

AD-A037 508

MINNESOTA SYSTEMS RESEARCH INC MINNEAPOLIS

F/G 5/10

A DESIGN FOR OBSERVER SAMPLING OF MEMBER INTERACTION IN A LARGE--ETC(U)

FEB 77 R E SYKES

N00014-75-C-0075

UNCLASSIFIED

TR-3

NL

1 of 1
AD
A037508



REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER TR #3	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A DESIGN FOR OBSERVER SAMPLING OF MEMBER INTERACTION IN A LARGE, SPONTANEOUS GROUP	5. TYPE OF REPORT & PERIOD COVERED Technical Report	
7. AUTHOR(s) Richard E. Sykes	6. PERFORMING ORG. REPORT NUMBER 1477R-3	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Minnesota Systems Research, Inc. 2412 University Avenue S. E. Minneapolis, MN 55414	8. CONTRACT OR GRANT NUMBER(s) N00014-75-0075	
11. CONTROLLING OFFICE NAME AND ADDRESS Organizational Effectiveness Research Program Office of Naval Research (Code 452) Arlington, VA 22217	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 170-790	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	12. REPORT DATE February 17, 1977	
	13. NUMBER OF PAGES 34	
	15. SECURITY CLASS. (of this report) Unclassified	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) interaction proximity sample size sampling interaction similarity sampling design informal groups attraction		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Some social psychological problems require data on frequency of interaction. In a large (greater than 20 persons) spontaneous group this raises methodological problems of great complexity. A technique used for sampling interaction by observers of such a large, spontaneous group is described. The technique was based on estimates made by Larntz (Appendix 3.1) of size of the counts necessary to infer the existence of informal groups from a large matrix of such counts.		

AD A037508

DDC
RECEIVED
MAR 31 1977
A

390 596

DDC FILE COPY

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

AD V035208

DDC LIFE COBA

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

TR #3--February 17, 1977

A DESIGN FOR OBSERVER SAMPLING OF
MEMBER INTERACTION IN A LARGE, SPONTANEOUS GROUP ¹

by

Richard E. Sykes, Ph.D.

Prepared as part of the project Informal Social Network Formation in Navy Training Units sponsored by the Organizational Effectiveness Research Program, Office of Naval Research, under Contract No. N00014-75-0075, NR 170-790.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Published by Minnesota Systems Research, Inc., 2412 University Avenue, S. E., Minneapolis, MN 55414

¹ The author would like to thank K. Larntz and G. Blank for their suggestions and advice in developing this design.

TABLE OF CONTENTS

Technical Report No.3

A Design for Observer Sampling of Member Interaction in a
Large, Spontaneous Group by Richard E. Sykes 1

Appendix 3.1

Determining Sample Size Necessary to Detect Groups from
Interaction Frequencies by Kinley Larntz 25

INTRODUCTION

Social psychological research problems sometimes require collection of data on intragroup communication. The investigator must know who talked to whom and how frequently. Two factors which have the potential for complicating the collection of such data are group size and type of communication pattern. Chapple, Bales and many others have collected such data, but usually in groups either so small or so orderly that one observer could monitor the entire group simultaneously.

If subjects are to be identified individually, then size is an important constraint. The observer must quickly be able to identify every subject individually. The larger the group, the more difficult it is to do this. No doubt there is some upper limit on size (so far as I know, unknown) for ready identification by the average observer. Size also effects geographical area. Larger groups require a larger area and constrain visual recognition across distances.

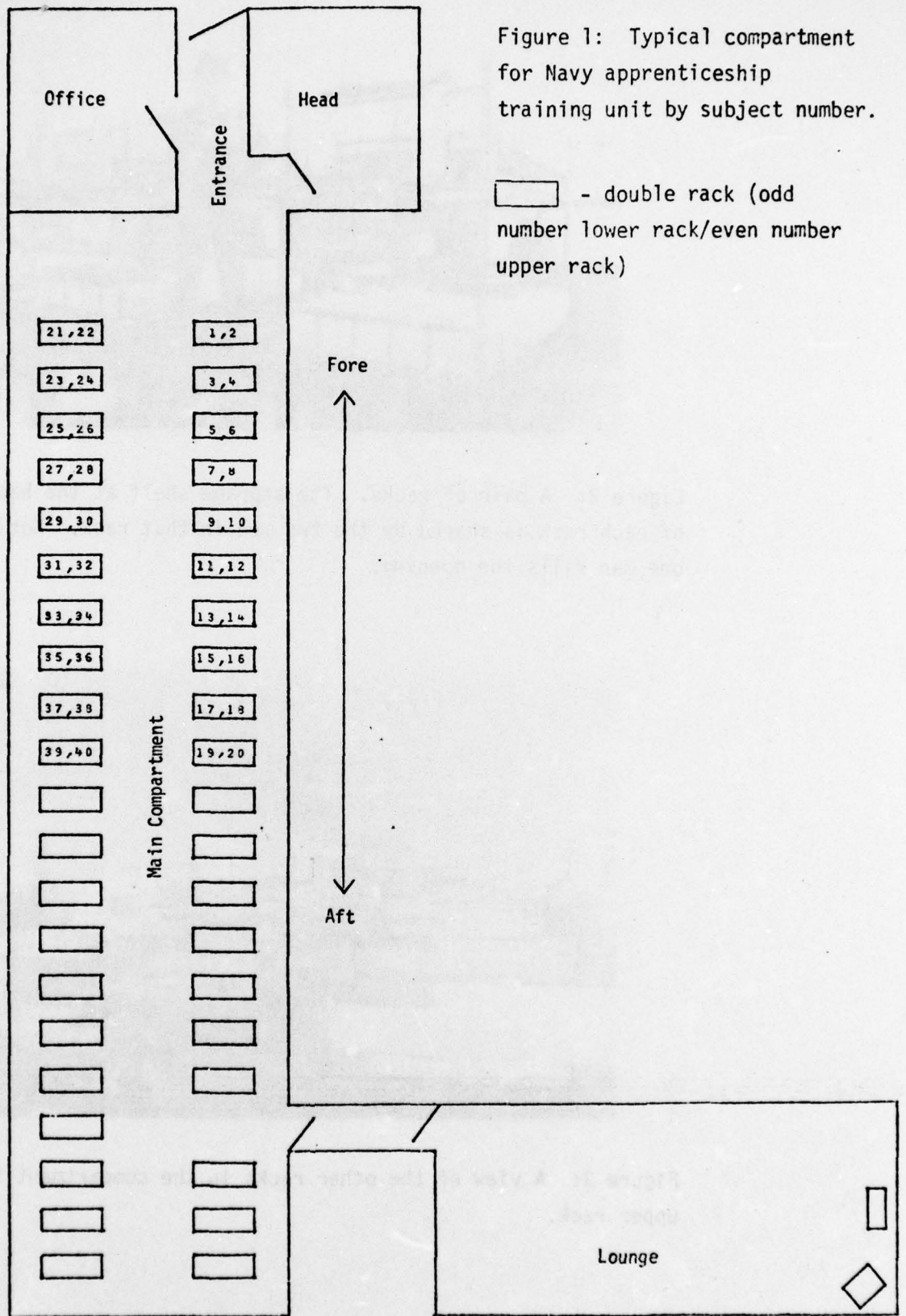
Even more important than size is communication pattern. For instance, a very large lecture class can be observed so long as the lecturer is the focus of communication. The lecturer's communicative acts and those addressed to him by auditors can be coded without great difficulty. When such a pattern breaks down and spontaneous, simultaneous dialogue begins within many dyads most category systems interpret this as a manifestation of disorder and provide a residual category to describe it such as "silence or confusion" (Flanders, 1970, 34).

There are many natural groups in which spontaneous, simultaneous communication within many dyads is not only normal but a manifestation of order, not disorder. In a large office many separate conversations occur simultaneously. This is also true of a party. Whatever the group or situation, multiple dyadic interaction requires the observer to adopt a different strategy of data collection than is used in small groups or groups with a centralized communication pattern.

THE PROBLEM IN THE NAVY STUDY

The focus of the Navy study was the influence of proximity and similarity on informal group formation. As in the Sherif studies I wished to study the emergence of informal groups in a natural setting. It was necessary to record the frequency of interaction of individuals as these smaller, informal groups differentiated within a larger one, and be able to document the emergence of patterns. One way to do this was to observe subjects in training units varying in size from 30 to 60 men interacting spontaneously and simultaneously in the compartment (barracks wing) in which they lived.

Figure 1 is a sketch of a typical compartment of such a unit. The men were assigned to double racks (bunks) (for typical examples see Figures 2-7). Each compartment held a maximum of 80 men in 40 racks. In addition to the main compartment itself, there was a small office and entrance hallway to the fore. Aft, the compartment connected to a lounge where men could view TV or use pop machines. At the fore of the compartment, on the starboard side, were doors leading to the head, showers and drying room. Each unit spent free time from 1530 until taps



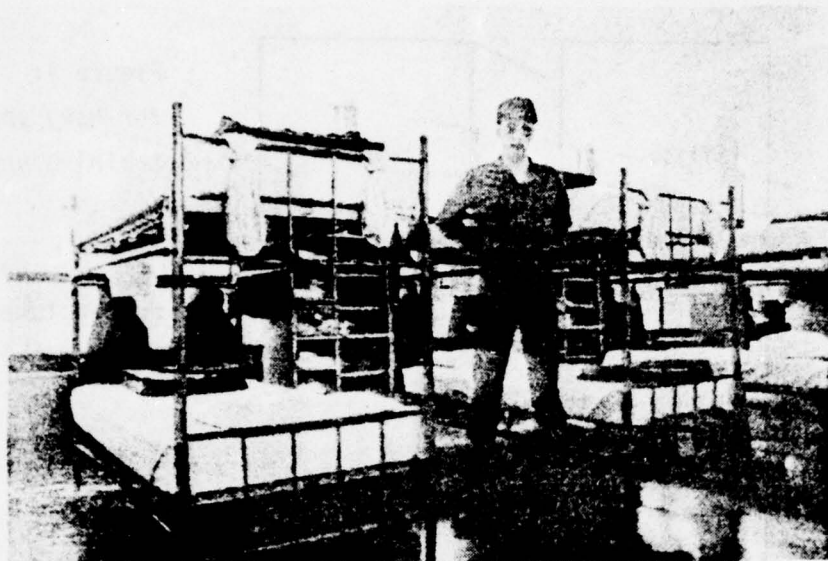


Figure 2: A pair of racks. The storage shelf at the back left of each rack is shared by the two men in that rack. Notice that one man fills the opening.

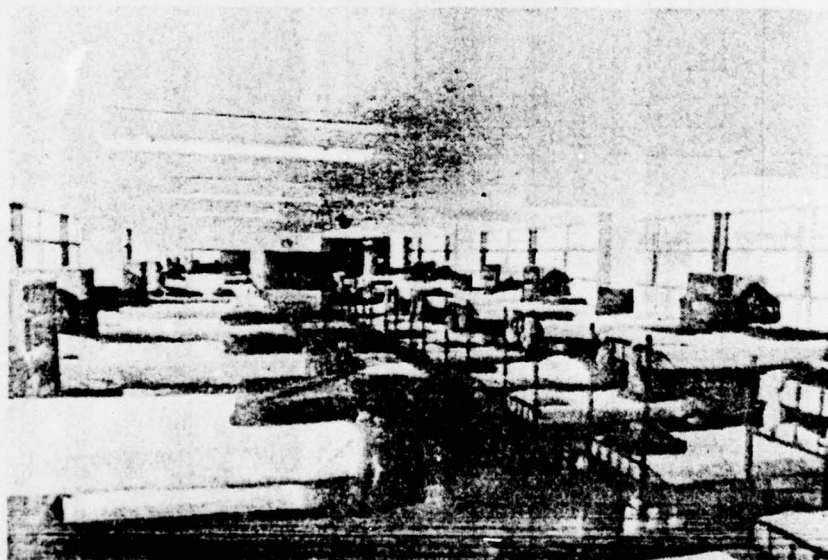


Figure 3: A view of the other racks in the compartment from an upper rack.

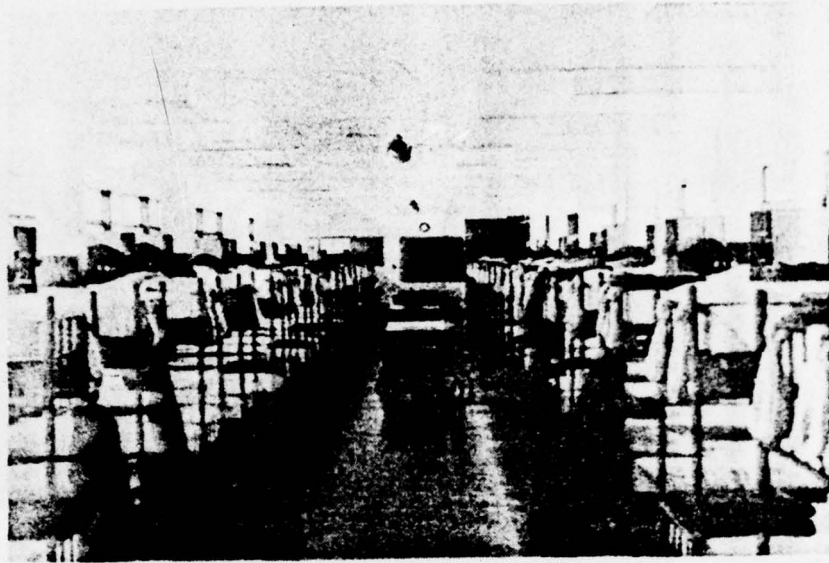


Figure 4: A perspective down the center of the compartment. Notice the tables where men may read, write, play cards (contrary to regulations) or sew.

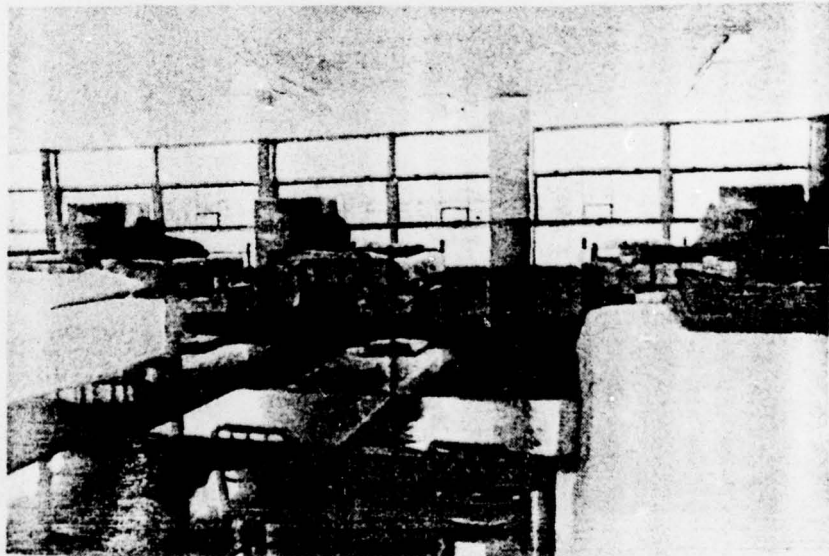


Figure 5: A side view of the compartment from the perspective of an upper rack. This is a typical view of observers when collecting interaction data.

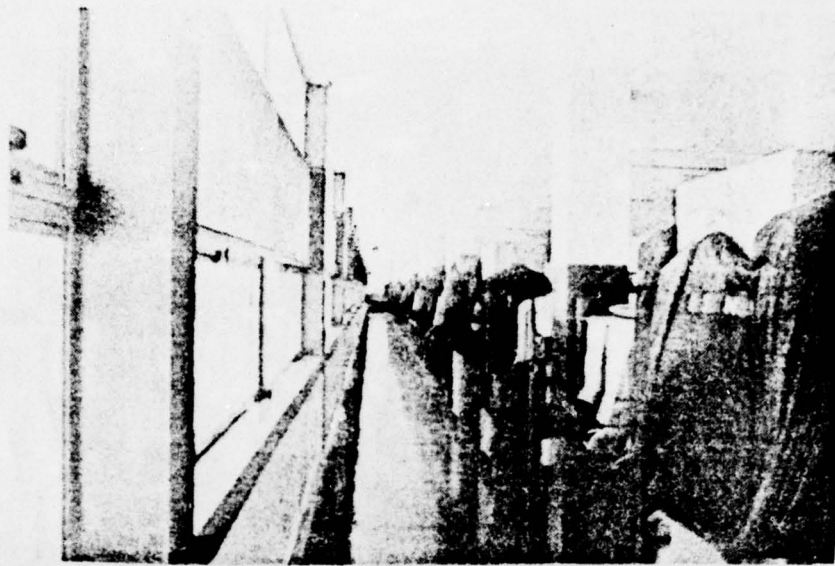


Figure 6: The passageway along the starboard bulkhead. Most observations were taken from the vantage of this relatively inconspicuous and little-used location.



Figure 7: The smoking and television lounge attached to the compartment. Notice the television set and Coke machine, the foci of activities. Many units spent well over a quarter of their free time watching television.

(2130) within this area. Interaction patterns developed within a new training unit during their free time. Interaction was usually spontaneous and simultaneous among several different groups of trainees.

Because of their number, physical location and the decentralized communication pattern it was not possible to observe all the men simultaneously. It was necessary to develop a sampling design based on a knowledge of the size of sample necessary for analysis and inference as estimated by Larntz (see the other paper included with this report). For reasons stated below it seemed important to collect data on interpersonal communication within other areas of the training battalion including the galley, classroom and classroom lounge (see Figures 8-9) where interaction was also simultaneous and spontaneous. Here too, it was necessary to develop a sampling design.

SAMPLING THE COMPARTMENT

In developing the design, consideration was given to several alternatives. In a previous study six subjects in a group of 80 had been sampled on an ordered, rotation basis. Using a list of the six, the observer had located the first subject on the list and recorded his communicative behavior at the time located, then the second, and so on. When the sixth subject was coded the observer started again at the beginning of the list. All six were coded within a time limit. This method entailed successive search within a geographical area containing several rooms, one of which was about 160 feet long and contained many people. Finding six subjects was very time consuming. Successively locating an entire unit would have been even more time consuming and would have limited the total number of samples to a small number. Any method involving search for individuals was impractical.

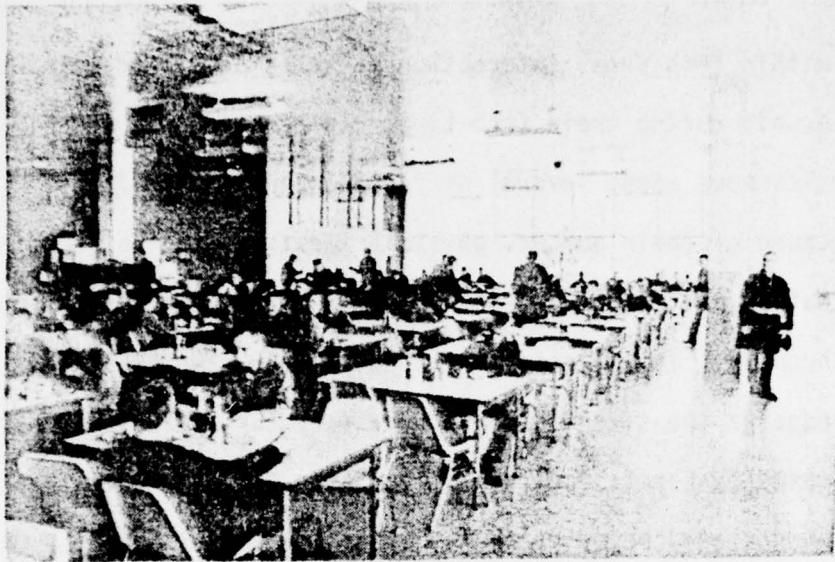


Figure 8: The galley where apprentices ate their meals. They seated themselves at vacancies at these tables after going through a cafeteria line. The unit marched to the galley as a group.

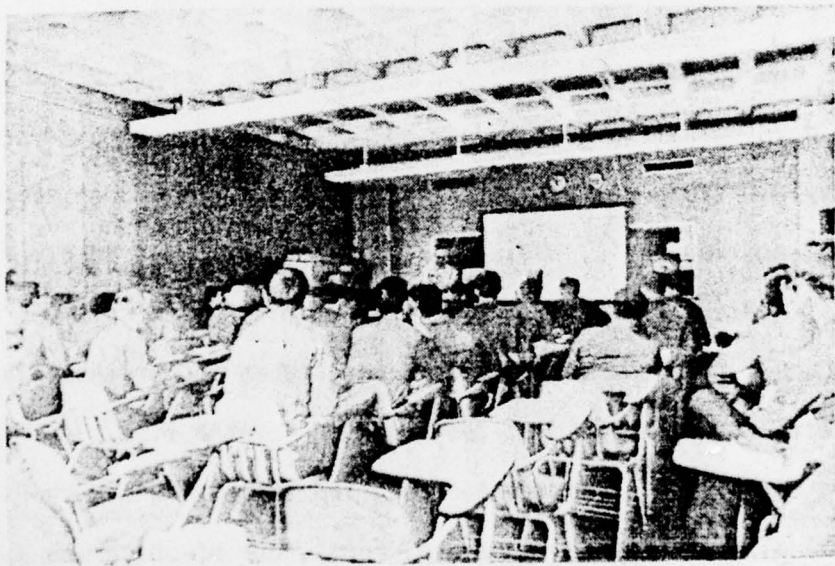


Figure 9: One of the classrooms. Each unit was in training for ten curriculum days.

The alternative we choose was a form of area sampling. The compartment and lounge were divided into ten sampling areas (see Figure 10). Each area was delimited, based on pretesting, considering as a criteria what size geographical area could normally be quickly scanned by an observer to determine: a) who was in the area; b) if someone was in the area, whether interaction was occurring; and c) if interaction was occurring, then with whom?

It was not practical to divide the area into units of equal size. Certain large areas were seldom used, while some small areas were intensively used. In Figure 10 the numbered sampling areas are displayed. Dividing lines between areas went through the center of racks. This may at first seem surprising, but is due to the fact that pretests showed that subjects lying flat on their backs in the middle of racks seldom interacted. They were more apt to sit up on one side of the rack or the other when talking to another subject.

One study unit occupied about half the racks in the compartment. All the racks in Area Seven were unoccupied. While subjects occasionally used this area, activity was not great considering its size. Similarly Area One was larger, including the entrance passageway and a small office. Much interaction took place in the equally-sized Area Two through Six. By walking down the starboard isle next to the bulkhead the observer could inspect one sampling area at a glance. He then coded the area: a) empty of subjects; or b) if occupied he identified the occupants and whether or not they were interacting; and c) if they were interacting he coded the interaction partner(s).

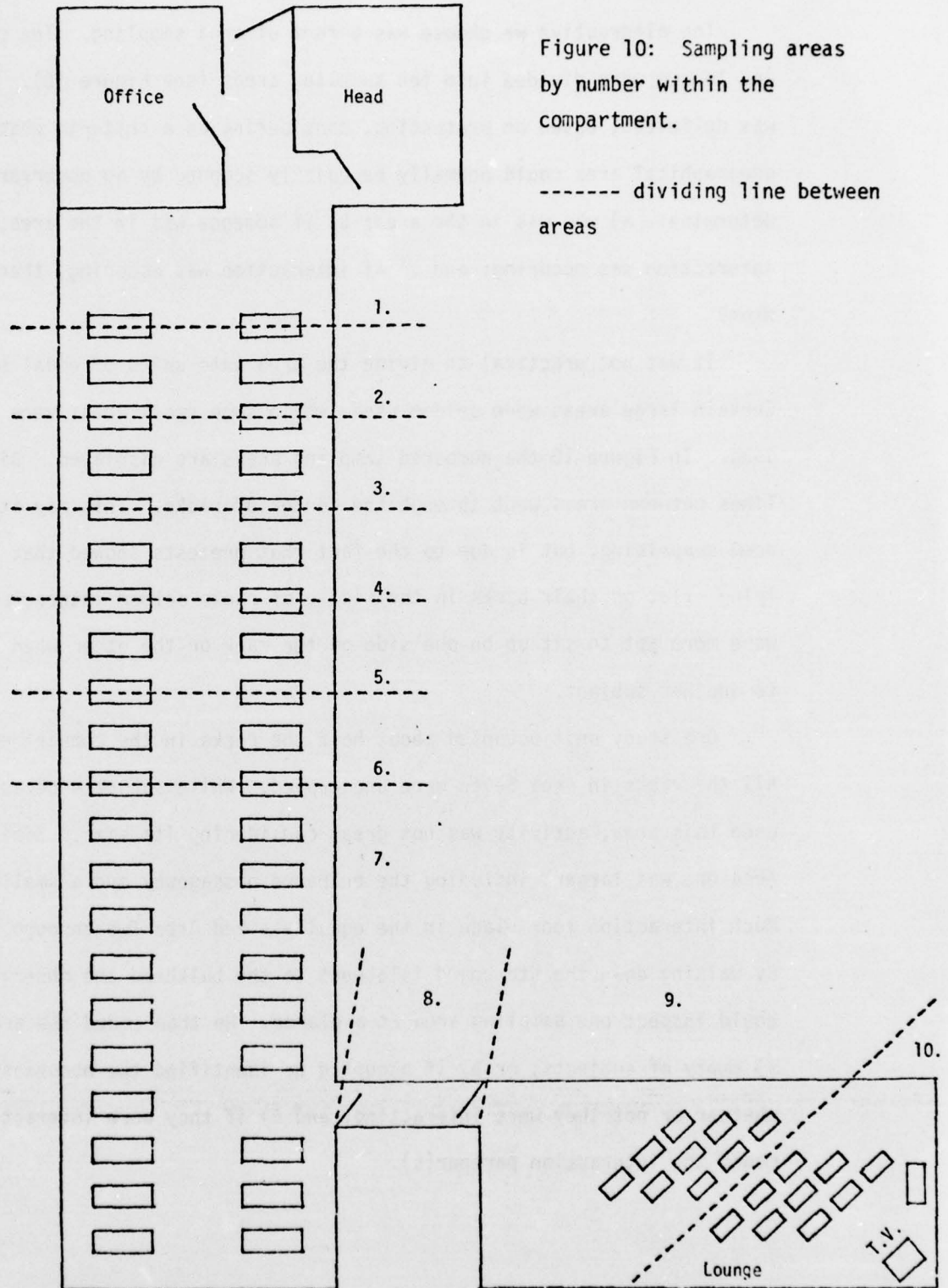


Figure 10: Sampling areas by number within the compartment.

----- dividing line between areas

The hallway into the lounge was considered a sampling area (Area 8). The lounge was divided into two areas. The chairs were lined up in several rows facing the television, which was mounted in a corner. A more or less diagonal line divided rows of chairs closest to the television from rows in the back. One sampling area in the lounge included that part of the area close to the television; the other that part which was more distant. While these two areas were often crowded the attention of the subjects was usually focused on the television. This focus effectively constrained interaction so that the observer could readily code for each area.

After the areas were delimited random lists (without replacement) of the ten areas were generated. After pretesting it was found that it was possible to sample all ten areas in 15 minutes. The observers were instructed to sample the ten areas in random order for 15 minutes. They would then sample the ten areas the next 15 minutes using a different random order. Sampling was carried out each weekday from 1530 to 2130 on the 15 minute schedule with a half hour break for chow. It was possible to get from 18 to 22 samples per day using this schedule.

It was necessary to enunciate several conventions for observers to follow in utilizing this design. The first convention pertained to those lying in the middle of their racks and therefore in the middle of the imaginary line dividing sampling areas. Such a person was always considered to be in the area aft of the dividing line. Thus, a person lying in the middle of the first rack was considered to be in area two.

The second convention pertained to interarea interaction. For obvious reasons it was necessary to avoid counting a subject twice, if for instance, two subjects were conversing across the boundary of an area. Since both

speaker and listener were coded it would have been possible to code both present in each area, thus compromising the independence of each count. To avoid this, when interarea interaction was occurring it was coded only if the other subject was to the aft of the area being coded. Thus, if subject A in Area Two talked to subject B in Area Three, when Area Two was sampled the interaction would be toward the aft and therefore coded. When Area Three was sampled, since the same dyad would be interacting toward the fore it was not counted. Counting the same interaction twice was thus avoided with this procedure.

Finally, by convention (or definition) we utilized a somewhat broader definition of "interaction" than many have used. Often "interaction" is limited to focused verbal behavior. There are several other essentially symbolic activities which are focused, but which may not involve continuous overt verbal behavior. In the first place interacting includes not just talking, but active listening. Since the observer scanned an area, taking a fix on its occupants and whether or not they were interacting, he did not listen to a sequence of turn taking in conversation. The observation of each area was somewhat like a snapshot, recording interaction as of that moment. Therefore, listening was considered as important as speaking. The first code indicated the speaker. The subsequent code specified the active listener(s).

Secondly, mutually focused symbolic activities in which two or more engaged such as playing cards or craps were considered to be interaction. Purely audience-type activities were not considered interaction. To qualify as interaction there had to be mutual participation in some common focus of interest which involved active participation by the subjects. Since, in this context, this unit was not a work group, this enabled us to code quite important forms of leisure interaction not limited to overt verbal behavior.

In Figure 11 is a printout from one observation period. Figure 12 is the matrix of counts from a full observation. The cells on the diagonal indicate the number of times the subject was coded as physically present in the compartment in some sampling area but not interacting. The off diagonal cells indicate the interaction counts with other subjects. Only the upper right triangular matrix is necessary because speaking is counted equivalent to listening. The count is symmetric. Marginals may be unequal for several reasons, not only due to normal sample variation, or observer error but due to subject absence from the compartment. For obvious reasons of privacy the head, shower and drying room areas were off limits to the observers. Subjects might also be on liberty during at least part of the period of observation (see "Liberty Sections" below) or occasionally, somewhere else in the barracks.

It was possible for the observer to code a subject twice during a 15 minute sampling period if the subject physically moved from one area to another during that time. The second coding indicated a different conversation from that coded in the first area, and therefore was at least partially independent of the first.

CLASSROOM SAMPLING

In every day life as one sequentially occupies one position after another - family member, student, employee, customer - one moves from one set of proxemic relations to another. Most studies have taken only one set of such relations

Figure 11:

An example of a data printout from three sampling periods in the compartment (see text under CODE for explanation)

	U0100 MMH,0000,				
	U0110 U40576,0019,				
1.	U0120 1532,0024,				
	U0130 CFH,0047,				
2.	U0140 1533,0061,		U0680 F18,2826,		
3.	U0150 95*115*21,	4.	U0690 C1,2898,	3.	U9540 CFH,3124,
4.	U0160 C2,1621,		U0700 H33,2902,		U9550 2102,3129,
	U0170 H22,1625,		U0710 1,2904,	4.	U9560 F17,3299,
	U0180 12,1629,		U0720 F14,2908,	4.	U9570 C3,3297,
4.	U0190 C7,1710,		U0730 H,2909,		U9580 F17,3299,
	U0200 F17,1710,		U0740 27,2912,	4.	U9590 C8,3367,
	U0210 C8,1700,	4.	U0750 C3,2983,		U9600 F100,3369,
	U0220 F29,1765,		U0760 F8,2986,	4.	U9610 C4,3449,
4.	U0230 C5,1861,		U0770 14,2989,		U9620 F13,3452,
	U0240 F11,1884,		U0780 24,2991,		U9630 H30,3455,
	U0250 14,1887,		U0790 22,2993,		U9640 5,3457,
	U0260 12,1869,		U0800 26,2995,	4.	U9670 C9,3629,
	U0270 H28,1894,		U0810 12,2999,		U9680 F9,3633,
	U0280 22,1897,	4.	U0820 C10,3165,		U9690 18,3636,
	U0290 33,1901,		U0830 F17,3168,	4.	U9700 F13,3640,
4.	U0300 C4,1964,		U0840 21,3171,	4.	U9710 C5,3713,
	U0310 H30,1969,		U0850 F9,3173,	4.	U9720 F32,3716,
	U0320 5,1971,		U0860 F16,3176,	4.	U9730 C7,3752,
	U0330 F12,1974,		U0870 31,3179,	4.	U9740 F23,3754,
	U0340 33,1977,		U0880 F7,3182,	4.	U9750 C2,3815,
4.	U0350 C1,2033,		U0890 10,3185,	4.	U9760 F3,3817,
	U0360 F20,2037,		U0900 H85,3190,	4.	U9770 C1,3864,
	U0370 24,2040,		U0910 32,3193,		U9780 H22,3869,
	U0380 23,2042,	4.	U0920 F25,3197,		U9790 2,3871,
	U0390 1,2044,		U0930 F4,3199,	4.	U9800 F17,3875,
4.	U0400 C9,2225,		U0940 C4,3281,		U9810 20,3878,
	U0410 H29,2229,	4.	U0950 F100,3283,		U9820 1,3879,
	U0420 27,2233,		U0960 C9,3389,	4.	U9830 C10,4029,
	U0430 F19,2236,	4.	U0970 H19,3393,		U9840 F15,4034,
	U0440 F13,2241,		U0980 27,3396,		U9850 H85,4040,
4.	U0450 C5,2330,		U0990 F29,3400,		U9860 29,4045,
	U0460 F31,2333,		U1000 F23,3404,		U9870 F85,4052,
4.	U0470 C5,2385,		U1010 H85,3407,		U9880 F8,4056,
	U0480 F14,2389,		U1020 11,3410,		U9890 F51,4063,
4.	U0490 C10,2619,	4.	U1030 C6,3512,	4.	U9900 F33,4068,
	U0500 F29,2624,		U1040 F100,3515,		U9910 24,4072,
	U0510 F17,2627,	4.	U1050 C5,3619,		U9920 19,4074,
	U0520 18,2630,		U1060 F12,3622,	4.	U9930 C8,4140,
	U0530 F8,2632,		U1070 10,3625,		U9940 F14,4142,
	U0540 2,2635,		U1080 30,3628,	5.	U9960 CCC,4182,
	U0550 F16,2641,		U1090 14,3632,	6.	U9970 2112,4184,
	U0560 9,2643,		U1100 18,3641,		10450 HFC,2652,
	U0570 H7,2647,	4.	U1110 C2,3707,		10460 33*2015,2664,
	U0580 10,2649,		U1120 F100,3709,		10470 26*2025,2675,
	U0590 21,2651,	4.	U1130 C6,3820,		10480 12*2015,2682,
	U0600 F4,2654,		U1140 F85,3823,	7.	10490 26*1959,2691,
	U0610 F25,2656,	4.	U1150 12,3826,		10500 HFC/HFC,2699,
5.	U0620 CCC,2659,		U1190 CCC,3888,		10510 ***,2703,
	U0630 1559,2665,	5.	U1200 1612,3892,		10520 040676,2710,
3.	U0640 CFH,2734,				10530 2240,2713,
	U0650 1600,2738,				10540 0000000
	U0660 95*116*21,				
4.	U0670 C7,2822,				

BEST AVAILABLE COPY

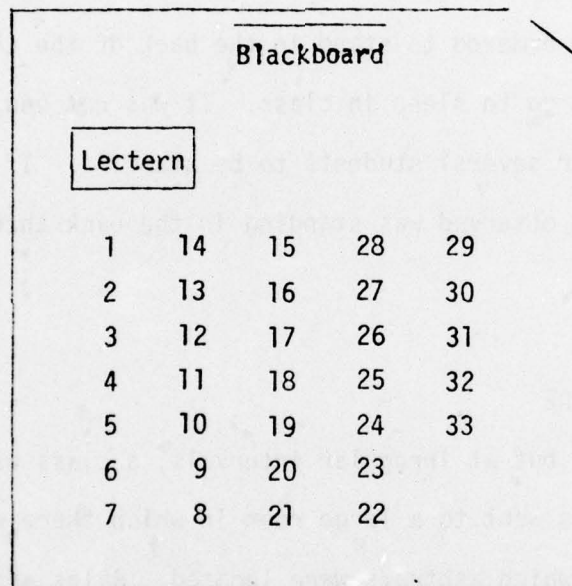
into account. For instance, Festinger, Schachter and Back took into account only the proxemic relations in the housing project. They did not take into account the relations the respondents' husbands had as students and employees during the many hours a day they were not at home. During apprenticeship training only two such sets of relations were fixed by assignment and occupied for considerable periods of time each day: compartment rack order and classroom desk order.

In designing the study, I considered whether these orderings should be different or as alike as possible. The effect of very different orderings was interesting. There is little attention in the literature to sequential, multiple orderings. This would have constituted an additional experimental condition impossible to accommodate within the overall design since it would have resulted in too many groups. I decided that while it was impossible to maintain classroom ordering equivalent to that in the compartment we should seek to maintain it as closely as possible. Since informal observation indicated covert interaction occurred mostly within rather than between rows the ordering was maintained within rows. A typical ordering is displayed in Figure 13. Rows were numbered from left to right (facing the front of the classroom). Subject 1 was assigned to the front seat in Row 1, subject 2 to the seat behind, and so on. The next row was filled from the back so that the next numbered subject after the last in Row 1 was directly across the aisle in Row 2.

Because our primary focus was on informal, spontaneous interaction it was necessary to observe apprentices very carefully in the classroom. For the most part their training consisted of six to eight hours of lectures per day. They were forbidden to talk to each other during lectures.

Figure 13:

A Typical Classroom Desk Ordering
by Subject Number



Communication was covert and fleeting, often whispered exchanges during lectures or tests. Since communication was covert it was difficult to code if the entire class was monitored at one time. Again, we decided to sample, this time by rows. The observer was given a randomly ordered list of rows (without replacement). He would observe the first row in the ordering for five minutes, then the next, and so on. When all the rows had been observed he would start again using another ordering. The observer sat at the back of the classroom to make these observations.

Students were ordered to stand in the back of the classroom if they thought they might go to sleep in class. It was not unusual at any particular time for several students to be standing. If a student(s) from the row being observed was standing in the back then he was included in that sample.

THE CLASSROOM LOUNGE

Periodically, but at irregular intervals, a class was given a smoke break. Apprentices went to a large room in which there were four high narrow benches on which ashtrays were located. Rules about throwing ashes or cigarettes on the floor were very strict so they clustered around these benches to talk and smoke. There were no chairs. During short breaks observers coded these clusters distinguishing separate conversation groups within them when such differentiation was possible. They did not code interaction but rather the apparent physical clustering of the subjects into interacting groups. Since this clustering was by choice, I thought it might give some clue to informal relations.

GALLEY SAMPLING

The base was served by a very large galley. Since nearly two thousand sailors had to be fed at each meal, scheduling was rigid. Each unit was scheduled for a particular time. The unit marched to the galley, went through a cafeteria line, and then individuals seated themselves at separate four to six man tables of their own choice. Their freedom to choose was constrained only by the number of tables already occupied.

Because I thought the sharing of meals might also provide a clue to the development of informal relations observers accompanied the unit to the galley. After it had passed through the cafeteria line and individuals had chosen their tables, the observer circulated through the galley coding the subjects by table. He made no attempt to count interactions but only to specify table groups.

LIBERTY SECTIONS

Each unit was divided into two liberty sections. On alternate days each section was on liberty from 1830 until 2130. The entire unit was together in the compartment from 1530 to 1830 and then only one section was in the compartment from 1830 to 2130, except for those who were eligible for liberty but chose to remain in the compartment. Some evenings, for disciplinary reasons, liberty was cancelled. While it was impossible to observe the group on liberty, it was possible to know which subjects they were and, to some extent, which subjects went together. Every subject who went on liberty was noted on a liberty list by the company clerk and his exact time of return was logged. This liberty list was copied at the end of each day. A fairly strong assumption was made that if subjects A, B and C all returned at exactly the same time that they had been on liberty together.

Interaction was not frequent in the classroom. From classroom lounge, galley and liberty list only grouping could be inferred. For none of these were frequencies large. Therefore, these data have only been used to supplement the interaction data gathered in the compartment.

MECHANICS OF DATA COLLECTION

This design proved practical for the Navy project. Subjects were contacted when they reported for training. The project was explained to them and their informed consent obtained. If they gave their consent to be part of the study, we sought their permission to take a polaroid snapshot. The photo was labelled with the subject's project identification number and name. This served as a learning tool for observers who also spent from one to three days with the subjects before the unit was officially formed and began training. In this way the 30 to 50 subjects were individually known to the observers by number and name before the formal data collection began.

During the approximately 14 days each unit was observed, data were recorded by both machine and pencil and paper techniques. A prototype, solid state Datamyte II was used during most of the study (see Figures 14-15). This device, which contains an automatic clock, has a keyboard which permits entry of numeric or alphabetic codes. These are stored, together with the times, in the device. At the end of the observation period the memory was emptied directly onto magnetic tape on a computer 400 miles away through an Execuport portable teletype terminal. The data were cleaned on line. The paper and pencil data were a backup in case of failure of the Datamyte II. Figure 11 is an example, not only of the observer code, but of Datamyte II output.



Figure 14: The Datamyte keyboard. It is rather like using a portable calculator.



Figure 15: This is the normal way to hold the Datamyte while in use. It is lightweight and easily grasped with one hand.

THE CODE

To record their data observers used an alphanumeric observational language which was recorded onto the memory of the Datamyte II through its keyboard. Various combinations of letters and numbers were the vocabulary of the observational language. Each "word" in the language was constrained by certain ordering rules, somewhat like a grammar, as well as by other conventions.

<u>Word</u>	<u>Meaning</u>	<u>Explanation</u>
HHH,	Beginning of observer work day.	The observations were made over several hours each day. Beginning and end of day codes marked the boundaries of the entire day's data.
1. mddy,	Date.	Month, day and year the data were collected.
ttt,	Time.	Datamyte II contains a clock counting in hundredths of minutes. To relate this to a regular 24-hour clock as used in the military it was necessary to enter the time of day by hand.
2. CFH,	Beginning of observation cycle.	In each sampling location (compartment, classroom, classroom lounge, galley) each complete cycle of samples had to be preceded by this indicator.
ttt,	Time	Time cycle began on 24-hour clock.
3. on*sn*ln,	Observer number, sample number, location number.	Every observer was given a number. Each sample cycle was assigned a unique number (1...n) from the start of each unit. Compartment, classroom, lounge and galley were also assigned numbers.
4. Cnn,	Sample area number within sample cycle.	For area numbers within the compartment see Figure 10.
Fnn,	Subject(s) physically present - no interaction.	Every subject physically present but not interacting in the sampling area was coded by number.

<u>Word</u>	<u>Meaning</u>	<u>Explanation</u>
Hnn, nn,	Subjects interacting.	One group of subjects interacting follows any H.
Hnn,nn,	Another group interacting	Often more than one group may be located in the same area. Every numbered subject within the group follows the H.
Cnn,	Sample area number.	The next area in the randomly ordered list of areas in the cycle.
etc.		
5. CCC,	End of observation cycle.	After the ten areas are sampled a cycle is complete.
tttt,	Time.	Time cycle ended on 24-hour clock.
6. HFC,	Liberty list.	After sailors return from liberty they sign in with the times of their return recorded beside their names. For clues to leisure companions this list was recorded by subject number and time.
nn*tttt,	Subject number and time returned from liberty.	
HFC, HFC,	End of liberty list.	
7. ***,	End of observer's work day.	This must appear at the finish of each day's data collection.
mmdyy,	Date.	
tttt,	Time.	

Specific Indicators

1n	Compartment - 21; Galley - 22; classroom - 23; classroom lounge - 24
F100	Sample area empty.
85	Non-unit apprentice (repeat as needed)
86	Navy person (not apprentice); i.e., unit counselor, instructor, etc.
99	Group as-a-whole.

Appendix 3.1

DETERMINING SAMPLE SIZE NECESSARY
TO DETECT GROUPS FROM
INTERACTION FREQUENCIES

by
Kinley Larntz, Ph.D.

INTRODUCTION

One objective of a study in which frequencies of dyad interactions are collected is to identify groups (i.e., cliques) that exist in the study sample. A group is defined here as a collection of individuals that interact at a rate higher than that usually expected for a random collection of individuals. Of course, the ability to identify groups will depend heavily upon how far above the norm the group interaction rate turns out to be, e.g., a group that interacts at ten times the normal rate is much easier to detect than one that interacts at ten percent above the normal rate.

In designing an experiment one of the principal questions a researcher must address is whether the study has a chance of accomplishing its objectives. If the identification of groups is an objective, then it is important to determine what constitutes an adequate sample size (i.e., how many observations are needed) for such identification. This report addresses the question of required sample size when groups interact at a rate of k ($k > 1$) times the normal rate.

NOTATION AND ASSUMPTIONS

The study sample consists of N individuals all of whom are free to interact with the others. The random pairing interaction model assumes that pairs of interacting individuals are formed at random but allows for differing propensities to interact ($\Pi_1, \Pi_2, \dots, \Pi_N$). For the pair (i, j) the relative frequency of interaction under the random pairing model is

$$P_{ij} = \frac{\Pi_i \Pi_j}{\sum_{i < j} \Pi_i \Pi_j} .$$

If n interactions are observed the expected number of (i, j) pairs is simply

$$\tilde{m}_{ij} = nP_{ij}$$

For details on estimation of the \tilde{m}_{ij} under the random pairing model, see Larntz and Weisberg (1976).

The data available for analysis will usually consist of the frequencies of pair (i, j) interaction which we will label X_{ij} . Thus $\sum_{i < j} X_{ij} = n$ is the total number of observations. Since observation usually occurs for a fixed length of time (rather than until n counts are recorded), a reasonable model for the counts X_{ij} is that they are independent random variables following a Poisson distribution, $X_{ij} \sim \text{Poisson}(m_{ij})$. The objective of identifying groups is translated into identifying pairs whose Poisson rate m_{ij} is higher than the rate under random pairing, namely \tilde{m}_{ij} .

A normal approximation for the distribution of X_{ij} is possible when m_{ij} is large. Small sample improvements have been proposed by many individuals, but one that has proven particularly useful is due to Freeman and Tukey (1950),

$$d_{ij} = \sqrt{X_{ij}} + \sqrt{X_{ij} + 1} - \sqrt{4m_{ij} + 1} \approx N(0, 1).$$

For the calculation of sample sizes this is replaced by the asymptotically

equivalent
$$2 \left[\sqrt{X_{ij}} - \sqrt{m_{ij}} \right] \approx N(0, 1).$$

However, in practice the group determinations are based on the Freeman-Tukey approximation.

Calculation of Required Sample Size

If there are N individuals in the study sample, there will be $\binom{N}{2}$ possible pairs, and thus when .05 level tests are conducted on each pair the overall error rate becomes much greater than .05. To adjust for this problem, a common procedure calls for making the individual tests at a reduced significance level,

$$\alpha^* = \frac{\alpha}{\binom{N}{2}}$$

The overall error is then bounded above by α . Thus, "significant" pairs will be identified by Freeman-Tukey deviates in excess of γ_{α^*} , where

γ_{α^*} is the one-tailed α^* critical value for the normal distribution.

Values of γ_{α^*} for N = 20, 40, 75 are given in the table below:

Adjusted Normal Critical Values

N	Number of Pairs	$\alpha = .10$	$\alpha = .10$	$\alpha = .10$
20	190	3.28	3.47	3.88
40	780	3.66	3.83	4.21
75	2775	3.97	4.13	4.49

To calculate the required sample size for a given situation, the alternative hypothesis and the power requirement must be specified. To illustrate the calculations we will use N = 40, $\alpha = .05$, required power of .90, and an alternative rate that is k = 3 times the random pairing rate. The form that the sample size requirement takes is calculation of \tilde{m}_{ij} . Thus, under the null hypothesis (random pairing) X_{ij} will have rate \tilde{m}_{ij} while under the alternative the

rate will be $k\tilde{m}_{ij}$. Using the normal approximation makes the power calculations easier since the requirement becomes one of finding \hat{m}_{ij} to satisfy

$$P_r \left[\sqrt{X_{ij}} \geq \sqrt{\tilde{m}_{ij}} + .5z_{\alpha^*} \mid E(X_{ij}) = k\tilde{m}_{ij} \right] = \text{Required Power}$$

If z_{1-p} is the normal one-tailed critical value for a test at level 1-p, then m_{ij} will be given by

$$\sqrt{k\tilde{m}_{ij}} - \sqrt{\tilde{m}_{ij}} - .5z_{\alpha^*} = .5z_{1-p}$$

or

$$\tilde{m}_{ij} = \left[\frac{.5(z_{1-p} + z_{\alpha^*})}{\sqrt{k} - 1} \right]^2$$

For the example case,

$$\tilde{m}_{ij} = \left[\frac{.5(1.28 + 3.83)}{\sqrt{3} - 1} \right]^2 = 12.18$$

Thus, if the random pairing model has expected value 12.18 in the (i,j) cell and the true rate of interaction is 36.54 we will have a 90% chance of detecting the increased rate at level $\alpha = .05$.

Tables of required \tilde{m}_{ij} are given for $N = 20, 40, 75$; $\alpha = .10, .05, .01$; $k = 2, 3, 4, 5, 7, 10$; and required power .50, .75, .90, .95, .99. The value computed above appears as the table value for $N = 40, \alpha = .05, k = 3, \text{ power} = .90$.

All of the computations are based on the normal approximation which is better for larger values of $\sqrt{k\tilde{m}_{ij}}$ than for small values. Thus, for very small values of $\sqrt{k\tilde{m}_{ij}}$, such as less than 5, the power requirement may not be attained exactly.

REFERENCES

Freeman, M.F. and Tukey, J.W. (1950).

Transformations related to the angular and square root.

Annals of Mathematical Statistics, 21, pp. 607-611.

Larntz, K. and Weisberg, S. (1976)

Multiplicative models for dyad interaction. To appear in

Journal of the American Statistical Association, 71, June 1976.

N = 20

Size of Test	Power					
	.50	.75	.90	.95	.99	
k = 2	.10	15.7	22.8	30.3	35.3	45.9
	.05	17.5	25.1	32.9	38.1	49.0
	.01	21.9	30.3	38.8	44.5	56.2

Size of Test	Power					
	.50	.75	.90	.95	.99	
k = 3	.10	5.02	7.32	9.70	11.3	14.7
	.05	5.62	8.03	10.5	12.2	15.7
	.01	7.02	9.70	12.4	14.2	18.0

Size of Test	Power					
	.50	.75	.90	.95	.99	
k = 4	.10	2.69	3.92	5.20	6.60	7.87
	.05	3.01	4.31	5.64	6.54	8.41
	.01	3.76	5.20	6.66	7.63	9.64

Size of Test	Power					
	.50	.75	.90	.95	.99	
k = 5	.10	1.76	2.57	3.40	3.97	5.15
	.05	1.97	2.82	3.69	4.28	5.50
	.01	2.46	3.40	4.36	4.99	6.31

Size of Test	Power					
	.50	.75	.90	.95	.99	
k = 7	.10	0.99	1.45	1.92	2.24	2.90
	.05	1.11	1.59	2.08	2.41	3.11
	.01	1.39	1.92	2.46	2.82	3.56

Size of Test	Power					
	.50	.75	.90	.95	.99	
k = 10	.10	0.58	0.84	1.11	1.30	1.68
	.05	0.64	0.92	1.21	1.40	1.80
	.01	0.80	1.11	1.42	1.63	2.06

NOTE: Tabled value is the cell expectation (\tilde{m}_{ij}) required under the random pairing model to detect a rate of $k\tilde{m}_{ij}$.

		<u>Power</u>				
N = 40	Size of <u>Test</u>	.50	.75	.90	.95	.99
k = 2	.10	19.5	27.4	35.6	41.0	52.3
	.05	21.4	29.6	38.0	43.7	55.3
	.01	25.8	34.8	43.9	50.0	62.3
k = 3		.50	.75	.90	.95	.99
	.10	6.25	8.79	11.4	13.1	16.7
	.05	6.84	9.49	12.2	14.0	17.7
	.01	8.27	11.2	14.1	16.0	20.0
k = 4		.50	.75	.90	.95	.99
	.10	3.34	4.71	6.10	7.04	8.97
	.05	3.67	5.09	6.53	7.49	9.49
	.01	4.43	5.98	7.54	8.57	10.7
k = 5		.50	.75	.90	.95	.99
	.10	2.19	3.08	3.99	4.60	5.87
	.05	2.40	3.33	4.27	4.90	6.21
	.01	2.90	3.91	4.93	5.61	7.00
k = 7		.50	.75	.90	.95	.99
	.10	1.24	1.74	2.25	2.60	3.31
	.05	1.35	1.88	2.41	2.77	3.50
	.01	1.64	2.21	2.78	3.16	3.95
k = 10		.50	.75	.90	.95	.99
	.10	0.72	1.01	1.30	1.50	1.92
	.05	0.78	1.09	1.40	1.60	2.03
	.01	0.95	1.28	1.61	1.83	2.29

NOTE: See Note on Page 7.

Power

N = 75		Size of Test	.50	.75	.90	.95	.99
k = 2	.10		23.0	31.5	40.2	45.9	57.8
	.05		24.9	33.7	42.6	48.6	60.8
	.01		29.4	38.9	48.5	54.8	67.8
k = 3			.50	.75	.90	.95	.99
	.10		7.4	10.1	12.9	14.7	18.5
	.05		8.0	10.8	13.7	15.6	19.5
	.01		9.4	12.5	15.5	17.6	21.7
k = 4			.50	.75	.90	.95	.99
	.10		3.9	5.4	6.9	7.9	9.9
	.05		4.3	5.8	7.3	8.3	10.4
	.01		5.0	6.7	8.3	9.4	11.6
k = 5			.50	.75	.90	.95	.99
	.10		2.6	3.5	4.5	5.2	6.5
	.05		2.8	3.8	4.8	5.5	6.8
	.01		3.3	4.4	5.4	6.2	7.6
k = 7			.50	.75	.90	.95	.99
	.10		1.45	2.0	2.5	2.9	3.7
	.05		1.57	2.1	2.7	3.1	3.9
	.01		1.86	2.5	3.1	3.5	4.3
k = 10			.50	.75	.90	.95	.99
	.10		0.84	1.16	1.47	1.69	2.1
	.05		0.91	1.24	1.56	1.78	2.2
	.01		1.08	1.43	1.78	2.0	2.5

NOTE: See Note on Page 7.

DISTRIBUTION LIST

MANDATORY

LIST 1

Office of Naval Research
(Code 452)
800 N. Quincy Street
Arlington, Virginia 22217

Director
U.S. Naval Research Laboratory
Washington, D.C. 20390
ATTN: Technical Information
Center

Defense Documentation Center
Building 5
Cameron Station
Alexandria, Virginia 22314

Library, Code 2029
U.S. Naval Research Laboratory
Washington, D.C. 20390

Science & Technology Division
Library of Congress
Washington, D.C. 20540

Psychologist
ONR Branch Office
495 Summer Street
Boston, Massachusetts 02210

LIST 2

Psychologist
ONR Branch Office
1030 E. Greet Street
Pasadena, California 91106

Research Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, Illinois 60605

PRINCIPAL INVESTIGATORS

LIST 3

Dr. Earl A. Alluisi
Old Dominion University
Res. Foundation
Norfolk, Virginia 23508

Dr. Judith Daly
Decisions & Designs, Inc.
Suite 100
8400 Westpark Drive
McLean, Virginia 22101

Dr. James A. Bayton
Department of Psychology
Howard University
Washington, D.C. 20001

Dr. H. Russell Bernard
Department of Sociology &
Anthropology
West Virginia University
Morgantown, West Virginia 26506

Dr. Arthur Blaiwes
Naval Training Equipment Center
Orlando, Florida 32813

Dr. Milton R. Blood
School of Business
Georgia Institute of Technology
Atlanta, Georgia 30332

Dr. David G. Bowers
Institute for Social Research
University of Michigan
Ann Arbor, Michigan 48106

Dr. John J. Collins
Vice President
Essex Corporation
6305 Caminito Estrellado
San Diego, California 92120

Dr. Harry R. Day
University City Science Center
Center for Social Development
3624 Science Center
Philadelphia, Pennsylvania 19104

Dr. C. Brooklyn Derr
Associate Professor, Code 55
Naval Post Graduate School
Monterey, California 93940

Dr. George T. Duncan
Carnegie-Mellon University
5000 Forbes Avenue
Pittsburgh, Pennsylvania 15213

Dr. Fred E. Fiedler
Department of Psychology
University of Washington
Seattle, Washington 98105

Dr. Samuel L. Gaertner
Department of Psychology
University of Delaware
2200 Wolf Hall
Newark, Delaware 19711

Dr. William E. Gaymon
Suite 200
1055 Thomas Jefferson St., NW
Washington, D.C. 20007

Dr. Paul S. Goodman
Graduate School of Industrial Admin.
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

Dr. Gloria L. Grace
System Development Corp.
2500 Colorado Avenue
Santa Monica, California 90406

Dr. J. Richard Hackman
Administrative Sciences
Yale University
56 Hillhouse Avenue
New Haven, Connecticut 06520

Dr. Leo A. Hazlewood
CACI, Inc.
1815 Fort Myer Dr.
Arlington, Virginia 22209

Dr. Edwin Hollander
Department of Psychology
State University of New York
at Buffalo
4230 Ridge Lea Rd.
Buffalo, New York 14226

Dr. Daniel F. Huck
General Research Corp.
Westgate Research Park
McLean, Virginia 22101

Dr. Charles L. Hulin
Department of Psychology
University of Illinois
Champaign, Illinois 61820

Dr. Rudi Klauss
Syracuse University
Public Administration Dept.
Maxwell School
Syracuse, New York 13210

Dr. Edward E. Lawler
Battelle Human Affairs Research
Center
4000 N.E. 41st Street
P.O. Box 5395
Seattle, Washington 98105

Dr. Arie Y. Lewin
Duke University
Duke Station
Durham, North Carolina 27706

Dr. Morgan W. McCall, Jr.
Center for Creative Leadership
5000 Laurinda Dr.
P.O. Box P-1
Greensboro, North Carolina 27402

Dr. Terence R. Mitchell
School of Business Administration
University of Washington
Seattle, Washington 98195

Dr. William H. Mobley
College of Business Administration
University of South Carolina
Columbia, South Carolina 29208

Dr. Thomas D. Morris
The Brookings Institution
1775 Massachusetts Avenue, NW
Washington, D.C. 20036

Dr. James P. Murphy
National Analysts
A Division of Booz-Allen &
Hamilton, Inc.
400 Market Street
Philadelphia, Pennsylvania 19106

Dr. Peter G. Nordlie
Human Sciences Research, Inc.
7710 Old Springhouse Rd.
McLean, Virginia 22101

Dr. Herbert R. Northrup
Industrial Research Unit
University of Pennsylvania
Philadelphia, Pennsylvania 19174

Dr. A. F. K. Organski
3068 Institute for Social Research
University of Michigan
Ann Arbor, Michigan 48104

Dr. Paul Pedersen
Society for Intercultural Education,
Training and Research
107 MIB, University of Pittsburgh
Pittsburgh, Pennsylvania 15260

Dr. Chester M. Pierce
Harvard University
Nichols House
Appian Way
Cambridge, Massachusetts 02138

Dr. Manuel Ramirez
Systems and Evaluations
232 Swanton Blvd.
Santa Cruz, California 95060

Dr. Irwin Sarason
Department of Psychology
University of Washington
Seattle, Washington 98195

Dr. John Ruhe
University of North Carolina
Dept. of Business Admin.
Charlotte, North Carolina 28223

Dr. Irwin Sarason
Department of Psychology
University of Washington
Seattle, Washington 98195

Dr. S. B. Sells
Texas Christian University
Fort Worth, Texas 76129

Dr. H. Wallace Sinaiko
Program Director
Manpower Research & Advisory Services
Smithsonian Institution
801 N. Pitt St. - Suite 120
Alexandria, Virginia 22314

Mrs. Alice I. Snyder
Mental Health Clinic
Naval Regional Medical Center
Pearl Harbor
FPO San Francisco 96610

Dr. Paul Spector
American Institutes for Research
Foxhall Square
3301 New Mexico Avenue N.W.
Washington, D.C. 20016

Dr. Richard Steers
Graduate School of Management &
Business
University of Oregon
Eugene, Oregon 97403

Dr. Lorand B. Szalay
American Institutes for Research
Foxhall Square
3301 New Mexico Avenue N.W.
Washington, D.C. 20016

Dr. Victor H. Vroom
School of Organizational Management
Yale University
56 Hillhouse Avenue
New Haven, Connecticut 06520

Dr. Paul Wall
Division of Behavioral Science
Tuskegee Institute
Tuskegee, Alabama 36088

Dr. Abraham R. Wagner
Analytical Assessments Corp.
357 South Robertson Blvd.
Beverly Hills, California 90211

Dr. J. Wilkenfeld
Department of Government &
Politics
College Park, Maryland 20742

Dr. Davis B. Bobrow
University of Maryland
Department of Government and Politics
College Park, Maryland 20742

MISCELLANEOUS

LIST 4

Military Assistant for
Human Resources
OAD (E & LS) ODDR&E
Pentagon 3D129
Washington, D.C. 20301

Air University Library/
LSE-8110
Maxwell AFB, Alabama 36112

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine
Corps (Code RD-1)
Washington, D.C. 20380

LCdr. C. A. Patin, U.S.N.
Director, Human Goals Dept.
Code 70, Naval Training Ctr.
Orlando, Florida 32813

Navy Personnel R & D Center
Code 10
San Diego, California 92152

Human Resources Management
Center Attachment
Naval Support Activity
c/o FPO New York, NY 09521
ATTN: TDC Nelson

Journal Supplement Abstract
Service
APA
1200 - 17th Street N.W.
Washington, D.C. 20036

HQ, USAF
AFMPC/DPMYAR
Randolph AFB, Texas 78148

Army Research Institute
Commonwealth Bldg.
1300 Wilson Blvd.
Rosslyn, Virginia 22209

Chief of Naval Personnel
Assistant for Research
Liason (Pers-Or)
Washington, D.C. 20370

Training Officer
Human Resource Management Ctr.
NTC, San Diego, California 92133

Office in Charge (Code L5)
Naval Aerospace Medical
Research Lab.
Naval Aerospace Medical Center
Pensacola, Florida 32512

HumRRO (ATTN: library)
300 N. Washington Street
Alexandria, Virginia 22314

Division Director for Social
Science
National Science Foundation
1800 G Street N.W.
Washington, D.C. 20550

AFOSR (NL)
1400 Wilson Blvd.
Arlington, Virginia 22209

Chief, Psychological
Research Branch
U.S. Coast Guard (C-P-1/62)
400 - 7th Street S.W.
Washington, D.C. 20590

Bureau of Naval Personnel
(Pers-6)
Assistant Chief of Naval
Personnel for Human Goals
Washington, D.C. 20370

Scientific Director
Naval Health Research Center
San Diego, California 92152

Capt. Bruce G. Stone, USN
(Code N-33)
Director, Educ. & Training
Research & Prog. Development
CNET Staff
Naval Air Station
Pensacola, Florida 32508

Director of Research
HumRRO Division #4 (Infantry)
P.O. Box 2086
Fort Benning, Georgia 31905

Mr. Luigi Petruzzo
2431 N. Edgewood Street
Arlington, Virginia 22207

Navy Materiel Command
Employee Development Office
Code SA-65
Room 150 Jefferson Plaza
Bldg. #2, 1429 Jeff Davis Highway
Arlington, Virginia 20360

Personnel Research and Development
Center
United States Civil Service
Commission
Bureau of Policies and Standards
Washington, D.C. 20415

Department of the Air Force
Air Force Institute of Technology (AU)
AFIT/SLGR (LT Col Umstot)
Wright-Patterson Air Force Base,
Ohio 45433