A MARKOV CHAIN ANALYSIS OF COLLEGE STUDENT INTERACTION *

by

KENNETH KAMMEYER
University of Kansas

CHARLES D. BOLTON
Portland State College

The analysis in this paper utilizes a stochastic model to examine the sequential pattern of topics discussed in college student bull sessions. The sequence of topics discussed by college students in their bull sessions is viewed as a process, wherein the topic discussed at one point in time influences the topic(s) discussed later. To put the matter somewhat more formally, the topic discussed at time t in a bull session is viewed as a probability function of the topic discussed at t-1 (or perhaps t-2 and t-1). The stochastic model to be used is a discrete-time Markov chain.[1] (Whether the first-order or second-order chain is more useful for describing the basic mechanism of the process will be determined by examining the data.)

The Markov chain model is a potentially useful mode of analysis whenever the phenomenon under investigation changes from one state to another over time, and there is reason to believe that the state at a given point in time is a function of a prior state (or states). With regard to the applicability of the model to the analysis of bull sessions, it seems reasonable to consider the successive topics of conversation among interacting individuals as states of that interaction. People customarily describe an interaction episode in terms of the sequence of topics discussed from the beginning to the end of the encounter. Our observations during the data gathering stage of this research assured us that college students could describe the content of their bull sessions in a topic to topic manner. From casual observations of the encounters among people we may also surmise that the topics of their conversation have some pattern or sequential form. We speak for example of "ice-breaking" topics: the topics of conversation that seem somehow easiest to discuss when individuals first meet. We also use the term "social amenities" when referring to culturally prescribed exchanges which have the promise of leading to more intense interaction.

The significance of the sequence of topics in college student bull sessions is both theoretical and substantive. From a theoretical point of view, social interaction, as a concept, is central to sociology and social psychology, and yet there has not been an overwhelming amount of study on the nature and content of interaction as it occurs in a natural setting. To be sure the interaction analysis system developed by Bales and his associates has been fairly widely used, and has had some considerable theoretical significance, but the interaction that is observed is almost always under more or less contrived conditions rather than in a natural setting.[2] The small group experiments that require subjects to communicate by means of written notes are even more unnatural, and even less useful as a means of studying the content of the process of interaction.

Some years ago there was a bit of interest in the content of conversations occurring in a natural setting, when several researchers employed an "eavesdropping" method of gathering data. In 1922 H. T. Moore collected segments of conversation overheard in the evening hours on Broadway.[3] Moore was not interested in the sequence of conversation topics, but rather with finding supportive evidence for the hypothesis that there are sex differences in conversations. Moore's study apparently stimulated M. H. Landis and H. E. Burtt to carry out similar research in a different and more varied locale,[4] and another researcher Carney Landis to make a cross-societal comparison of conversational content by conducting a similar study in England.[5] These studies, all of which were carried out in the 1920's, had as their main focus the sex differences in conversations, but did not examine the sequence of topics. Late in the 1920's Stuart M. Stoke and Elmer D. West conducted a study of college student bull sessions in which they gathered fairly extensive and detailed data on the topics discussed among college students.[6] Again they were primarily interested in making a descriptive statement and noting the sex differences in topics discussed.

College student bull sessions are again the point of interest in the present research. We undertook a study of college student bull sessions because we believed that what happens in the nonacademic side of student life is just as worthy of systematic study as the more frequently studied academic side. The bull session per se was taken as a focus of the study because it has so often been cited as a crucial aspect or feature of the college experience. For the student it perhaps serves two primary functions: it is a useful, if not crucial, augmentation to his intellectual development, and it is an aid in his psychological or identity development. From the standpoint of the social researcher the study of the student bull session has yet another function. The content of the informal student bull session is perhaps the best indicator of the student's true interests and values, and thus a good indicator of the effectiveness of a college education.

The internal analysis of the content of student bull sessions, is facilitated by a Markov chain analysis and promises to provide some answers to relevant questions about this form of interaction. For example, how much, if at all, does the first topic of a bull session affect the next and subsequent topics? If may be of interest to know what initial topic will be most likely to lead to an intellectual bull session.[7] These and a number of other substantive questions about the sequential patterns of bull session topics will be answered in the analysis below, but first a word about the manner in which the data were acquired and organized.

The Data

For this analysis we had information on 1,583 college student bull sessions which occurred on the Davis campus of the University of California during 1961 and 1962. We had data on the topical content (including the order of the topics), the sex composition and size of each bull session group.

The basic data were gathered by two quite different methods. Somewhat over a thousand of the bull session reports came from students in introductory sociology classes who acted as participant observers in their own naturally occurring bull sessions. The remaining data on 558 bull sessions came from time budget interviews with a sample of 210 undergraduate students. The two methods of data gathering yielded very different quantities of sessions, (the time budget technique revealing many more per week), but the general topic content did not differ greatly.[8]

For the participant observers and interviewers a bull session was defined as follows:

> a conversation between two or more persons which takes place in a relatively stationary situation. It (1) is informal and spontaneous, (2) lasts for fifteen minutes or longer, (3) has more intensity than a casual conversation, (4) involves one or more topics on which there is some difference of opinion, or feeling of a need for clarification, or about which a mutual sharing of experiences and ideas takes place.

In order to be counted as a bull session, at least one undergraduate student had to be a participant. Also, interaction between a single dating couple was not considered a bull session.

A rather elaborate 17 category content analysis system was developed to code the more than 4,000 separate topics. However, for this paper it will be necessary only to describe the four major headings under which these 17 were organized. There are first the "Personal" topics which include any discussion having a basically personal reference to one or more of the individuals participating in the session or someone known to the participants personally. For example, one of the major subcategories among the personal topics was "Boy-girl relations." This sub-category caught up all of the many discussions of dating, going steady, pinnings, etc. Other personal topics included "Personal relations with others," e.g. relations with roommates, parents, housemothers, etc., and "Short" and "Long range orientations and decisions," e.g. plans and preferences about clothing, hair styles (short range), occupations, careers, and marriage (long range).

The second major category is the "Campus " topic. Campus topics include the familiar talk about grades, courses, classes, athletic events, campus events and campus institutions.

The third major category we have labeled "Intellectual" topics. Intellectual topics include discussions of political, economic, religious or philosophical issues; of societal or psychological behavior and problems (if not on a personal level); of the arts, and of scientific or technical matters.

The fourth and final category is "Popular Culture." Topics classified as Popular Culture are primarily those related to national sports events (the baseball world series, a heavyweight boxing match), popular commercial movies, or movie and recording stars. The category "Popular Culture" is used in a somewhat restricted manner in that we have limited it to manifestations of what might be called "fandom."

In the following analysis it will be convenient to refer to these four major categories in an abbreviated form of notation. The following will be used: P = Personal, C = Campus, I = Intellectual and P.C. = Popular Culture.

The Analysis[9]

The first question to be answered in this analysis is whether or not the topics in student bull sessions are independent of one another or are influenced by the immediately preceding topic. In the data at hand we can ask if topic 2 is independent of topic 1, if topic 3 is independent of topic 2, if topic 4 is independent of topic 3, etc. Since five is the maximum number of topics recorded for any of the bull sessions, we have four matrices of transitional probabilities (topic 1 to topic 2, topic 2 to topic 3, etc.). The most efficient way to answer the first question would be to combine the four matrices of transition probabilities, but this would not be realistic until it has first been determined that the transition probabilities are similar enough that we can assume the same basic mechanism is operating throughout the entire process. We must first test the hypothesis that the transition probability matrices are constant, i.e., the transition probabilities from topic 1 to topic 2, do not differ significantly from the transition probabilities from topic 2 to topic 3, etc.[10]

The transitional probabilities are combined in the four matrices of Table 1, in order to test the hypothesis of constancy. None of the calculated x^2 values in Table 1 is large enough to reject the null hypothesis of constant transitional probabilities. The combined x^2 value for all four matrices is 26.41, making the probability of observing these differences under the assumptions of the null hypothesis equal to .85. The combined matrix now shown as Table 2 will be used to test the hypothesis that the topic discussed at time t is related to the topic discussed at time t-1.

Even casual inspection of the transitional probabilities shown in Table 2 reveals that the topics discussed at time t are not independent of the topics discussed at t-1. One of the most striking examples may

Table 1. Transition Probabilities - First-Order Chain-Test for Constancy

	Topic (t-1)	Personal	Campus	Intellectual	Popular Culture	Total	
(A)	P	.734 (520)	.140 (99)	.106 (75)	.020 (14)	1.000 (708)	$x^2 = 7.47$
	P	.721 (339)	.168 (79)	.083 (39)	.028 (13)	1.000 (470)	df = 9
	P	.695 (148)	.169 (36)	.108 (23)	.028 (6)	1.000 (213)	$P < .70$
	P	.638 (44)	.217 (15)	.116 (8)	.029 (2)	1.000 (69)	$P > .50$
(B)	C	.479 (147)	.401 (123)	.091 (28)	.029 (9)	1.000 (307)	$x^2 = 9.05$
	C	.500 (79)	.380 (60)	.076 (12)	.044 (7)	1.000 (158)	df = 9
	C	.488 (39)	.450 (36)	.050 (4)	.012 (1)	1.000 (80)	$P < .50$
	C	.567 (17)	.267 (8)	.167 (5)	.000 (0)	1.001 (30)	$P > .30$
(C)	I	.233 (47)	.084 (17)	.644 (130)	.040 (8)	1.001 (202)	$x^2 = 5.64$
	I	.287 (45)	.102 (16)	.586 (92)	.025 (4)	1.000 (157)	df = 9
	I	.267 (16)	.100 (6)	.633 (38)	.000 (0)	1.000 (60)	$P < .80$
	I	.333 (6)	.056 (1)	.556 (10)	.056 (1)	1.001 (18)	$P > .70$
(D)	PC	.415 (17)	.146 (6)	.146 (6)	.293 (12)	1.000 (41)	$x^2 = 4.25$
	PC	.424 (14)	.152 (5)	.121 (4)	.303 (10)	1.000 (33)	df = 9
	PC	.526 (10)	.211 (4)	.105 (2)	.158 (3)	1.000 (19)	$P < .90$
	PC	.750 (3)	.000 (0)	.000 (0)	.250 (1)	1.000 (4)	$P > .80$

Note: column heads "Personal" through "Total" are under the spanning heading "Topic t".

TOTAL: $x^2 = 26.41$ df = 36 P = .85

Table 2. Combined Transition Probabilities - First-Order Chain

Topic t-1	Topic t				
	P	C	I	PC	Total
P	.720 (1051)	.157 (299)	.099 (145)	.024 (35)	1.000 (1460)
C	.490 (282)	.395 (227)	.085 (49)	.030 (17)	1.000 (575)
I	.261 (114)	.091 (40)	.618 (270)	.030 (13)	1.000 (437)
PC	.454 (44)	.155 (15)	.124 (12)	.268 (26)	1.001 (97)
Total	(1491)	(511)	(476)	(91)	(2569)

be seen by noting that if the topic at t-1 is a Personal topic the probability that the next topic will also be a Personal topic is .720. By contrast, if the topic at t-1 is an Intellectual topic, the probability of the topic at t being Personal is only .261.[11] The calculated x^2 value for the differences between the observed frequencies in Table 2 and expected frequencies under the null hypothesis of independence is 904.2, which easily allows us to reject the null hypothesis. Perhaps a more relevant test would be some measure of our ability to predict the topic at t on the basis of our knowledge of the t-1 topic. Although the above calculated transition probabilities would obviously be our most reasonable predictions if we were to conduct some future independent trial, it is customary for post factum analyses such as this to use a measure of association based on some notion of prediction. The measure of association, lambda, is a familiar technique for describing the improvement in predictive power.[12] The lambda value for Table 2 is .145 which may be interpreted to mean that we can improve our prediction of the topic at t by about 15% if we know the topic at t-1. Surely not an impressive improvement, but in large part this is due to the nature of the measure we have used, since lambda is based on the notion that the the prediction of the t topic is maximized by predicting the maximum t column value for each category of t-1. The value of lambda is relatively low because both Campus and Popular Culture topics are followed most often by Personal topics (cells C → P, n = 282, and PC → P, n = 44). The fact that campus topics are followed by campus topics, and popular culture topics are followed by popular culture topics to a degree far greater than expected is not given any weight by the measure lambda.[13]

To this point we have assumed that the Markov chain is a first-order model and under this assumption have shown that the transition probabilities are constant. By combining the matrices of transition probabilities for a first-order chain we have shown that this would be a useful model for predicting the topics at time t on the basis of topics discussed at t-1.

It may be, however, that considerable advantage could be gained by using a second-order model of the Markov chain. Using this model one would predict (on the basis of transitional probabilities) the topics discussed at t by knowing the topics discussed at t-1 <u>and</u> t-2. One would ask of the model, "Will it help to improve our prediction of topic 3, if we know the sequence of the two preceding topics rather than just the one immediately preceding topic?" If the answer is yes, the second-order model will generally be more appropriate, although it has the obvious disadvantage of not being able to predict topic 2.

Again before examining the data to determine the appropriateness of the second-order model we must analyze the several transitional probability matrices of the second-order model to see if they are constant probabilities. This analysis involves sixteen 3 X 4 matrices similar in form to matrices A, B, C, and D of Table 1. One of the 16 matrices is shown in Table 3 to give the general form. Only 11 of the

Table 3. Transition Probabilities - Second-Order Chain-Test for Constancy (one example)

Topic t-2	Topic t-1	Topic t: P	C	I	PC	Total
P	P	.721 (238)	.179 (59)	.076 (25)	.024 (8)	1.000 (330)
P	P	.736 (106)	.167 (24)	.069 (10)	.028 (4)	1.000 (144)
P	P	.708 (34)	.188 (9)	.083 (4)	.021 (1)	1.000 (48)

$x^2 = .34$ df = 6 $P > .99$

16 matrices had sufficient cases that the x^2 value could be meaningfully calculated, but of those eleven the total x^2 value was 53.59 with combined degrees of freedom of 66 (11 X 6). The probability of this value of x^2 under the null hypothesis of independence is .86, so we cannot reject the null hypothesis. We may therefore assume a constancy of the second-order transition probabilities, and as before, this allows us to combine the matrices. In this case the 16 matrices may be reduced to four. They are shown in Table 4.

Table 4. Combined Transition Probabilities - Second-Order Chain

	Topic t-2	Topic t-1	Topic t P	C	I	PC	Total
(A)	P	P	[368.6] 378	[90.2] 92	[48.6] 39	[14.6] 13	522
	C	P	[112.3] 109	[27.5] 31	[14.8] 13	[4.4] 6	159
	I	P	[36.7] 31	[9.0] 5	[4.8] 15	[1.5] 1	52
	PC	P	[13.4] 13	[3.3] 2	[1.8] 3	[0.5] 1	19
			$x^2 = 30.05$#	df = 9	P < .001		
(B)	P	C	*[59.9] 70	[46.2] 36	[9.3] 10	[3.6] 3	119
	C	C	*[67.5] 51	*[49.7] 63	[10.0] 10	[3.8] 4	128
	I	C	* [7.6] 12	[5.8] 3	[1.2] 0	[0.4] 0	15
	PC	C	[3.0] 2	[2.3] 2	[0.5] 1	[0.2] 1	6
			$x^2 = 20.07$#	df = 9	P < .05		
(C)	P	I	*[19.7] 33	[6.8] 9	[41.1] 25	[1.5] 2	69
	C	I	[8.0] 8	[2.7] 3	[16.7] 16	[0.6] 1	28
	I	I	*[37.3] 25	[12.8] 10	[78.0] 94	[2.8] 2	131
	PC	I	[2.0] 1	[0.7] 1	[4.2] 5	[0.1] 0	7
			$x^2 = 25.56$#	df = 9	P < .01		
(D)	P	PC	[9.2] 12	[3.1] 2	[2.4] 2	[4.4] 3	19
	C	PC	[5.8] 6	[1.9] 3	[1.5] 1	[2.8] 2	12
	I	PC	[4.3] 3	[1.4] 3	[1.1] 2	[2.1] 1	9
	PC	PC	[7.7] 6	[2.6] 1	[2.0] 2	*[3.7] 7	16
			$x^2 = 10.84$	df = 9	P < .30		
	Combined:		$x^2 = 86.52$	df = 36	P < .001		

*The cells marked in this fashion contributed most to the x^2 values.
#The number of cases is below the level considered acceptable for the x^2.

In Table 4 the expected frequencies under the conditions of the null hypothesis are shown in brackets above the observed frequencies. The expected frequencies are shown in this case because they enter into the following discussion. The total calculated x^2 value for the four matrices in Table 4 is 86.52 (with 36 degrees of freedom) which is significant at the .001 level. The highly significant x^2 level indicates that the topic at t-2 (in combination with the topic at t-1) does influence the topic at t and thus the second-order Markov chain would be an appropriate model. However, careful inspection of the expected and observed frequencies in Table 4 reveals that only a small number of sequences contribute to the size of the computed x^2 value, and further that some of these same sequences are rarely occurring events. For example, consider the following sequence of topics: t-2 = Intellectual, t-1 = Personal, t = Intellectual in matrix A of Table 4. There are 15 cases of this sequence occurring out of a total of 752 three topic sequences in which a Personal topic is the second of three topics. The expected frequency for that cell is for only five cases. As a result, over two-thirds of the total x^2 value for matrix A is produced by this single deviation of observed from expected frequencies (21.68 out of 30.05). In this particular sequence of topics, an Intellectual topic preceding at t-2, a Personal topic at t-1 greatly increases the probability that the topic at t will be Intellectual again. This is a clear manifestation of a three stage sequence or a second-order chain, and it would be relatively easy to advance an <u>ad hoc</u> interpretation or explanation for this sequential pattern of bull session topics. Nevertheless, considering that this particular sequence represents such a small proportion of three sequence patterns in which the t-1 topic is Personal, it is obvious that it adds little to our general understanding of the flow of conversation topics.

In matrices B, C, and D of Table 4 it may further be observed that much of the difference between expected and observed frequencies is contributed by the three homogeneous sequences: C-C-C, I-I-I and PC-PC-PC. We see in these sequences the not surprising fact that two topics in a general category tend to lead to a third in the same category.

In summary, with regard to the appropriateness of the second-order model, we feel that even though the statistical test suggests that the second-order Markov chain is an appropriate model, a closer inspection of the data reveals that only in a relatively small number of sequences is the t-2 topic an important determinant of the t topic. In Table 4 cells marked with an asterisk indicate where the observed frequencies differ most from expected frequencies so that the reader may examine the most important three topic sequences. In Table 5 we will present the matrices of transition probabilities for the second-order chain, along with the first-order transition probabilities (repeated from Table 2 for easier comparison). Again in Table 5 we will mark with an asterisk those cells which appear to have the most significant variations from the average. Several of these will be referred to at the end of the paper.

Table 5. Matrices of First and Second Order Transition Probabilities

	Topic t-2	Topic t-1	Topic t P	C	I	PC
(A)	P	P	.724	.176	.075	.025
	C	P	.686	.195	.082	.038
	I	P	.596	*.096	*.288	.019
	PC	P	.684	.105	.158	.053
	First Order	P	*.720	.157	.099	.024
(B)	P	C	*.588	.303	.084	.025
	C	C	*.398	*.492	.078	.031
	I	C	*.800	.200	.000	.000
	PC	C	.333	.333	.167	.167
	First Order	C	.490	*.395	.085	.030
(C)	P	I	*.478	.130	.362	.029
	C	I	.286	.107	.571	.036
	I	I	*.191	.076	*.718	.015
	PC	I	.143	.143	.714	.000
	First Order	I	.261	.091	*.618	.030
(D)	P	PC	.737	.105	.105	.158
	C	PC	.500	.250	.083	.167
	I	PC	.333	.333	*.222	.111
	PC	PC	.375	.062	.125	*.438
	First Order	PC	.454	.155	.124	*.268

*The cells marked in this fashion have the most significant variations from the average.

Although the second-order transition probabilities are presented in Table 5 it is our conclusion that the description or prediction of topical sequences in student bull sessions is served nearly as well by the first-order model. One advantage of the first-order model is that the number of cases remains larger in the matrix cells, which in turn allows us to examine the transition probabilities for various strata in the population. This is the objective in the following section.

Transition Probabilities by Sex Composition and Group Size

As Goodman points out, it may be that the population under study is heterogeneous with respect to some characteristic and the transition probabilities that may be valid for one stratum or segment of the population may not be valid for the other strata or segments.[14] Only by separating the different strata or segments of a population, and engaging in separate analysis can one determine whether or not the same transition probabilities are relevant for all strata. In the case of college bull sessions, one of the most significant characteristics of a session is its sex composition. The bull session may occur among only females, or males, or in mixed sex groups (both males and females). The transition probabilities were examined for the three categories: all female, all male and mixed sex groups. As implied above, a first-order model was used in the following analysis. The method of analysis follows Goodman's presentation and again utilizes the tests he suggests.[15] If the tests of significance indicate that the transition probabilities differ by the sex composition of bull sessions, then we may conclude that there is a difference in the process of interaction for these categories. Table 6 shows the four matrices of transition probabilities and the results of the relevant statistical tests.

Three of the four matrices in Table 6 have differences in transitional probabilities that are statistically significant. In the case where the t-1 topic is Intellectual there is no apparent difference among the sex composition categories. The combined x^2 value for all four matrices is 54.03. With a total of 24 degrees of freedom this value of x^2 is significant at the .001 level of significance. We may therefore conclude that the sex composition of a college bull session group influences the topical flow of such sessions. We will make additional comments about the substantive nature of these differences in the concluding section of this paper.

We examined one other compositional characteristic of college student bull sessions in the same manner as just described for sex composition. In this analysis we looked at the size of the bull session group to determine if the number of people participating might influence the transitional probabilities. The sessions ranged in size from two to seven participants and the average size was between three and four. The null hypothesis was: The transition probabilities will not differ

Table 6. Transition Probabilities - First-Order Chain - Stratified by Sex

Topic t-1	Topic t P	C	I	PC	Total
(A)					
Males Personal	.674 (180)	.165 (44)	.112 (30)	.049 (13)	(267)
Females Personal	.734 (700)	.161 (153)	.087 (83)	.081 (17)	(953)
Mixed Personal	.707 (169)	.138 (33)	.134 (32)	.021 (5)	(239)

$x^2 = 14.92$ df = 6 $P < .05$

Topic t-1	P	C	I	PC	Total
(B)					
Males Campus	.481 (62)	.364 (47)	.085 (11)	.070 (9)	(129)
Females Campus	.529 (183)	.382 (132)	.075 (26)	.014 (5)	(346)
Mixed Campus	.370 (37)	.480 (48)	.120 (12)	.030 (3)	(100)

$x^2 = 17.91$ df = 6 $P < .01$

Topic t-1	P	C	I	PC	Total
(C)					
Males Intellectual	.241 (19)	.101 (8)	.595 (47)	.063 (5)	(79)
Females Intellectual	.258 (66)	.102 (26)	.617 (158)	.023 (6)	(256)
Mixed Intellectual	.284 (29)	.059 (6)	.637 (65)	.020 (2)	(102)

$x^2 = 5.58$ df = 6 $P < .50$

Topic t-1	P	C	I	PC	Total
(D)					
Males Popular Culture	.348 (16)	.217 (10)	.087 (4)	.348 (16)	(46)
Females Popular Culture	.639 (23)	.111 (4)	.111 (4)	.139 (5)	(36)
Mixed Popular Culture	.357 (5)	.000 (0)	.286 (4)	.357 (5)	(14)

$x^2 = 15.62$ df = 6 $P < .02$

by the number of people in the group. The x^2 values were not sufficiently large to reject the null hypothesis and we must therefore conclude that the number of people in the group does not influence the transitional probabilities.

Summary and Conclusions

The summarizing and concluding remarks of this paper will center around two general issues: 1) the findings relative to the Markov chain model and its general applicability to this kind of datum, and 2) the substantive information about college student bull sessions produced by this mode of analysis. With regard to the findings specific to the model itself:

1. Assuming a first-order chain, the transition probabilities from topic 1 to topic 2, topic 2 to topic 3, etc. were found to be constant (Table 1).
2. The first-order chain revealed the non-random character of bull session topics, and thus the applicability of this form of sequential analysis (Table 2).
3. Assuming a second-order chain the transition probabilities were found to be constant (Table 3 and subsequent discussion).
4. The second-order chain appeared to be an appropriate model on the basis of statistical tests (Table 4), but closer analysis revealed that it had only a limited usefulness.
5. An analysis of transitional probabilities by the sex composition of the group revealed some differences in the sequential pattern (Table 6).

Substantive findings about the sequential nature of bull session topics have not thus far received a great deal of attention, but of course that is the raison d'etre of the entire analysis, so let us see what the foregoing analysis may reveal about this form of social interaction.

1) At a general level we see that the topics students talk about in their bull sessions are not random, but the topic discussed at one point is influenced by those that have preceded it (Tables 2 and 5).

a) If the topic at t-1 is a Personal topic the probability that the next topic will also be a Personal topic is much greater than if any of the other topic categories had preceded it. If t-1 is Personal, the transition probability to another Personal is .720; if Campus it is .490; if Popular Culture it is .454; if Intellectual it is .261 (Table 2).

b) An Intellectual topic at t-1 greatly increases the probability of the topic at t also being Intellectual - (.618 compared to .099, .085 and .124 for Personal, Campus and Popular Culture at t-1 respectively) (Table 2).

c) Both Campus topics and Popular Culture topics at t-1 have a higher probability of leading to a Personal topic at t than to themselves again (Table 2). This probably shows that Campus and Popular Culture topics serve as ice-breakers leading to more personally oriented discussions.

d) Despite the point made in (c) above, the discussion of any topic category at t-1 greatly increases the probability that the topic will be discussed at time t. Note that for Popular Culture this is particularly true. For any other topic category at t-1 the probability of the next topic being Popular Culture is about .03, while Popular Culture itself has a probability of .268 of being followed by another Popular Culture topic. (Table 2).

e) The sequences I → P, and I → PC (at t-2 and t-1 respectively) leads to an above average probability of an Intellectual topic at t. But I → C has no corresponding effect (matrices A and D compared to matrix B - Column Intellectual - Table 5). It would appear from this that turning to a Campus topic has the effect of diminishing the probability of returning to an Intellectual topic.

2) The transitional probabilities differ by the sex composition of the bull session group (Table 6).

a) All-male bull session groups are more likely to go from Personal topics to Popular Culture topics than either all-female or mixed-sex groups (Table 6, matrix A).

b) All-female groups are more likely to go from Campus to Personal topics than mixed-sex groups; while mixed-sex groups are more likely to go from Campus to Intellectual topics than all-female groups (Table 6, matrix B).

c) Mixed-sex groups are less likely to go from Intellectual to Campus topics than either all male or all female groups (Table 6, matrix C).

d) All-male groups are more likely to go from Intellectual topics to Popular Culture topics than either all-female or mixed-sex groups (Table 6, matrix C).

e) All-male groups are more likely to go from Popular Culture to Campus topics more than either all-female or mixed-sex groups.

f) All-female groups are more likely to go from Popular Culture topics to Personal topics. The all-female groups are at the same time much less likely than the all-male or mixed-sex groups to go from one Popular Culture topic to another (Table 6, matrix C). These two factors in combination suggest that for college girls discussions of Popular Culture serve as "lead-ins" to matters of more personal concern.

g) For mixed-sex groups Popular Culture topics lead to Intellectual topics more than for all-male and all-female groups (Table 6, matrix D - Note, however that this difference is based on a very small number of cases).

The application of a Markov Chain Model to the data on college student interaction, has allowed us to examine this form of social behavior carefully and systematically. Though we have noted a fairly large number of empirical regularities which are useful for describing or perhaps even predicting certain kinds of behavior, it is also true that no overarching or generalizing principle has emanated from this analysis. This is not particularly a criticism of the technique for the model cannot produce dramatic patterns which are not in the empirical data. Bull sessions are interactions which are contributed to by two or more, and sometimes as many as six or seven, individuals. Factors of personality and social relationships (e.g., status, degree of acquaintance and heterogeneity), as well as intrinsic inter-connections between topics, enter into the determination of the flow of topics of conversation. It is only by careful experiments -- whether artificial or _post facto_ in natural settings -- that the relative influence of these several kinds of factors can be assessed.

FOOTNOTES

* This paper was read by Fred Costello, Rolf Kjolseth and Joseph Zelan. We appreciate the suggestions they made. Robin Anderson Pardini helped us to organize the data.

1. John G. Kemeny, J. Lauri Snell, and Gerald L. Thompson, Introduction to Finite Mathematics, Englewood Cliffs, New Jersey: Prentice-Hall, 1957, p. 171 ff.

2. Robert F.Bales, Interaction Process Analysis: A Method for the Study of Small Groups, Cambridge, Massachusetts: Addison-Wesley Press, 1950.

3. Henry T. Moore,"Further Data Concerning Sex Differences," Journal of Abnormal and Social Psychology, 17 (July-September, 1922), pp. 210-214.

4. M. H. Landis and H. E. Burtt, "A Study of Conversation," Journal of Comparative Psychology, 4 (February, 1924), pp. 81-89.

5. Carney Landis, "National Differences in Conversation," Journal of Abnormal and Social Psychology, 21 (January-March, 1927), pp. 354-357.

6. Stuart M. Stoke and Elmer D. West, "Sex Differences in Conversational Interests," Journal of Social Psychology, 1-2 (February, 1931), pp. 120-126; also in Jerome M. Seidman (ed.), The Adolescent: A Book of Readings, New York: Dryden Press, 1953, pp. 247-255.

7. The categorizations, "personal," "campus," and "intellectual" will be discussed below.

8. Using the random sample data as our basis for judgement, we would note that the inclusion of the participant observer data does overstate somewhat the amount of intellectual discussion which we believe actually occurs among college students. For a fuller discussion of the data gathering methods and an evaluation of the same, see Charles D. Bolton and Kenneth C. W. Kammeyer, The University Student: A Study of Student Behavior and Values, New Haven, Connecticut: College and University Press, 1967.

9. The mode of analysis employed here follows the procedures suggested by Leo A. Goodman in his paper, "Statistical Methods for Analyzing Processes of Change," American Journal of Sociology, 68 (July, 1962), pp. 57-78.

10. Ibid., pp. 71-74. It might be noted that after any topic of a bull session the next state may be a Personal topic, Campus topic, Intellectual topic, or Popular Culture topic, or the next state may be no topic at all, i.e., termination. If termination is viewed as a state,

then the process would be most aptly described as an "absorbing state Markov chain." An absorbing state is some state into which the process goes and from which it cannot return. (Kemeny, Snell, and Thompson, op. cit., pp. 326-332). However, in this analysis we will not consider "stop" or "termination" as a state, since we are interested primarily in the topical content of the interaction among students. It might be noted in passing, however, that as one goes from topic 1 to state 2, to state 3, etc., there is an increasing probability that the next state will be a terminated state, regardless of the content of the prior topic.

11. We would admit that some of this difference may be attributed to the method of coding the topics discussed in the bull sessions. There was probably some tendency for the coder to develop a psychological set in coding consecutive topics, and short of some rather elaborate precautionary procedures this is always likely to occur in some degree. We did take the precaution of having all topics double coded by coders working independently, or the coding was checked by one of the principal investigators. We feel that the differences are so pronounced that they are more than mere artifacts of the data processing method.

12. Leo A. Goodman and William H. Kruskal, "Measures of Association for Cross Classification," Journal of the American Statistical Association, 49 (December, 1954), pp. 740-742. See also Morris Zelditch, A Basic Course in Sociological Statistics, New York: Henry Holt and Company, 1959, pp. 174-177.

13. Goodman and Kruskal's Tau b is based on a prediction model, and also indicates about a 15% improvement in prediction. Ibid., pp. 759-760.

14. Goodman, "Statistical Methods for Analyzing Processes of Change," op. cit., pp. 63-65.

15. Ibid., pp. 75-77.