 12 | 1 | 0 | 0 | 0 | 12 | 1 | 0 | 0 |
| 13 | 1 | 0 | 0 | 0 | 13 | 1 | 0 | 0 |
| 14 | 1 | 0 | 1 | 100 | 14 | 1 | 1 | 100 |

preceding them in the chain, and the reason the dropout had been selected to receive the document. These four variables were tabulated for dropouts versus non-dropouts. None of the resulting contingency tables achieved the .05 level of statistical significance by chi-square test; we are therefore led to accept the null hypothesis of no difference between the two groups, at least on this limited set of variables. Of course, a definitive answer to the question of whether dropouts are really random must wait until the determinants of chain length are understood, or until a way is found to force all chains to completion.[4]

SUBPOPULATION COMPARISONS. A possible paradigm for future research using the tracing procedure described here involves systematic variation of the relationship between the starting and target populations. One such study, using Negro and White starting and target groups, has already been completed by Korte and Milgram (in press). In the present study, which involved only a single target person, three starting populations were used (Nebraska random, Nebraska stockholders, and Boston random.) The relevant experimental questions were whether the proportion of completed chains or mean chain lengths would vary as a function of starting population.

*Chain Length.* Letters from the Nebraska subpopulations had to cover a geographic distance of about 1300 miles in order to reach the target, whereas letters originating in the Boston group almost all started within 25 miles of his home and/or place of work. Since social proximity depends in part on geographic proximity, one might readily predict that complete chains originating in the Boston area would be shorter than those originating in Nebraska. This presumption was confirmed by the data. As Table 2 shows, chains originating with the Boston random group showed a mean length of 4.4 intermediaries between starters and target, as opposed to a mean length of 5.7 intermediaries for the Nebraska random group. ($p \leqq .001$ by

[4] Professor Harrison White of Harvard University has developed a technique for adjusting raw chain length data to take account of the dropout problem. His method assumes that dropouts are "random," in the following sense. An intermediary who knows the target sends him the folder, completing the chain, with probability 1. Otherwise, an intermediary throws away the folder with fixed probability $1-\alpha$, or sends it on with probability $\alpha$. If sent on, there is a probability $Q_i$ (which depends on number of removes from the origin) that the next intermediary knows the target. The data is consistent with a value for $\alpha$ of approximately 0.75, independent of remove from the origin, and hence with a "random" dropout rate of 25 per cent. The limited data further suggest that $Q_i$ grows in a "staircase" pattern from zero (at zero removes from the starting population) to approximately one-third at six removes, remaining constant thereafter. Based on these values, the hypothetical curve of completions with no dropouts resembles the observed curve shifted upward; the median length of completed chains rises from 5 to 7, but no substantial alteration is required in conclusions drawn from the raw data.

TABLE 2

*Lengths of Completed Chains*

| | Frequency Distribution | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number of Intermediaries | | | | | | | | | | | | |
| Population | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | Total |
| Nebraska Random | 0 | 0 | 0 | 1 | 4 | 3 | 6 | 2 | 0 | 1 | 1 | 0 | 18 |
| Nebraska Stock | 0 | 0 | 0 | 3 | 6 | 4 | 6 | 2 | 1 | 1 | 1 | 0 | 24 |
| Boston Random | 0 | 2 | 3 | 4 | 4 | 1 | 4 | 2 | 1 | 0 | 1 | 0 | 22 |
| All | 0 | 2 | 3 | 8 | 14 | 8 | 16 | 6 | 2 | 2 | 3 | 0 | 64 |

| Means | |
|---|---|
| Starting Population | Mean Chain Length |
| Nebraska Random | 5.7 |
| Nebraska Stockholders | 5.4 |
| All Nebraska | 5.5 |
| Boston Random | 4.4 |
| All | 5.2 |

a one-tailed Mann-Whitney U test.) Chain length thus proved sensitive to one demographic variable—place of residence of starters and target.

The Nebraska stockholder group was presumed to have easy access to contacts in the brokerage business. Because the target person was a stockbroker, chains originating in this group were expected to reach the target more efficiently than chains from the Nebraska random group. The chain-length means for the two groups, 5.7 intermediaries for the random sample and 5.4 for the stockholders, differed in the expected direction, but the difference was not statistically significant by the Mann-Whitney test. The stockholders used the brokerage business as a communication channel more often than did the random group; 60.7 per cent of all the participants in chains originating with the stockholder group reported occupations connected with finance, while 31.8 per cent of participants in chains originating in the Nebraska random group were so classified.

*Proportion of Completions.* As indicated in Table 3, the proportions of chains completed for the Nebraska random, Nebraska stockholder, and Boston subpopulations were 24 per cent, 31 per cent and 35 per cent, respectively. Although the differences are not statistically significant, there is a weak tendency for higher completion rates to occur in groups where mean length of completed chains is shorter. This result deserves brief discussion.

Let us assume that the dropout rate is constant at each remove from the start. If, for example, the dropout rate were 25 per cent then any chain would have a 75 per cent probability of reaching one link, $(.75)^2$ of reaching two links, etc. Thus, the longer a chain needed to be in order to reach completion, the less likely that the chain would survive long enough to run its full course. In this case, however, chain-length differences among the three groups were not sufficiently large to produce significant differences in completion rate. Moreover, if the dropout rate declines as chains grow long, such a decrease would off-set the effect just discussed and weaken the observed inverse relation between chain length and proportion of completions.

TABLE 3

*Proportion of Completions for Three Starting Populations*

| | Starting Population | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Nebraska Random | | Nebraska Stock. | | Boston | | Total | |
| Complete | 18 | (24%) | 24 | (31%) | 22 | (35%) | 64 | (29%) |
| Incomplete | 58 | (76%) | 54 | (69%) | 41 | (65%) | 153 | (71%) |
| | 76 | (100%) | 78 | (100%) | 63 | (100%) | 217 | (100%) |

$\chi^2=2.17$, df.$=2$, $p>.3$, N.S.

COMMON CHANNELS. As chains converge on the target, common channels appear—that is, some intermediaries appear in more than one chain. Figure 3 shows the pattern of convergence. The 64 letters which reached the target were sent by a total of 26 people. Sixteen, fully 25 per cent, reached the target through a single neighbor. Another 10 made contact through a single business associate, and 5 through a second business associate. These three "penultimate links" together accounted for 48 per cent of the total completions. Among the three, an interesting division of labor appears. Mr. G,

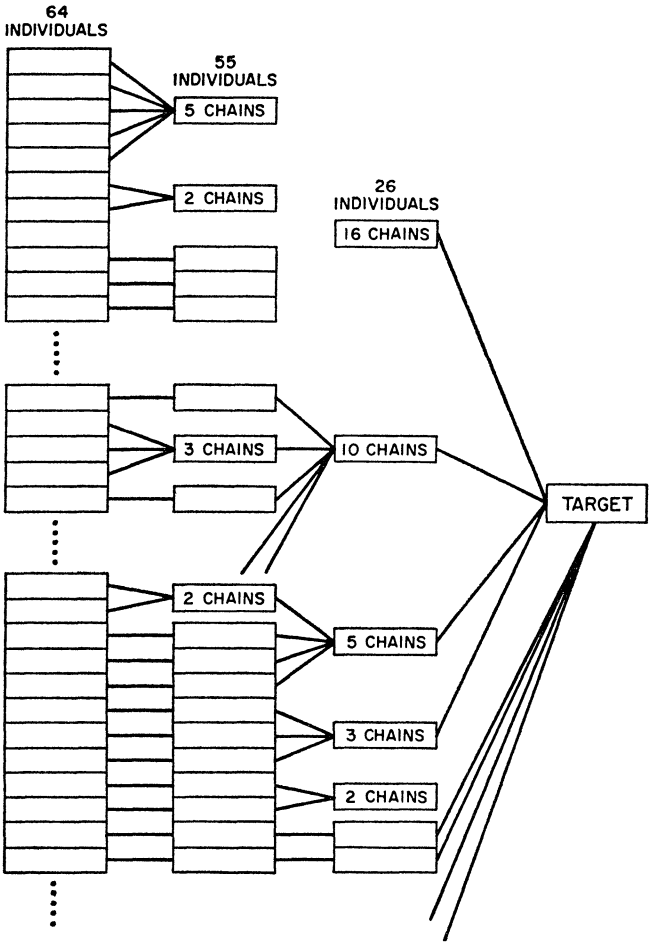


FIGURE 3

*Common Paths Appear as Chains Converge on the Target*

who accounted for 16 completions, is a clothing merchant in the target's hometown of Sharon; Mr. G funnelled toward the target those chains which were advancing on the basis of the target's place of residence. Twenty-four chains reached the target from his hometown; Mr. G accounted for ⅔ of those completions. All the letters which reached Mr. G came from residents of Sharon. By contrast, Mr. D and Mr. P, who accounted for 10 and 5 completions, respectively, were contacted by people scattered around the Boston area, and in several cases, by people living in other cities entirely. On the other hand, whereas Mr. G received the folder from Sharon residents in a wide variety of occupations, D and P received it almost always from stockbrokers. A scattering of names appear two or three times on the list of penultimate links; seventeen names appear once each.

Convergence appeared even before the penultimate link. Going one step further back, to people two removes from the target, we find that the 64 chains passed through 55 individuals. One man, Mr. B, appeared 5 times, and on all occasions sent the document to Mr. G. Other individuals appeared two or three times each.

Additional Characteristics of Chains. Eighty-six per cent of the participants sent the folder to persons they described as friends and acquaintances; 14 per cent sent it to relatives. The same percentages had been observed in an earlier pilot study.

Data on patterns of age, sex and occupation support the plausible hypothesis that participants select recipients from a pool of individuals similar to themselves. The data on age support the hypothesis unequivocally; the data on sex and occupation are complicated by the characteristics of the target and the special requirement of establishing contact with him.

Age was bracketed into ten-year categories and the ages of those who sent the document tabled against the ages of those to whom they sent it. On inspection the table showed a strong tendency to cluster around the diagonal, and a chi-square test showed the association to be significant at better than the .001 level.

Similarly, the sex of each sender was tabled against the sex of the corresponding recipient. Men were ten times more likely to send the document to other men than to women, while women were equally likely to send the folder to males as to females ($p<.001$). These results were affected by the fact that the target was male. In an earlier pilot study using a female target, both men and women were three times as likely to send the document to members of the same sex as to members of the opposite sex. Thus there appear to be three tendencies governing the sex of the recipient: (1) there is a tendency to send the document to someone of one's own sex, but (2) women are more likely to cross sex lines than men, and (3) there

is a tendency to send the document to someone of the same sex as the target person.

The occupations reported by participants were rated on two components—one of social status and one of "industry" affiliation, that is, the subsector of the economy with which the individual would be likely to deal. The coding system was *ad hoc,* designed to fit the occupational titles supplied by participants. Tabling the status and "industry" ratings for all senders of the document against those of respective recipients, we observed a strong tendency for people to select recipients similar to themselves on both measures ($p<.001$ for both tables). However, the strength of the relationship for industry seemed to be largely due to a tendency for the folder to stay within the finance field once it arrived there, obviously because the target was affiliated with that field. Moreover, the participants in the study were a heavily middle-class sample, and the target was himself a member of that class. Thus there was no need for the document to leave middle-class circles in progressing from starters to target.

When separate contingency tables were constructed for complete and incomplete chains, the above results were obtained for both tables. Similarly, when separate tables were constructed for chains originating in the 3 starting populations, the findings held up in all 3 tables. Thus, controlling for completion of chains or for starting population did not affect the finding of demographic homogeneity within chains.

## CONCLUSIONS

The contribution of the study lies in the use of acquaintance chains to extend an individual's contacts to a geographically and socially remote target, and in the sheer size of the population from which members of the chains were drawn. The study demonstrated the feasibility of the "small world" technique, and took a step toward demonstrating, defining and measuring inter-connectedness in a large society.

The theoretical machinery needed to deal with social networks is still in its infancy. The empirical technique of this research has two major contributions to make to the development of that theory. First, it sets an upper bound on the minimum number of intermediaries required to link widely separated Americans. Since subjects cannot always foresee the most efficient path to a target, our trace procedure must inevitably produce chains longer than those generated by an accurate theoretical model which takes full account of all paths emanating from an individual. The mean number of intermediaries observed in this study was somewhat greater than five;

additional research (by Korte and Milgram) indicates that this value is quite stable, even when racial crossover is introduced. Both the magnitude and stability of the parameter need to be accounted for. Second, the study has uncovered several phenomena which future models should explain. In particular, the convergence of communication chains through common individuals is an important feature of small world nets, and it should be accounted for theoretically.

There are many additional lines of empirical research that may be examined with the small world method. As suggested earlier, one general paradigm for research is to vary the characteristics of the starting person and the target. Further, one might systematically vary the information provided about the target in order to determine, on the psychological side, what strategies people employ in reaching a distant target, and on the sociological side, what specific variables are critical for establishing contact between people of given characteristics.

## REFERENCES

Abelson, R. P.
1967 "Mathematical models in social psychology." Pp. 1–54 in L. Berkowitz (ed.) Advances in Experimental Social Psychology, Vol. III. New York: Academic Press.

Bailey, N. T. J.
1957 The Mathematical Theory of Epidemics. New York: Hafner.

Coleman, J. S., E. Katz and H. Menzel
1966 Medical Innovation: A Diffusion Study. Indianapolis: Bobbs-Merrill.

Fararo, T. J. and M. H. Sunshine
1964 A Study of a Biased Friendship Net. Syracuse: Youth Development Center, Syracuse University.

Foster, C. C., A. Rapoport and C. J. Orwant
1963 "A study of a large sociogram II. Elimination of free parameters." Behavioral Science 8(January):56–65.

Gurevitch, M.
1961 The Social Structure of Acquaintanceship Networks. Unpublished doctoral dissertation, Cambridge: M.I.T.

Korte, C. and S. Milgram
Acquaintance Links Between White and Negro Populations: Application of the Small World Method. Journal of Personality and Social Psychology **(in press).**

Milgram, S.
1967 "The small world problem." Psychology Today 1(May):61–67.
1969 "Interdisciplinary thinking and the small world problem." Pp. 103–120 in Muzafer Sherif and Carolyn W. Sherif (eds.) Interdisciplinary Relationships in the Social Sciences. Chicago: Aldine Publishing Company.

Pool, I. and M. Kochen
A Non-Mathematical Introduction to a Mathematical Model. Undated mimeo. Cambridge: M.I.T.

Rapoport, A.
1953 "Spread of information through a population with socio-structural bias." Bulletin of Mathematical Biophysics 15(December):523–543.
1963 "Mathematical models of social interaction." Pp. 493–579 in R. D. Luce, R. R. Bush and E. Galanter (eds.) Handbook of Mathematical Psychology, Vol. II. New York: John Wiley and Sons.

Rapoport, A. and W. J. Horvath
1961 "A study of a large sociogram." Behavioral Science 6(October):279–291.