This is to certify that the

dissertation entitled

THE RATE OF INTERACTIVE INFORMATION DIFFUSION

presented by

Connie L. Bauer

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Communication

Richard V. Farace
Major professor

Date 4/9/82

THE RATE OF INTERACTIVE INFORMATION DIFFUSION

By

Connie L. Bauer

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Communication

1982

Accepted by the faculty of the Department of Communication, College of Communication Arts and Sciences, Michigan State University, in partial fulfillment of the requirements for the Doctor of Philosophy degree.

Director of Dissertation

Guidance Committee: , Chairperson

ABSTRACT

THE RATE OF INTERACTIVE INFORMATION DIFFUSION

By

Connie L. Bauer

The interactive information diffusion process is the person-to-person spread of information in a system of interacting members, i.e., communication network. For successful diffusion, the overall behavior of the accumulated number of knowers is described by the logistic equation. However, the logistic equation does not explain or predict the variations in the rate of spread for different messages in different networks. The purposes of this research were to: (a) develop a theory (model) which relates selected variables to the overall behavior of the process (i.e., logistic equation), and (b) conduct a study to test the theory (model).

Since the purpose of the theory is to explain and predict the <u>rate</u> of the overall interactive diffusion process, the logistic b parameter became the dependent variable. The six selected exogenous variables are: (a) network density, (b) network reachability, (c) anchor centrality, (d) anchor reachability, (e) message temperature, and (f) message transmission resistance. Based on the six variables' boundary conditions, the theory specifies a multiplicative functional form of the variables. To test the theory, a nonexperimental field study was conducted. A single message was diffused in multiple networks (N = 33).

The results of the data analyses indicate that the multiplicative functional form is appropriate. In addition, the results indicate that the two message variables (message temperature and message transmission resistance) can be considered constants in the study. Contrary to the proposed theory, the amount of network density was not found to be an important predictor variable. In addition, anchor reachability was found to have a positive rather than negative relationship with the logistic b in the multi-exogenous equation. The major outcome of this research has been to provide a basis and direction for developing and testing a theory of the rate of interactive information diffusion process.

ACKNOWLEDGMENTS

This dissertation is the materialization of a dream. Needless to say, over the past five years my dream has been influenced by a number of people whose contributions I would now like to recognize.

I would first like to thank Dr. Richard V. Farace, my advisor, for his many contributions, guidance, and patience. Also, I'd like to thank the other members of my doctoral committee--Dr. Edward L. Fink, who inspired this dissertation; Dr. Frank Fear, who encouraged me to maintain a balance between the theoretical and practical aspects; and Dr. Peter R. Monge, who taught me patience and determination to overcome all obstacles.

Gaining entry into an organization to gather research data can often be quite difficult. Hence, a special thanks to Ann Johnson, Resident Assistant, who helped me gain permission to survey dorm residents and who also helped to recruit other Resident Assistants to participate in the study. I also want to thank Audrey Delidow and Stan Avery for volunteering their time and energy to help gather the data.

Over the past five years, many people have contributed to my progress in a variety of ways. However, two people will always stand out in my memories. One person is Dr. Michael Burgoon who lifted a "straw that almost broke the camel's back." During fall term, 1981, when I was not a registered student, Dr. Burgoon gave me two visitor's parking permits. In addition, I'd like to thank him for helping me to keep my fighting spirit alive and well. The second person is Lucille Johnson who became a close friend. For the past nine months, she helped me to maintain an outward cloak of sanity and cheered me on to the "finish line."

Lastly, but most importantly, I want to thank Lu Horgen who shared the dream with me that this day would come. Lu helped me to keep the dream alive through all the storms and droughts of a graduate student's life.

TABLE OF CONTENTS

Chapter | Page

LIST OF TABLES

Table		Page

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER I

OVERVIEW AND PROBLEM STATEMENT

Introduction

This study is concerned with the process by which information spreads to the members of a social system. My initial interest in person-to-person information spread developed while working for a small company some years ago. The company tested electronic components for manufacturers according to the manufacturer's specifications. Although the company was quite small and informal, there seemed to be a continuous stream of complaints by various departments about not receiving necessary information in a timely fashion. For example, when someone in sales accepted an order, the information often took an unusually long time to reach the order entry clerk, production scheduling, production, and shipping. Information was not circulating as it should have nor as fast as it should have for the company to meet its deadline commitments. The company's frequent response to the problem was to fire a supervisor or restructure top management (four times in 18 months). The outcome of this research is expected to help a communication manager to understand and rectify similar information problems.

The general process of person-to-person communication in a social system has interested social scientists as far back as Durkheim

(1885/1966; Nisbet, 1974), Tarde (1903), and Simmel (1922/1955; 1950). One sub-area of person-to-person communication is the spread of new information in a social system, i.e., a set of interacting members (Rogers with Shoemaker, 1971). There are numerous studies which describe the person-to-person information diffusion process and individual variables related to source, receiver, message, and social system characteristics. However, there does not appear to be a theory concerning the rate or speed of person-to-person information spread. Therefore, the general purpose of this research is to develop a theory which explains and predicts the speed (rate) at which information diffuses in a social system.

Before proceeding further, two types of information diffusion processes will be distinguished. The two major processes are based on the two basic types of transmission. One is the broadcast process, in which mass media are the source of transmission. The overall process behavior (i.e., cumulative number of knowers over time) is characterized by a growing asymptotic curve (e.g., Feldman & Lie, 1974; Funkhouser, 1970; Funkhouser & McCombs, 1972; Gray & von Broembsen, 1974; Lave & March, 1975). The second is the interactive process, which is based on person-to-person transmission. Here the overall behavior is characterized by the logistic (S-shaped) curve of the cumulative number of knowers over time (Brown, 1968; Dodd, 1953, 1955; Gray & von Broembsen, 1974; Feldman & Lie, 1974; Hamblin, Miller, & Jacobsen, 1973; Lave & March, 1975; Rogers with Shoemaker, 1971).

Numerous information diffusion studies have reported on the heavy reliance by our society on mass media sources for certain types of information (e.g., Budd, MacLean, & Barnes, 1966; Deutschman & Danielson, 1960; Greenberg, 1964a, 1964b; Larson & Hill, 1954; Schneider & Fett, 1978). However, there is still much information which is transmitted primarily by interpersonal channels within a social system, e.g., job information (Granovetter, 1973). Such information may be of interest only to a particular social system and not the populace at large. For example, much of the information diffused in an organization is relevant only to that particular organization or, at most, other organizations within a particular industry. Often information may not be considered newsworthy enough, the subject matter too sensitive, or just simply not reach a mass media source for transmission.

While we know from our everyday interactions that we receive and transmit much information (e.g., Larson & Hill, 1954; Lionberger & Hassinger, 1954), only recently has human communication research begun to investigate the diffusion process in a social system of interacting members (e.g., Rogers, 1979; Rogers & Kincaid, 1981). Because much of our information is communicated via interpersonal sources in our formal (work) and informal (social) groups, the focus of this research is on the interactive process of information diffusion. Information from the mass media will be considered as a possible initial source of information which may stimulate interactive diffusion.

Literature Overview

The vast body of interactive information diffusion literature can be classified into three categories. The largest category is composed of studies which focus on message characteristics, usually in the broadcast process (e.g., Budd, MacLean, & Barnes, 1966; Deutchman & Danielson, 1960; Fathi, 1973; Greenberg, 1964a, 1964b; Greenberg, Brinton, & Farr, 1965; Haroldson & Harvey, 1979; Hill & Bonjean, 1964; Hyman & Sheatsley, 1947; Larson & Hill, 1954; Schneider & Fett, 1978; Spitzer, 1964-1965). Studies concerned with message characteristics for either the broadcast or interactive diffusion process will be reviewed in the following chapter.

A second, but much smaller, category of interactive information diffusion studies are those which focus on the social sysem's structural characteristics. The social system of interest in interactive information diffusion is the communication system or network, i.e., communication patterns in a group of interacting members. The analysis for describing communication patterns is referred to as network analysis.

Within the interactive information diffusion literature, there are few studies which identify the network (system) characteristics which influence the rate and pattern of information diffusion (Dodd, 1958-1959; Dodd & McCurtain, 1965; Erbe, 1962; Friedkin, 1979; Granovetter, 1973; Larson & Hill, 1958; Lionberger & Hassinger, 1954; Rogers & Bhowmik, 1971). Innovation adoption is also an interactive process, producing a logistic curve of the cumulative number of adopters over time and includes information diffusion (broadcast and interactive) about the innovation as a subprocess (Rogers with Shoemaker, 1971). Hence, innovation diffusion studies which focus on network characteristics (e.g., Becker, 1970; Coleman, Katz, & Menzel, 1957; Czepiel, 1975; Davis, 1963; Guimaraes, 1972; Lin & Burt, 1976; Liu & Duff, 1972; Menzel & Katz, 1955-56; Rogers, 1979; Rogers & Kincaid, 1981; Shoemaker, 1971) are relevant to the network aspect of interactive information diffusion. Although information diffusion is considered a subprocess of innovation diffusion, no adoption studies have been found which examine characteristics of the information about the innovation. For structural characteristics, network analysis in information and innovation diffusion literature will be reviewed under the general heading of networks.

The third category of literature is mathematical modeling which has evolved out of epidemiology. The spread of contagious diseases is also an interactive process which results in the overall behavior represented by the logistic curve (i.e., cumulative number of infected over time). Mathematical models of epidemiology have been extensively developed (an example of which is the work of Bailey, 1975). Because of the richness of mathematical epidemiology models, there have been numerous comparisons of the interactive diffusion process (both information and innovations) to the process of contagious diseases (e.g., Bartholomew, 1973; Brown, 1968; Cane, 1966; Coleman, 1964; Daley & Kendall, 1964; Dietz, 1967; Goffman, 1966; Goffman & Newill, 1964; Lave & March, 1975; Monin, Benayoun, & Sert, 1976). While this line of inquiry does not appear to have generated any theory which explains or predicts the rate of interactive information diffusion, it has generated interest in developing similar types of mathematical models for interactive information diffusion.

In summary, the broadcast information diffusion literature (mass media studies) provides information about message characteristics. The interactive diffusion of information and innovation (adoption) literature provides some information on possible relevant network characteristics. The literature on mathematical modeling provides descriptive and predictive models of the overall behavior of interactive information diffusion.

Problem Statement

Theoretical Problem

There are enough studies which validate the logistic curve as a description of the overall behavior of the interactive information diffusion process. However, there is no model (or theory) which explains and predicts the different overall rates of the logistic curve resulting from different messages diffused in different social systems (i.e., networks).

Given the studies on message characteristics, network structure, and mathematical modeling, the next step in advancing the field is to integrate the various findings. Therefore, there is a need to develop a comprehensive theory which relates both message (information) and network (system) variables to the variation in rates of the overall interactive information diffusion process.

Practical Problem

Quite often human communication scientists are asked by practitioners for advice on ways to quickly disseminate information within a social system of interacting members (e.g., organizations, neighborhoods, or communities). This is often the case when information, for one reason or another, cannot or will not be transmitted via the mass media. Relying on the interactive process, a practitioner would want to know who in the network should be given the message for optimum spread; what is the best form (wording) of the message; and how long will the message take to diffuse completely through the network of interest. Although there have been many studies conducted on message and network characteristics, communication scientists must rely on piece-meal studies to advise the practitioner. Resolution of the theoretical problem above will produce a theory to help guide the practitioner to disseminating information via person-to-person contacts. It will also help the practitioner to predict the amount of time the diffusion process will require for a given message in a particular network.

Purpose of Study

The purpose of this research is to develop and test a theory to explain and predict the rate of interactive information diffusion based on network and message variables. The theory to be developed will be composed of: (a) conceptual definitions of the theoretical variables, (b) the bivariate relationships of the endogenous variable with each of the exogenous variables and boundary conditions, (c) the multi-exogenous equation representing the theory, and (d) the measurement model for each of the theoretical variables.

Organization of the Dissertation

The following chapter reviews the major literature on interactive information diffusion, in four sections. The first part of the chapter presents literature on the historical background of the problem, and is followed by literature on the mathematical modeling of the interactive information process. The third section of the chapter discusses relevant network literature. The fourth part of Chapter II is a review of message variables suggested to affect the rate of interactive information diffusion.

Following the literature review, Chapter III presents a theory of the rate of interactive information diffusion. Chapter IV describes a field study designed to test the theory presented in Chapter II, and Chapter V contains the results of the study. The final chapter, Chapter VI, is a discussion of the research results, a critique of the research with suggestions for future research, a discussion of the research outcomes and theory application, and an overall summary.

CHAPTER II

LITERATURE REVIEW

Theoretical Background

Interactive diffusion occurs in a variety of person-to-person transmission contexts. The "object" transmitted or passed from one person to another may be information (in the case of interactive information diffusion), a new practice, technology, idea, or a contagious disease. For the transmission of information or an innovation, people must come into communication contact, e.g., face-to-face, telephone, or mail. For the transmission of a contagious disease, an infected person must come into physical proximity with another person (face-to-face), or a vector must first contact the infected person, and then contact a healthy person (e.g., mosquito in malaria). In all three cases, a person passes the "object" to one or more other people. Thus, the "object" spreads in a chain-like reaction, resulting in an overall S-shaped or logistic growth pattern, i.e., cumulative number of knowers, adopters, or infected over time.

Because contagious diseases, innovations, and information all spread interactively and yield the same overall logistic growth curve, their research and theoretical histories overlap. However, the greatest overlap is between interactive information and innovation diffusions. Not only do interactive information and innovation diffusions occur in a social system of interacting members, but both depend on communication interaction (Rogers with Shoemaker, 1971). The first two stages in the innovation diffusion process are awareness and knowledge of the innovation. Thus, information diffusion (interactive and broadcast) is considered a <u>sub</u>process of innovation diffusion (Rogers with Shoemaker). Innovation diffusion also includes subsequent subprocesses of persuasion, decision (adoption, rejection, or discontinuance), and implementation of the innovation. Because of these other subprocesses, innovation diffusion research focuses primarily on the influence rather than on the information-exchange patterns (i.e., networks). However, the same variables are often used to describe the two different types of networks.

Many research areas have used studies of their own substantive areas to contribute to the body of interactive information and innovation diffusion literature. These areas include sociology, human communication, agriculture, marketing, advertising, political science, anthropology, education, and medical sociology; the field of geography has focused on the spatial diffusion of information and innovations (e.g., Hägerstrand, 1953).

While many discipline areas have been involved, this section focuses on the major evolutionary background of interactive information diffusion. Those areas of research which have made significant contributions are discussed. These areas and their major relationships to one another are presented in Figure 1. The directed lines indicate the direction of contributions from one research area to another.

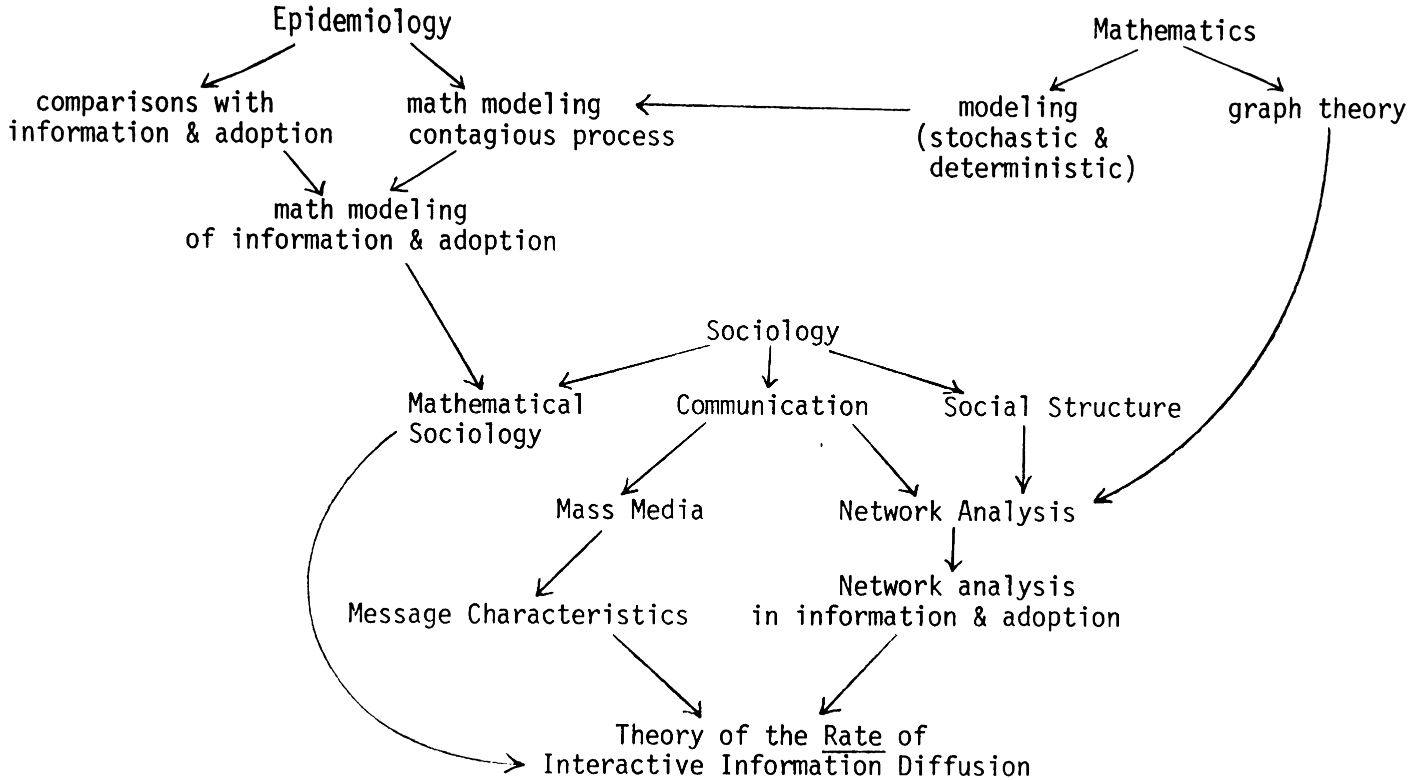


Figure 1. Evolutionary background of interactive information diffusion.

This is a very simplified overall evolutionary scheme and not a strict chronological ordering. The purpose of Figure 1 and the following discussion is to give the reader a general background overview to the problem addressed by this research.

As mentioned above, many social science areas have been interested in the interactive information diffusion process. However, major theoretical contributions by social scientists have been largely based on developments in the fields of epidemiology and mathematics. Because contagious diseases have had devastating effects on the human population, biologists have long been interested in the spread of such diseases. To understand the process and to predict the rate of spread, mathematical epidemiologists have developed highly sophisticated models (deterministic and stochastic). This area of research is exemplified by the work of Bailey (1957, 1975). Because of the theoretical modeling developments in epidemiology and the similar overall behaviors (i.e., S-shaped cumulative curve), numerous attempts have been made to apply epidemiological models to interactive information and innovation diffusions (e.g., Bartholomew, 1973; Brown, 1968; Cane, 1966; Coleman, 1964; Daley & Kendall, 1964; Dietz, 1967; Goddman, 1966; Goffman & Newill, 1964; Lave & March, 1975; Monin, Benayoun & Sert, 1976; Rapoport, 1956).

The work by Dodd (1953, 1955) triggered many comparisons of interactive information diffusion with contagious disease. Dodd, a sociologist, was the first to suggest that the accumulated curve of interactive information diffusion is logistic (S-shaped). However, equating the components of interactive information diffusion (e.g., knowers, nonknowers, number of tellings, etc.) to the components of the contagious disease process (e.g., infected, susceptibles, infectious period, etc.) has not resulted in any theoretical advances in interactive information diffusion. The major contribution of epidemiology to interactive information diffusion has been to serve as a theoretical role model for social scientists, in particular, mathematical sociologists (e.g., Coleman, 1964; Hamblin, Jacobsen, & Miller, 1973; Sørensen, 1978).

Parallel research to epidemiology and mathematical sociology was conducted in the new emerging area of human communication. The major focus by communication scientists has been largely on the broadcast process of information diffusion (i.e., mass media research). The contribution by mass media research to interactive information diffusion has had a minor focus on message characteristics which may influence the rate of spread (e.g., Budd, Maclean, & Barnes, 1966; Deutschman & Danielson, 1960; Funkhouser, 1970; Greenberg, Briton, & Farr, 1965; Haroldsen & Harvey, 1979). However, no study systematically examines message characteristics and resulting changes in diffusion rates for either the broadcast or interactive processes.

National news is primarily transmitted by the mass media. However, mass media researchers have commented on the frequent co-occurrence of interactive diffusion with the broadcast process as far back as Lazarsfeld and Katz (Katz & Larzarsfeld, 1955; Lazarsfeld, Berelson, & Gaudet, 1948). To explain these observations mass media researchers focused on the two-step and subsequent multi-step flow

hypotheses of information proposed by Lazarsfeld et al. (1948) and Menzel and Katz (1955), respectively. This line of research highlighted the role of opinion leadership in bridging the broadcast and interactive information diffusion processes. However, only a few studies have examined the social structure involved in interactive diffusion (Dodd, 1958/59; Garabedian & Dodd, 1962; Larson & Hill, 1954, 1958; Lionberger & Hassinger, 1954).

Except for the classical study by Coleman, Katz, & Menzel (1957) and a few subsequent others (Becker, 1970; Czepiel, 1974; Rogers, 1979; Rogers & Kincaid, 1981), innovation diffusion has also suffered from a lack of research focus on the social structure. Almost 20 years ago, Katz, Levin, and Hamilton (1963) criticized innovation diffusion research for a lack of focus on ways in which different kinds of structural arrangements in a group influence the diffusion of a given item. For further progress in the field, Katz et al. stated that:

> interpersonal channels of communication must be viewed as elements of social structure; (6) that work is urgently needed on the comparative study of the same item diffusing in different social structures . . . (p. 252).

The same criticism and recommendation could have been made of interactive information diffusion; it is still valid today for both processes.

The absence of studies on communication structure has largely been due to a lack of appropriate analytic methodology. However, largely based on the work of Harary, Norman, and Cartwright (1965), both communication scientists and sociologists have developed what has come to be known as network analysis. The principles of network analysis are largely drawn from graph theory, a branch of mathematics. Network analysis provides a way of examining the communication structure within which information diffuses via the interactive process. With recent developments in computer programs for analyzing networks (e.g., Richards, 1975), network analysis allows examination of communication structural (i.e., network) variables which may influence interactive information diffusion.

In conclusion, the one generally accepted finding about interactive information diffusion is that, over time, the accumulated number of knowers produces the logistic or S-shaped curve. The logistic curve represents the overall behavior of the process. The two major components of the process are: (a) the communication structure (network) in which a (b) message diffuses. Only limited research has been conducted on message/information variables and network variables which may influence the speed at which a message diffuses in a network. As yet, there has been virtually no attempt to integrate past research findings into a theory which would explain and predict the growth rate of the logistic curve for interactive information diffusion.

The following sections of this chapter present a review of the literature on mathematical modeling, network structure, and message characteristics. The section on mathematical modeling reviews the history of the logistic curve and its associated equation along with a theoretical modeling strategy for relating component variables (network and message) to the overall behavior. The sections on network structure and message characteristics review the literature on variables which influence the rate of interactive information diffusion.

Mathematical Modeling

This section presents a brief discussion of deterministic versus stochastic models, followed by a discussion of the logistic equation. The last part of this section presents the modeling strategy adopted for relating network and message variables to the logistic equation.

Deterministic vs. Stochastic

Most mathematical work on information diffusion (interactive and broadcast) has been on stochastic models (e.g., Bartholomew, 1973, 1976; Funkhouser, 1970; Funkhouser & McCombs, 1972; Gray & von Broembsen, 1974; Karmeshu & Pathria, 1980; Rapoport, 1953a, 1953b, 1979; Rapoport & Rebhun, 1952; Sørensen, 1978; Taga & Ish, 1959). Stochastic models focus on the probability of a given person being affected, i.e., the probability of individuals moving from one state to another (e.g., nonknower to knower or nonteller to teller). Deterministic models focus directly on change in the variable values or on the number of people affected (Monin et al., 1976; Sørenson, 1978).

Deterministic models can include a random term which is introduced into the solution of the defining equation. In contrast, the defining equation in stochastic models uses changes in probability

levels as the dependent variable (Sørensen, 1978). While stochastic models may be more predictively accurate, deterministic models often provide good approximations, since the deterministic form can often be taken as the expected or mean value of the probability distribution (Coleman, 1964, p. 527; Olinick, 1978, p. 11). Thus, the deterministic model can reflect in a simpler form the same basic process as does the stochastic model (Coleman, 1964, p. 527; Olinick, 1978, p. 327). The major difficulty with stochastic models is that with increased complexity they quickly become intractable (Coleman, 1964; Olinick, 1978; Sørensen, 1978). Therefore, deterministic models are preferred.

The commonly used deterministic model of complete interactive diffusion of information is the logistic equation (e.g., Bartholomew, 1976; Dodd, 1953, 1955; Feldman & Lie, 1974; Goffman, 1966; Goffman & Newill, 1964; Landahl, 1953; Landau & Rapoport, 1953; Lave & March, 1975; Monin et al., 1976; Sørensen, 1978). It represents the S-shaped or logistic curve and is simpler than its stochastic counterpart. The logistic equation for interactive information diffusion is:

$$n = \frac{k}{1 + e^{-(a + bt)}} \qquad (1)$$

where n is the cumulative proportion of knowers of the information, k is the total proportion of possible knowers (i.e., size of the social system), t is time, a is a constant scaling parameter to be estimated, and b is the parameter which summarizes the rate of growth, also to be estimated.

The logistic equation is a description of the overall behavior on the global level by condensing data into two parameters (a and b) (Feldman & Lie, 1978). Since it is the basic model of interactive information diffusion, the logistic equation is discussed further in the next section.

Logistic Equation

The logistic curve, which is an S-shaped symmetrical, cumulative growth curve, has been found to be applicable to populations with limited growth (von Bertalanffy, 1950). The rate of growth is not constant as in unlimited exponential growth, but is dependent on the maximum possible size of the population, e.g., k in the logistic equation (1).

Historically, the logistic equation was first developed in the 1840's by a Belgium mathematician, Pierre-Francois Verhulst, to model population growth (Olinick, 1978). Because the existing census data at that time were inadequate to form any effective test of his model, Verhulst's discovery of the logistic curve and its application to population growth was forgotten for almost 80 years. It was rediscovered in the 1920's independently by two American biologists, Raymond Pearl and Lowell J. Reed (Olinick, 1978).

In addition to population or biological growth, the logistic curve has also been found applicable in describing the cumulative growth of autocatalytic chemical reactions (Coleman, 1964; Hamblin et al., 1973), the epidemic of communicable diseases (e.g., Bailey, 1957, 1975; Bartholomew, 1973; Dietz, 1967; Monin, Benayoun & Sert, 1976; Serfling, 1952), the diffusion of innovations (e.g., Brown, 1968; Feldman & Lie, 1974; Griliches, 1957; Hamblin et al., 1973; Rogers with Shoemaker, 1971), and the interactive diffusion of information (e.g., Dodd, 1953, 1955; Bartholomew, 1973, 1976; Feldman & Lie, 1974; Funkhouser, 1970; Gray & von Broembsen, 1974; Landau & Rapoport, 1953; Lave & March, 1975; Rapoport, 1979/80; Rogers with Shoemaker, 1971). The first theoretical work on and application of the logistic equation to the S-shaped curve for successful interactive information diffusion was done by Dodd (1953, 1955).

Complete interactive information diffusion results in logistic growth rather than exponential, decaying exponential, or any other growth curve because of: (a) the interactive or chain-reaction process in a (b) finite population (network) where someone who knows the information tells it to other people (e.g., Dodd, 1953, 1955; Feldman & Lie, 1974; Lave & March, 1975). The growth limit (k) is the size of the communication network. In the early stage, the interactive process will be relatively efficient because there is a high probability of talking to someone who does not have the information. Thus, the growth rate of the proportion of knowers increases and the lower half of the logistic curve is exponential growth. Toward the end of the process, the growth rate of the cumulative proportion of knowers is very slow. By this time most of the contacts will be with people who already have received the information rather than with the few remaining people who have not yet received it (Lave & March, 1975).

The upper half of the logistic curve is asymptotic growth resulting from a decreasing rate of the proportion of knowers.

Two other mathematical functions, the Gompertz and the normal ogive, have also been used to describe the S-shaped curve of successful diffusion (both information and innovation). The Gompertz equation has been largely rejected because it describes an asymmetrical S-shaped curve, yet the accumulated S-shaped curves of innovation and interactive information diffusion data are generally quite symmetrical (Hart, 1945; Hamblin et al., 1973).[1] Both the logistic and normal ogive curves are symmetrical. Working with diffusion of innovation data, Pemberton (1936) was the first to suggest that the normal ogive, the integrated normal frequency distribution, was the equation that best described the S-shaped curve.[2] However, the logistic equation is based on a chain-reaction process in a finite population and is, therefore, the preferred equation for mathematically describing successful innovation diffusion and successful interactive information diffusion (Feldman & Lie, 1974). The logistic equation is also preferred over the normal ogive because it can be easily transformed to a linear (in parameters) equation for a least-squares estimation of the two parameters:

$$\ln \left(\frac{n}{k - n}\right) = a + bt, \qquad (2)$$

or, in linear regression form:

$$n^* = a + bt, \qquad (3)$$

where $n^* = \ln(n/k - n)$.[3]

The logistic curve is illustrated in Figure 2 where k represents the total possible population of knowers, i.e., the maximum possible limit of growth for interactive information diffusion. Across different diffusion studies, high b parameters represent faster overall growth rates.

Mathematical Modeling Strategy

The equation chosen to model the overall behavior of interactive information diffusion is the single deterministic logistic equation (1) with one exogenous variable, i.e., time. The logistic equation is adequate for describing the overall process's behavior for any one particular interactive diffusion study. However, it does not provide any explanation as to why the overall growth rate of the accumulated knowers (i.e., the b parameter) differs across messages and across communication networks. The focus of this research is to develop a theory which will explain why the overall diffusion time as represented by the b parameter is greater or lesser for one interactive information diffusion study than another.

Sørensen (1978) presents two main strategies for modifying simple models to improve their theoretical adequacy. The first strategy is to model the variation in parameters of simple models. This is done by writing parameters as functions of independent variables (i.e., deterministic modeling) or by modeling the distribution of parameters (i.e., stochastic modeling). The second strategy is to modify the dependent variable. Modifying the dependent variable can be dealt with by introducing time-dependent variables as independent

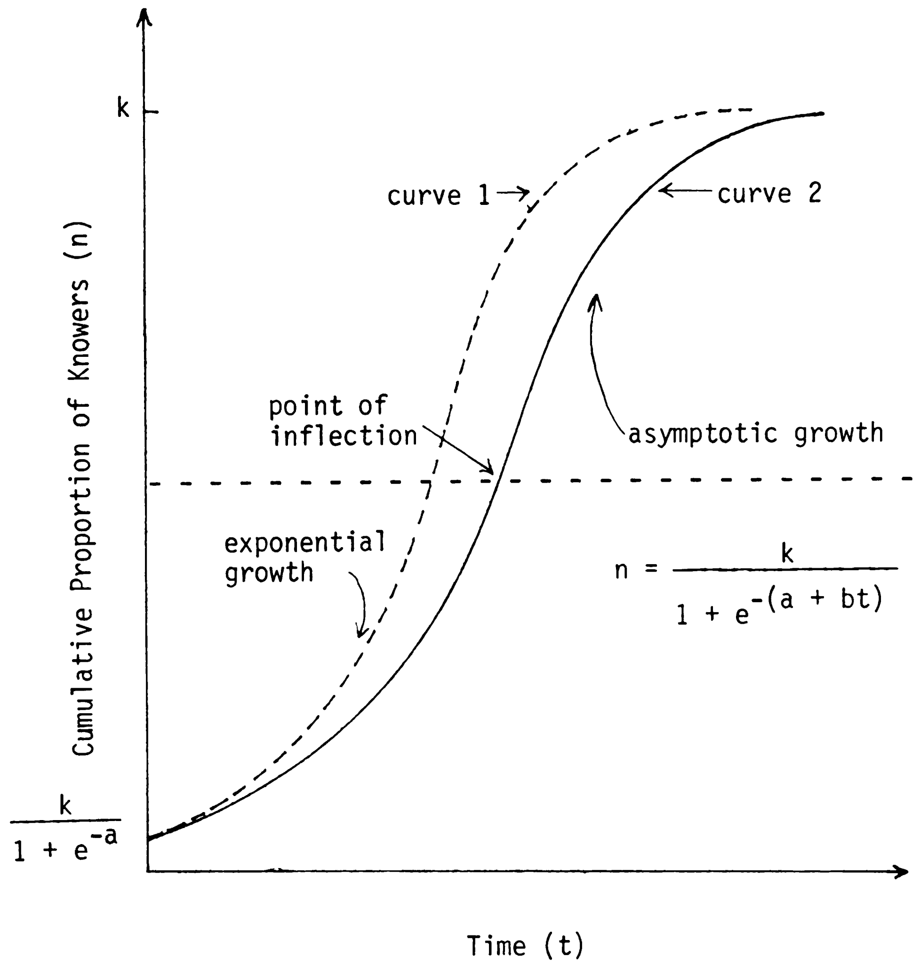


Figure 2. Logistic growth curve where n is the cumulative proportion of knowers, k is the limit or total possible proportion of knowers. Curve 1 represents a more rapid spread of information than does curve 2, i.e., b_1 b_2.

variables or by combining several dependent variables, e.g., simultaneous equations model.

The first strategy focuses on variables to explain and predict the variation in parameters of the basic model. For interactive information diffusion, the parameter of interest is the b in equation (1). The b parameter summarizes the accumulated growth rate (or speed) of diffusion for a specific message in a specific communication network. The questions of interest in this research are: why does a message diffuse faster (or slower) in one network than another and why will one message diffuse faster than another in the same network? Theoretically, these questions ask what variables are needed to explain the variation of the logistic b parameter. Thus, the logistic b parameter is the dependent variable of interest.[4]

Writing the logistic b parameter as a function of independent variables has been applied by Hamblin et al. (1973) to explain the different rates of adoption with a wide variety of innovations. Hamblin et al. model the logistic b parameter as a function of the level of reinforcement for adoption. Thus, they consider the logistic b a variable to be explained. In addition, Chadha and Chitgopekar (1971) also use this strategy to explain changes in the potential market (k in the logistic equation) for the adoption of residence telephones.

For interactive information diffusion, there does not appear to have been any attempts at applying Sorensen's first strategy to the logistic b parameter. Thus, similar to the work by Hamblin et al., a set of independent variables will be sought to explain and predict the variation of the logistic b parameter resulting from different messages diffusing in different networks. The following two sections of this chapter review the literature on network structure and message characteristics, respectively. From this view, variables are selected or developed to explain the variation of the logistic b parameter. The model (or theory) relating the selected variables is presented in Chapter III along with the measurement models for each selected variable.

Structural Variables

As described above, interactive information diffusion refers to the spread of a message (information) in a social system of interacting members. The social system is defined as a communication network where communication is the basis of interaction. The location of the point where the message enters the network has also been suggested as important to the overall diffusion rate (Bavelas, 1948; Becker, 1970; Czepiel, 1975). This section reviews the literature for possible network and message entrance location variables which might most affect the overall rate of interactive diffusion, i.e., the logistic b parameter. The discussion of structural variables is divided into two parts: Network structure and network location of message entrance.

Network Structure

Communication networks are defined "as the set of stable person-to-person relationships through which information flows . . ."

(Monge, Edwards, & Kirste, 1978, p. 312). The network structure is the actual configuration of the communication relationships. Individuals in the network are referred to as nodes and the communication relationship between two people (nodes) is called a link (Richards, 1976). The strength of the link is based on frequency of communication interaction, duration of interaction, and/or intensity of interaction (Richards, 1976). Thus, a network is a configuration of a set of nodes interconnected by communication links.

The network structure influences the speed at which the message moves from person to person because it imposes constraints on the possible paths the message can traverse (e.g., Bavelas, 1948). However, only a small number of studies actually compare different networks (social systems) and the flow of information (e.g., Dodd, 1953, 1955; 1958/59; Larson & Hill, 1954, 1958; Lionberger & Hassinger, 1954). Dodd found that as city size and population density increased, the rate and extent of diffusion of airborne leaflets increased. Larson and Hill found that the rate of message spread is greater in more stable networks (i.e., communities) than in less stable ones. In comparing sources of farm information, Lionberger and Hassinger found that neighborhood residents most frequently named friends and neighbors as their most important sources while non-neighborhood residents most often named the mass media.

Most networks are compared for adoption rates or innovativeness levels (Allen, 1970; Becker, 1970; Coleman et al., 1957; Czepiel, 1975; Guimaraes, 1972; Liu & Duff, 1972; Rogers, 1979; Rogers &

Kincaid, 1981; Shoemaker, 1971). Innovativeness has been found to increase with increases in system openness (Allen, 1970) and communication integration (Allen, 1970; Giumaraes, 1972; Shoemaker, 1971). Increased adoption rates have been found to be positively associated with centrality (Becker, 1970; Czepiel, 1975), integration (Coleman et al., 1957), heterophilous relationships (Liu & Duff, 1972; Rogers & Kincaid, 1981), weak ties (Liu & Duff; Rogers & Kincaid), and system openness (Rogers & Kincaid). In addition, linkage distance (i.e., reachability), has been found to be negatively related to the adoption rate (Rogers & Kincaid). Because there are only a few network studies on interactive information diffusion, network studies on the adoption process are also included in this section.

Many ways have been developed to characterize various aspects of a network. They can be classified as: (a) link of dyadic properties (e.g., symmetry, strength, and reciprocity), (b) network roles (e.g., isolates, group members, bridge, and liaison), (c) network variables at the individual or node level (e.g., centrality, choice status, and reachability), (d) network variables at the group or clique level (e.g., connectedness, clique size, openness and dominance), and (e) network variables at the network level (e.g., density, reachability, openness, and anchorage). The variables of interest are those which provide description of structure at the network level rather than at the individual or dyadic (link) level. Since all network level variables are defined in terms of links, the link or dyadic properties are subsumed in the network level variables.

Network variables at the node level along with network anchorage are discussed in the section on network location of message entrance.

In the diffusion literature, several network related variables are suggested as influencing the rate of diffusion. A list of the variables found in the literature is provided in Table 1. Almost half focus on the node or dyadic levels rather than on the larger network level. The node and dyadic variables include: centrality, heterophilous relations, interaction frequency, valence of dyadic relationship, and weak ties. Centrality can be both a node and network level variable. Except for centrality, these variables appear to be very similar on the operational level.

Interaction frequency and valence are all measures of link strength (Richards, 1976), where weak ties (low link strength) fall at the lower end of the link strength continuum (Granovetter, 1973; Liu & Duff, 1972; Rogers, 1979). While heterophilous relations often refer to the degree of dissimilarity of values, education, and social status (Rogers & Bhowmik, 1971; Rogers with Shoemaker, 1971; Rogers & Kincaid, 1981), they can be operationalized as the amount of interaction frequency (Rogers, 1979; Rogers & Kincaid, 1981). For example, two people with a high degree of dissimilarity in attitudes and values tend to communicate with one another less frequently than do two people with a low degree of dissimilarity. Thus, heterophilous relations fall at the lower end of the link strength continuum of interaction frequency the same as weak ties (Granovetter, 1973; Liu & Duff, 1972; Rogers, 1979). While not identical, heterophilous

TABLE 1. Network Factors Influencing the Rate of Interactive Information Diffusion (from the Literature)

Factors	Researchers
1. Centrality	Bavelas (1948); Becker (1970); Czepiel (1975).
2. Clique size	Dodd & McCurtain (1965); Garabedian & Dodd (1962).
3. Connectedness	Bott (1955); Rogers (1979); Rogers & Kincaid (1981).
4. Density	Czepiel (1975); Erbe (1962); Granovetter (1973).
5. Heterophilous relations	Barnett (1975); Davis (1963); Erbe (1962); Liu & Duff (1972); Rogers & Bhowmik (1971); Rogers with Shoemaker (1971); Rogers & Kincaid (1981).
6. Interaction frequency	Erbe (1962).
7. Integration	Coleman, Katz, & Menzel (1957); Erbe (1962); Guimaraes (1972); Lin & Burt (1976); Rogers (1979); Shoemaker (1971).
8. Valence of dyadic relationship	Davis (1963); Fathi (1973).
9. Weak ties	Davis (1963); Friedkin (1980); Granovetter (1973); Liu & Duff (1972); Rogers (1979); Rogers & Kincaid (1981).
10. System openness	Allen (1970); Rogers (1979); Rogers & Kincaid (1981)
11. Reachability (Connectedness II)	Rogers & Kincaid (1981).

relations and weak ties appear to be closely related in some causal way. For the purpose of describing communication networks based on interaction frequency, heterophilous relations and weak ties are considered to be functionally the same. Weak ties are also related to density and network (or clique) openness and will be discussed below.

Centrality has been operationalized in several different ways. In tracing the history of the concept of centrality, Freeman (1978/79) found that there was little agreement on the definition and measurement of centrality. The only consensus Freeman found is that centrality is considered an important structural attribute of networks. Based on his literature review, Freeman categorizes the various centrality definitions into those based on: (a) direct communication activity or degree measures (e.g., number of direct links), (b) closeness (or distance)-based measures, and (c) betweenness or communication control-based measures. In the diffusion literature, centrality is used as either a communication activity (or degree) or a closeness (or distance)-based measure.

Centrality is used as a direct communication activity-based measure by Becker (1970) and Czepiel (1975). Alternatively, Bavelas (1948), who originated the term centrality (Freeman, 1978/79), defines it based on distance. Although Bavelas, Becker, and Czepiel use point centrality (i.e., node centrality), the variable can be used on the network level as discussed below (Freeman). Point centrality will be further discussed in the section on network location of message entrance.

When centrality is used as an activity-based measure at the network level, it is operationally the same as density (Czepiel, 1975; Erbe, 1962; Granovetter, 1973). Density was first operationally defined by Prihar in 1956 as the ratio of the actual number of links to the total number possible in a network (Barnes, 1969). Barnes was the first to refer to this measure as density. Thus, density is a direct communication activity-based measure of centrality at the network level.

As discussed by Barnes (1969), connectedness also has a variety of different definitions. However, based on the categories set forth by Freeman (1978/79), connectedness is used as a direct activity-based measure by Bott (1955), Rogers (1979), and Rogers and Kincaid (1981). As an activity-based measure on the network level, connectedness is the same variable as density. However, connectedness is also used as a closeness-based measure by Rogers and Kincaid, which they refer to as connectedness II. Rogers and Kincaid's connectedness II measure is the same variable as reachability, a distance-based measure.

The same situation also occurs for integration, a sociological term. Integration is used as an activity-based measure by Coleman et al. (1957), Erbe (1962), Guimaraes (1972), Lin and Burt (1979), and Rogers (1979). As an activity-based measure, integration is the same variable as density. Alternatively, Shoemaker (1971) uses integration as a distance-based measure.

In summary, there are two major types of centrality variables used in the diffusion literature. They are based on either direct communication activity or distance. For purposes of distinguishing between the two and for consistency with most of the communication network literature, the direct activity-based variable, hereafter, will be referred to as density. The distance-based variable will be referred to as reachability.

A completely connected network has no path constraints since every node directly connects every other node. An increase in density means there is less constraint on the paths which a message can take; therefore, a message should diffuse faster in a network with higher density.

At the group level only clique size and connectedness have been suggested as affecting the rate of interactive diffusion. Based on Monte Carlo simulations, Dodd and McCurtain (1965) and Garabedian and Dodd (1962) found that as clique size increased, the rate of information diffusion increased within a network. For the simulations, this is explained by the accompanying increase in the number of connections (links) the larger clique has with the rest of the network. Thus, in the simulations, increases in clique size also resulted in an increase in network density.

The strength of links connecting a clique to the rest of the network (or other cliques in the network) is often weaker (i.e., weak ties) than those of intraclique links (Friedkin, 1980; Granovetter, 1973; Liu & Duff, 1972; Rogers, 1979; Rogers & Kincaid, 1981). As mentioned above, weak ties indicate heterophilous relations and facilitate the flow of new information from one clique to another (Granovetter, 1973; Liu & Duff, 1972; Rogers, 1979). Alternatively, stronger ties within a clique facilitate intraclique diffusion. Thus, the greater the number of weak ties between cliques within a network, the more cliques overlap (Rogers, 1979) and the greater the density of the network as a whole. Therefore, the number of weak ties within a network is subsumed in the summary variable density.

The number of weak ties has also been suggested as a measure of the degree of openness of the clique or network (Rogers, 1979; Rogers & Kincaid, 1981). Openness is defined as the "degree to which a group has linkages with its environment" (Farace, Monge, & Russell, 1977, p. 202). When the degree of openness is applied to a clique, it is related to the density of the network. However, degree of openness as applied to a network does not appear to be a useful variable to account for the rate of diffusion in the network. Its usefulness seems to be more relevant to diffusion across networks.

Reachability, which has been briefly discussed above, can be regarded as a distance-based centrality measure, i.e., an overall measure of the shortest length or distance of a network. Conceptually, reachability is defined as the "extent to which other system members can be connected with a minimum of intermediaries . . ." (Farace et al., p. 202), i.e., the geodesic (Barnes, 1969; Freeman, 1977). Reachability takes into account both direct and indirect pathways; whereas, density takes into account only direct paths. In addition, if the links are asymmetric, the reachability score can reflect the paths that must be followed when the "one-way streets" of asymmetric relations are present (Farace et al.). Thus, a message diffusing in a network with low reachability should spread faster than in a network with high reachability.

Based on the literature and the above discussion, it appears that a network's density and reachability are two very useful network variables and subsume most of the variables at the group/clique, dyad, and individual levels. In a larger context, density (mass divided by volume) and reachability (distance) are important variables (characteristic properties of objects) in theories of physical and biological systems (e.g., fluid dynamics, kinetic theory of gases, and the diffusion of molecules) (Giancoli, 1980). Therefore, to explain and predict the rate of interactive diffusion of information, the amount of network density and the length of network reachability will be the two variables used to describe the network structure.

Network Location of Message Entrance

The orientation point or reference for a network is referred to as its anchorage (Farace et al.; Mitchell, 1969). For information diffusion, the node where the message enters the network will be referred to as the anchor. Because it is theoretically possible for information to be introduced into a network via any node having external contact, the network location of the anchor should affect the rate of diffusion in the network (Bavelas, 1948; Becker, 1970; Czepiel, 1975). For example, if an anchor with eight links is compared with an anchor with only two links, the accumulated number of knowers will grow faster if the node with the eight links is the message point of entrance.

It has been suggested that the most likely anchor will be a node with a weak, as opposed to a strong, tie to the environment (e.g., Friedkin, 1980; Granovetter, 1973; Liu & Duff; Rogers, 1979; Rogers & Kincaid). While weak ties may be the links which facilitate information diffusion across networks (or cliques), the anchor's location in the network will be important for the network diffusion rate.

To maximize the initial transmission of a message in a network, the optimum anchor location would be the most central node (i.e., the node with the most direct links to other network members). In both the information and innovation diffusion literature, the most central position is related to opinion leadership which has been found to be influential in innovation diffusion (Becker, 1970; Czepiel, 1975; Rogers, 1979; Rogers & Kincaid; Rogers with Shoemaker). However, the most central position based on the number of direct links is only one of many possible locations in a network. Used in this way, centrality is a communication activity variable describing one node's activity with all other network nodes (Freeman, 1978/79). The location of the anchor based on communication activity will be referred to as the amount of anchor centrality.

Not only does it seem important to know the anchor's direct communication activity, but also the minimum communication distance necessary for the message to reach the other network members.

Bavelas (1948) suggests that the minimum amount of time for complete information spread throughout a network is achieved if the spread starts with the most central node. Bavelas's measure of length of centrality is based on distance. Thus, the average shortest distance of the anchor to all other nodes would provide additional information on the location of the anchor. This distance-based variable of location for the anchor will be referred to as the length of anchor reachability.

Using both the amount of anchor centrality and the length of anchor reachability should provide sufficient location information about the network entrance point of the message. If one selectes the node where a message will be introduced, the optimum anchor for the most rapid spread would be the node with the highest amount of anchor centrality and the lowest length of anchor reachability.

Message Characteristics

Certain characteristics of the message or information have been suggested as influencing the rate of diffusion. Before proceeding to a discussion of characteristics, the entity being diffused is first discussed.

In the information diffusion literature, the terms "message" and "information" are not defined nor distinguished from one another. Often they are used interchangeably, resulting in the same operationalizations (e.g., Deutschman & Danielson, 1960; Funkhouser & McCombs, 1972; Liu & Duff, 1972; Schneider & Fett, 1978; Spitzer, 1964-65). The lack of definitions for message and information in the diffusion literature is justifiable on the basis that a commonly used word in "everyday" communication does not need to be defined. If a word is used in a more restricted or technical sense than its common usage, then the word warrants definition. For example, Shannon and Weaver (1949) used the word "information" in a more restricted and technical sense than its everyday meaning. Therefore, they provided a definition for their meaning to distinguish information from its "everyday" meaning/use. The context in which message and information are used in the diffusion literature indicates that their everyday meaning is implied and that they are not used in any restricted or technical sense.

A list of message characteristics which have been suggested as influencing the rate of information diffusion is provided in Table 2. The variables are divided into two categories for purpose of discussion. Upon close examination, the first eight characteristics in Table 2 are functionally equivalent. In fact, quite often the terms are used interchangeably. For example, news value is used interchangeably with age of a message (Landau & Rapoport, 1953) and importance (e.g., Budd, MacLean, & Barnes, 1966; Deutschman & Danielson, 1960; Hill & Bonjean, 1964). Utility is used interchangeably with importance (Greenberg, 1964b) and relevance (Schneider & Fett, 1978).

The only empirical investigation on message characteristic relationships was conducted on interest and utility; they were highly correlated (Greenberg, Brinton, & Farr, 1965). Even comparative

TABLE 2. Message Characteristics Influencing the Rate of Interactive Information Diffusion (from the Literature)

	Characteristics	Researchers
1.	Age of message	Landau & Rapoport (1953).
2.	Comparative novelty	Greenberg, Brinton, & Farr (1965).
3.	Complexity (relative) of information	Greenberg, Brinton, & Farr (1965); Lave & March (1975).
4.	Importance	Budd, MacLean, & Barnes (1966); Deutschman & Danielson (1960); Greenberg (1964b); Greenberg, Brinton, & Farr (1965); Haroldsen & Harvey (1979); Lave & March (1975).
5.	Interest value	Bartholomew (1976); Funkhouser (1970); Funkhouser & McCombs (1972); Gray von Broembsen (1974); Greenberg, Brinton & Farr (1965); Hyman & Sheatsley (1947); Landahl (1953); Lave & March (1975).
6.	News value	Budd, MacLean, & Barnes (1966); Deutschman & Danielson (1960); Faithi (1973); Hill & Bonjean (1964); Landau & Rapoport (1953).
7.	Relevance	Schneider & Fett (1978).
8.	Shocking value	Haroldsen & Harvey (1979).
9.	Tabooness of message topic	Rogers (1979); Rogers & Kincaid (1981).
10.	Utility	Greenberg, Brinton, & Farr (1965); Schneider & Fett (1978).

novelty and shocking value appear to be similar, based on degree of "unexpectedness." For the full set of message variables, conceptual definitions are not offered by the researchers so that one message characteristic cannot be distinguished from other closely related ones. The only measurements made are for interest (Funkhouser, 1970; Funkhouser & McCombs, 1972; Hyman & Sheatsley, 1947) and shocking value (Haroldsen & Harvey, 1970).

The first eight message characteristics seem to indicate a common underlying variable. A clue to this may be in the frequent slang expressions: "I have a hot tip for you" or "I have some hot news for you." Slang is based on immediate experience and offers us an immediate index to changing perceptions (McLuhan, 1964). Thus, the expression, "I have some hot news for you" indicates the speaker's immediate experience with the information. The two expressions seem to indicate highly charged affect (e.g., excitement and enthusiasm) in the speaker regarding the information about to be transmitted to another person.

Just as increases in temperature cause particles to more faster, we would expect a person to want to quickly pass on a "hot" message to many of his/her contacts. The perception of a message as "hot" can be called message temperature. The higher the temperature of the message, the faster it should spread. The first eight message characteristics in Table 2 which appear to be functionally equivalent can be reformulated as indicators of the underlying message temperature variable. For example, a new message would be expected to have a higher message temperature than an older one. Novelty, importance, relevance, utility, and news value of the information would also be expected to contribute to the message's temperature. In addition, increasing the interest and shocking value of a message should also increase the message's temperature. Therefore, the first eight variables in Table 2 are viewed as observed variables (indicators) of the unobserved variable called message temperature. This approach provides a framework for organizing functionally similar variables.

The complexity of information has been suggested as influencing the rate of spread (Greenberg et al., 1965; Lave & March, 1975). However, no empirical studies were found in the diffusion literature supporting this. Yet, it seems reasonable that complex information will not diffuse as fast as simple information because of the greater difficulty in transmitting (i.e., "telling") the complex message. Related to complexity of information is the amount of time required to "tell" the information. If a more complex message takes longer to tell than a simple message, we would expect fewer tellings per person. Both complexity and amount of time appear to be indicating a resistance factor in passing on a message. In a similar light, the tabooness of the message topic also seems to contribute to the resistance in telling the information (Rogers, 1979; Rogers & Kincaid). Generalizing from both characteristics, there appears to be a general factor of resistance to telling a message.

Support of a resistance to transmit factor is the MUM effect (keeping mum about undesirable messages to the recipient) studied by Rosen and Tesser (Rosen & Tesser, 1970; Tesser & Rosen, 1975). They define the MUM effect as the "reluctance to communicate information that one could assume to be noxious for a potential audience" (Rosen & Tesser, 1970, p. 260). Thus, the tabooness of a message can be considered noxious for a potential receiver. Such a perceived message tends to be communicated less frequently and less quickly than non-noxious messages (Tesser & Rosen, 1975). An unobserved variable which could capture a general reluctance or hesitance to transmit will be referred to as the message transmission resistance.

It should be mentioned that the topic of the message influences the configuration of the network (Farace et al., 1977; Monge et al., 1978; Bavelas, 1948). Thus, for comparing the rates of spread for different messages in the same network, the topic of the message should be held constant.

In conclusion, no study was found which examines the influences of message characteristics on the rate of interactive information diffusion in a network. In addition, very little research defines (conceptually and operationally) the message characteristic of interest. In fact, in almost all of the literature reviewed, message characteristics are ad hoc discussions. However, two major dimensions of information can be identified from the literature. The two resulting variables considered important to the rate of interactive information diffusion are message temperature and message transmission resistance.

Summary

The first section of this chapter presented a theoretical background for interactive information diffusion. The research history of interactive information diffusion is closely tied to the research history of innovation diffusion. Epidemiology theory has provided a role model for theoretical development of interactive information diffusion. Alternatively, graph theory and related computer programs have contributed a methodology (i.e., network analysis) for studying the communication structure through which information spreads. Mass media research was the only area found to discuss message characteristics. However, there does not appear to be a theory which relates communication network variables and message variables to the variations in the overall rate of interactive information diffusion.

The section on mathematical modeling presented the deterministic logistic equation and its history. The logistic equation has been used by numerous researchers to represent the overall S-shaped accumulated growth behavior of interactive information diffusion. In the last part of the section, a strategy for modifying the logistic equation to include independent variables was presented.

The third and fourth sections of this chapter reviewed literature on network structure and message characteristics. From the discussion of network structure, two variables were selected for an overall description of a network:

1. Amount of network density, and
2. Length of network reachability.

The discussion on networks led to a discussion on the network location of the message's entry point (anchor). The following two variables were selected to describe the anchor point of the message:

3. Amount of anchor centrality, and

4. Length of anchor reachability.

In the discussion on messages, two distinct variables were developed for the message (or information) being diffused:

5. Amount of message temperature, and

6. Amount of message transmission resistance.

These six variables were selected to explain and predict the logistic equation b parameter for interactive information diffusion.

The next chapter presents a theory of the rate of interactive information diffusion composed of the above six exogenous variables.

FOOTNOTES--CHAPTER II

[1]The Gompertz equation is:

$$n_t = a^{b^t}$$

where n is the cumulative number of knowers or adopters and t is time (Feldman & Lie, 1964).

[2]The normal ogive equation is:

$$n_t = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-(u-t_0)^2/2\sigma^2)} du$$

where n_t is the cumulative number of knowers at time t, σ is the standard deviation of n, u is the population mean of n, e is the base of natural logarithms (Feldman & Lie, 1974). The rate equation for the normal ogive is the cumulative density function for the normal distribution.

[3]Transforming the logistic equation into simple linear regressions form (Brown, 1968):

$$n = \frac{k}{1 + e^{-(a + bt)}}$$

$$\frac{n}{k} = \frac{1}{1 + e^{-(a + bt)}}$$

$$k = n + ne^{-(a + bt)}$$

$$\frac{k-n}{n} = e^{(a + bt)}$$

$$\frac{n}{k - n} = e^{(a + bt)}$$

$$\log_e\left(\frac{n}{k - n}\right) = a + bt$$

$$n^* = a + bt$$

where $n^* = \log_e \left(\frac{n}{k - n}\right)$. The last equation is linear in parameters and can be used for OLS estimation of the parameters (a and b).

[4]The growth rate, dn/dt, of the logistic equation is the velocity (speed) of the information spread per time unit. Modeling dn/dt as a function of exogenous variables focuses on explaining the growth rate at each time interval for a single diffusion process. The research focus is on explaining the variations in the overall or summary growth rate resulting from many different diffusions, rather than on accounting for each time interval rate in a single diffusion. Thus, the research focus is on the logistic b parameter rather than on the first derivative (dn/dt) of the logistic equation.

[5]Message distortion will not be addressed in this paper.

CHAPTER III

THEORY

This chapter is divided into two sections: theoretical model and measurement models. The strategy for developing the theoretical model is to model the logistic b parameter as a function of exogenous variables (discussed in Chapter II). The purpose of the theory is to explain and predict variations in the logistic b parameter as a result of different messages diffusing in different networks. Therefore, the logistic b is written as a function of the six exogenous variables selected from the literature review. The theoretical model section provides: (a) conceptual definitions of the variables, (b) the bivariate relationship of each exogenous variable with the logistic b parameter, (c) the multi-exogenous equation representing the theory, and (d) the constraints of the theoretical model. The section on measurement models provides the operational definition and boundary condition for each observed variable.

Theoretical Model

Conceptual Definitions

The logistic b is a summary measure of the growth rate of the proportion of knowers over time. It is an estimated parameter in the logistic equation (1) and is the dependent variable to be explained in this study.

Network links. The first four variables below (network density, network reachability, anchor centrality, and anchor reachability) are based on communication network links. A network link represents a certain minimum level of regularly occurring direct communication activity between two network members (nodes).

Amount of network density is defined as the average amount of direct communication activity in a network.[1] This can be interpreted as the average degree to which network members are interlinked, i.e., the extent of interconnectedness (e.g., Farace, Monge, & Russell, 1977). Maximum network density would indicate that the network members are completely interconnected with maximum communication activity.

Length of network reachability is defined as the average minimum distance between all pairs of network members. The average geodesic (i.e., minimum distance) can be interpreted as the minimum diameter of the network (Bavelas, 1948; Harary, Norman, & Cartwright, 1965). Reachability takes into account all direct and indirect pathways the message can follow. If network density is at its maximum, then network reachability is at a minimum. However, the relationship between the two variables is not linear.[2]

Amount of anchor centrality is defined as the anchor person's average amount of out-going direct communication activity with all other network members. The anchor is the first network node to receive a message entering the network. Because the message radiates out from the anchor, anchor centrality is based on only out-going linkages. Anchor centrality reflects one node's average out-going communication activity; whereas, network density reflects the average amount of direct communication activity in the network as a whole.

Length of anchor reachability is defined as the average minimum distance from the anchor to all other network members.

Amount of message temperature is defined as the average level of arousal from the message among the members of the network.

Amount of message transmission resistance is defined as the average level of reluctance or hesitancy about passing on the message among the members of the network.

Bivariate Relationships

The bivariate relationships of the logistic b parameter with each of the above six exogenous variables (along with boundary conditions for the relationship) are given next:

> As network density (ξ_1) increases, the logistic b (η) increases. If network density is zero, there will be no interactive diffusion.

The overall growth rate of the accumulated proportion of knowers should be positively related to the average amount of direct communication activity in the network. That is, the greater the amount of direct communication activity, the greater the growth rate of knowers in the network. If there is little direct communication activity occurring, messages should take longer to spread through the network. By definition, if there is no direct communication activity among a set of people, there is no network, and therefore, interactive information diffusion is impossible.

As network reachability (ξ_2) increases, the logistic b (η) decreases (i.e., a slower rate of spread). As network reachability approaches infinity, the logistic b approaches zero, i.e., no diffusion.

The overall growth rate of accumulated knowers should be negatively related to the communication distance the message has to travel. Holding the speed of the message constant, the longer the average distance that the message must travel, the more time it will take for the message to spread through the network (i.e., speed is a measure of distance divided by time).

As anchor centrality (ξ_3) increases, the logistic b (η) increases. If anchor centrality is zero (i.e., an isolate is the anchor), there will be no interactive diffusion.

Anchor centrality is based on direct out-going communication activity of the first network message sender (i.e., anchor). Therefore, the more direct out-going communication an anchor has, the faster should be the initial growth rate of accumulated knowers, which should be reflected in a higher overall growth rate for the network. If an anchor has no direct out-going communication activity with any other network member (i.e., if s/he is an isolate), then anchor centrality will be zero and the message will not spread beyond the anchor person.

As anchor reachability (ξ_4) increases, the logistic b (η) decreases. If anchor reachability is infinite (i.e., an isolate is the anchor), there will be no diffusion.

Holding the speed of the message constant, the greater the average distance from the anchor to all other nodes, the greater the diffusion time required. Thus, the overall growth rate of the accumulated knowers will be slower. If the average communication distance from the anchor to all other network members is infinite, then there will be no interactive diffusion.

> As message temperature (ξ_5) increases, the logistic b (η) increases. If there is absolutely no message temperature, there will be no diffusion and, therefore, no b parameter to estimate. Also, if the message temperature is extremely high, the message could diffuse almost instantaneously which would most likely not produce the logistic curve.

The more excited, enthusiastic, or interested network members are about the information, the more people they are likely to pass it on to and the sooner they are likely to pass it on. Holding the network and anchor variables constant, increases in message temperature should result in a faster overall growth rate of accumulated knowers. If there is absolutely no excitement, enthusiasm, or interest in the information, people will have little inclination to pass on the information. However, if the message temperature is extremely high, the anchor of a small network will soon pass on the message to all other network members. This would likely produce some other curve than the logistic growth.

> As message transmission resistance (ξ_6) increases, the logistic b (η) decreases. Message transmission resistance probably has a critical point, beyond which no diffusion will occur.

The more people are reluctant to pass on the information, the fewer people they will tell and the longer the time delay between receiving and sending the information. Thus, the overall growth rate of accumulated knowers should be much slower. However, there is probably some upper level of resistance beyond which people will simply not tell one another. For example, an anchor person who is very sensitive to group norms may consider the message taboo and not pass it on to others. This appears to have been the case for family planning methods in many Korean villages prior to the national information campaign in the 1960's (Rogers & Kincaid, 1981). Even today in the U.S., abortion information does not circulate in certain networks. Thus, any amount of message transmission resistance will have the effect of slowing the rate of message spread and, therefore, slowing the growth rate of accumulated knowers.

Multi-exogenous Equation

Combining these six bivariate relationships produces the following multi-exogenous proposition for the logistic b parameter:

> The logistic b parameter (η) will increase as network density (ξ_1) increases, network reachability (ξ_2) decreases, anchor centrality (ξ_3) increases, anchor reachability (ξ_4) decreases, message temperature (ξ_5) increases, and message transmission resistance (ξ_6) decreases.

The multi-exogenous equation for the logistic b parameter is:

$$\eta = \gamma_0 \xi_1^{\gamma_1} \xi_2^{-\gamma_2} \xi_3^{\gamma_3} \xi_4^{-\gamma_4} \xi_5^{\gamma_5} \xi_6^{-\gamma_6} \zeta , \tag{4}$$

or, alternatively:

$$\eta = \gamma_0 \frac{\xi_1^{\gamma_1} \xi_3^{\gamma_3} \xi_5^{\gamma_5}}{\xi_2^{\gamma_2} \xi_4^{\gamma_4} \xi_6^{\gamma_6}} \zeta , \tag{5}$$

where γ_0 is a scaling constant, γ_i are parameters to be estimated, and ζ is the error of prediction.

The scaling constant, γ_0, could represent some external input (e.g., input by the mass media) where more than one network member receives the message from the environment.

Constraints

Based on the boundary conditions of the exogenous variables (e.g., greater than zero) and the bivariate relationships, the multi-exogenous equation representing the above proposition is written as multiplicative. That is, if any one of the positively related exogenous variables is zero or any one of the negatively related exogenous variables approaches infinity, there will be no interactive diffusion. Thus, the multiplicative form is required to estimate the equation's parameters. In equation (5), as the denominator increases, η will decrease. Alternatively, as the numerator increases, η will increase.

When exogenous variables are multiplicative, the error term is often also multiplicative rather than additive. Multiplicative error terms are frequently found to be heteroscedastic and nonnormally distributed (e.g., Danes, 1978; Laroche, 1977; Welch, 1978). Transforming both sides of equation (4) by taking logarithms will create an equation which is linear in parameters with an additive logarithmic error term, which is assumed to be normally distributed (i.e., Log-normal). Linearizing equation (4) by taking natural logarithms (ln) results in:

$$\ln\eta = \ln\gamma_0 + \gamma_1\ln\xi_1 - \gamma_2\ln\xi_2 + \gamma_3\ln\xi_3 - \gamma_4\ln\xi_4 + \gamma_5\ln\xi_5 - \gamma_6\ln\xi_6 + \ln\zeta. \qquad (6)$$

Substituting equation (4) into equation (1) yields:

$$\eta = \frac{k}{1 + e^{-[a + (\gamma_0\xi_1{}^{\gamma_1}\xi_2{}^{-\gamma_2}\xi_3{}^{\gamma_3}\xi_4{}^{-\gamma_4}\xi_5{}^{\gamma_5}\xi_6{}^{-\gamma_6}\zeta)t]}} \qquad (7)$$

The difference between equations (4) and (7) is that equation (7) allows the six exogneous variables to vary during any one diffusion study (i.e., one message in one network). Equation (4) constrains the exogenous variables to be constant (i.e., stationary) during any one diffusion study but allows them to vary across studies (i.e., across messages and/or across networks). Thus, equation (7) could be a nonrecursive model for a single message diffused in one network, while equation (4) is a recursive model for multiple messages and/or multiple networks. It may be the case that ultimately the diffusion of a single message in one network may lead to subsequent adoption of some new idea or technology, which, in turn, may lead to altering the network structure. However, the purpose behind developing the theory of the rate of interactive information diffusion is to account for the variations in the overall rates from the spread of different messages in different networks. For this purpose, equation (4) best represents the theory.

Measurement Models

The measurement model for each of the seven theoretical (unobserved) variables in equation (4) is presented below. The measurement model consists of both the operational definition and boundary condition for each observed variable. For all seven theoretical variables, the measurement model is a single indicator.

Logistic b parameter is estimated from the linearized logistic equation (3):

$$n^{*} = \hat{a} + \hat{b}t \qquad (3)$$

where $n^{*} = \ln(n/K - n)$ and t is time.

Strength of network links must be operationally defined before the measurement models for the two network and two anchor variables can be presented. For interactive information diffusion, link strength is operationally defined as the number of hours in a typical week that one person directly communicates with another person. The link strength scale is, therefore, bounded at zero and 168 hours (i.e., 24 hours times seven days). A link exists if the strength is greater than zero. This definition allows relatively weak links (i.e., weak ties) to exist along with strong links for tracing the paths of a message in a network.

The measurement models for network density, network reachability, anchor centrality, and anchor reachability are based on a value directed graph (vigraph) (Peay, 1980). A vigraph allows directed links with the link strength values retained. Two vigraph networks with the same configuration and network size can usually be distinguished from one another based on the link strength values. On the other hand, two ordinary graphs or two directed graphs with the same configuration and network size cannot be distinguished from one another. Basing the network measures on vigraphs allows a greater possibility of distinguishing between two networks with the same size and the same number of links. Thus, vigraphs are used for both network and anchor variables.

Amount of network density is operationally defined as:

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{k} S_{ij(adj)}}{n(n - 1)}, \qquad (8)$$

where $S_{ij(adj)}$ is the strength of the link between two adjacent nodes i and j (i.e., a direct link), $\sum_{i=1}^{n} \sum_{j=1}^{k} S_{ij(adj)}$ is the sum of rows and columns in the adjacency matrix (i.e., sum of all directional link strengths in the network), n is the number of nodes, and $n(n - 1)$ is the number of off-diagonal elements in the adjacency matrix.

Equation (8) is an extension of the density measurement model used for ordinaty graphs (e.g., Czepiel, 1975; Friedkin, 1981; Granovetter, 1976; Richards, 1976) and directed graphs (e.g., Edwards & Monge, 1977; Guimareas, 1972; Rogers & Kincaid, 1981) where the strength of actual links is summed rather than the number of links. The density measure can be interpreted as the average direct communication activity of the average direct link strength between any two adjacent nodes in a network. Dividing $\sum_{i=1}^{n} \sum_{j=1}^{k} S_{ij(adj)}$ by $n(n - 1)$

results in an average score, allowing for the comparability of networks of differing sizes (Freeman, 1978/79). Based on link strength scale of zero to 168, amount of network density can range between zero and 168. A maximum density score of 168 means that all the nodes in a network have maximum direct communication activity (i.e., 168 hours per week) with all other nodes. While a maximum network density score of 168 is pragmatically impossible, it is theoretically possible.

Length of network reachability is defined as the average geodesic (shortest distance) from node i to node j. For ordinary or directed graphs, the geodesic is the minimum number of links or steps from node i to node j (Harary, Norman, & Cartwright, 1965; Peay, 1980). The length of network reachability for a directed graph (digraph) is (e.g., Edwards & Monge, 1977):

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{k} L_{ij}(\text{dist})}{n(n - 1)}, \quad \text{if } i=j,\ L = 0, \tag{9}$$

where $L_{ij(\text{dist})}$ is the minimum number of links from node i to node j in the distance matrix.

For a metric distance between two adjacent nodes in a communication network, Farace and Mabee (1980) suggest the inverse of the link strength ($1/S_{ij}$). The assumption is that the stronger the link strength, the shorter the communication distance. Alternatively, the weaker the link strength, the greater the communication distance.

Taking the inverse of the link strength converts strong link strengths into short distances and weak link strengths into long distances.

The link strengths (S_{ij}) are retained in the adjacency matrix for the computation of the amount of network density. Thus, extending the operational definition from digraphs to vigraphs, the length of network reachability is defined as:

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{k} 1/S_{ij(dist)}}{n(n - 1)}, \text{ if } i=j,\ 1/x = 0, \qquad (10)$$

where $1/S_{ij(dist)}$ is the distance (inverse of link strength) from node i to node j in the distance matrix, and n(n-1) is the number of finite off-diagonal elements in the distance matrix.

On close examination, there is a problem with equation (10) and equation (9) for directed graphs. Since digraphs and vigraphs do not assume symmetrical links, it is possible to have a directed path from node i to node j, but no path from node j back to node i. Thus, node j is reachable from node i, but node i is not reachable from node j. When node i is not reachable from node j, the distance is infinite (Barnes, 1969; Harary et al., 1965). An infinite distance for a cell element in the distance matrix results in an incomplete matrix of finite numbers. Since zero on the main diagonal represents the distance from every node to itself (Harary et al.), zero would be inappropriate to represent infinite reachability. Using equation (9) and (10), i.e., dividing by n(n-1), implies a full matrix of finite numbers.

The number n(n-1) for a full matrix represents the number of off-diagonal finite elements for computing an average (mean) reachability score for a network. Therefore, n(n-1) can be generalized to Σf(D), which is the sum of the number of off-diagnoal finite elements in the distance matrix.[2] Replacing n(n-1) with Σf(D), equation (10) for a vigraph becomes:

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{k} 1/S_{ij(dist)}}{\Sigma f(D)}, \quad \text{if } i = j,\ 1/S = 0. \qquad (11)$$

Using Σf(D) as the denominator restricts the interpretation of the resulting score to the average geodesic of the nodes that are reachable.

Based on link strengths of one to 168 for existing links, the distance between two adjacent nodes can range from .0060 (1/168) to one (1/1). The length of network reachability can range from greater than zero to infinity. A zero or infinite score would imply that no interactive diffusion of a message occurred. Although, as the amount of network density increases, the length of network reachability decreases, the relationship between the two variables is nonlinear. The inverse of link strength for computing distance is a monotonic nonlinear transformation. In addition, the sum of the inverse of each link is not the same as the inverse of the link strength (i.e., $\Sigma 1/S \neq 1/\Sigma S$). Therefore, network reachability is a different structural measure than network density and the reachability score cannot be directly derived from the density score, i.e., they are not multicollinear.

Amount of anchor centrality is operationally defined as:

$$\frac{\sum_{j=1}^{k} S_{ij(adj)}}{(n - 1)}, \qquad (12)$$

where $\sum_{j=1}^{n} S_{ij(adj)}$ is the sum of the out-going link strengths from the anchor (node i) to all other nodes in the adjacency matrix and (n-1) is the number of other network members.

Like the amount of network density, the amount of anchor centrality score can range from zero to 168. Thus, the amount of anchor centrality score can be interpreted as the anchor's average number of hours per week of out-going direct communication with other network members.

Length of anchor reachability is operationally defined as:

$$\frac{\sum_{j=1}^{k} 1/S_{ij(dist)}}{(n - 1)}, \quad \text{if } i=j,\ 1/S = 0, \qquad (13)$$

where $\sum_{j=1}^{k} 1/S_{ij(dist)}$ is the sum of the out-going distances (i.e., inverse link strength) from the anchor (node i) to all other network members.

Because a message will spread out from the anchor, all other network nodes that ultimately receive the message will be reachable from the anchor. The length of anchor reachability score can range from greater than zero to infinity. Both a zero or infinite score would imply that an isolate (i.e., a node with no communication links to other network members) is the anchor and, therefore, no interactive message diffusion has occurred. (A zero score represents the distance to oneself while a score of infinity represents the unreachability to other nodes.)

Amount of message temperature is operationally defined as the average amount of excitement, enthusiasm, or interest a message generates in the network. The measurement scale and general question to be asked of network members are:

> If 0 (zero) is NOT AT ALL, and 100 is AVERAGE:
>
> How EXCITED, ENTHUSIASTIC, or INTERESTED are you in the message news?

The above is a ratio scale bounded at zero with 100 given as a modulus (reference point). To create the average score for the people in the network, individual members' scores are summed and divided by the number of network members (N).[4] Thus, amount of message temperature scores for networks can range from zero to infinity. A score of less than 100 would indicate a network's excitement, enthusiasm, or interest in the message is less than average.

Amount of message transmission resistance is operationally based on the same measurement scale used for amount of message temperature above with the following question:

> How HESTITANT (for any reason) are you to pass the information on to someone you know might be interested in the news?

To create the average score for the network, individual members' scores are summed and divided by the number of network members (N). Amount of message transmission resistance scores for networks can also range from zero to infinity.

Single indicators for the two message variables were chosen based on the study's design and the sample which are presented in the next chapter.

Summary

This chapter has presented a theory (model) of the rate of interactive information diffusion. As part of the theory, both the theoretical and measurement models were presented. The theoretical model consists of a single multi-exogenous equation with the logistic b parameter as the dependent variable. The theory specifies that the functional form of exogenous variable relationships is multiplicative, or alternatively, the linearized form is the log-log equation.

The measurement model for each variable is a single indicator. The measurement models for amount of network density, length of network reachability, amount of anchor centrality, and length of anchor reachability are all based on the analysis of value directed graphs (vigraphs). Analysis of vigraphs allows the greater precision in distinguishing between networks and between anchors than vigraphs. However, to allow for a possible incomplete distance matrix (i.e., nodes of infinite reachability), the often used formula for network reachability was modified by replacing n(n-1) with the more general $\Sigma f(D)$ in the denominator.

The next chapter will present a study conducted to test the theory presented in this chapter. The research is a non-experimental field study of the diffusion of a single message in multiple networks.

FOOTNOTES--CHAPTER III

[1]For brevity, all exogenous variables are referred to by their key identifying name, rather than their full name after they are conceptually defined, e.g., amount of network density is subsequently referred to only as network density.

[2]This will be explained in the section on measurement models.

[3]Two other possible alternatives are to replace the infinite element with: (a) the median reachability between nodes i and j or (b) some finite upper limit value (e.g., the maximum reachability between nodes i and j) (personal conversation with Edward L. Fink, January 1982). The difficulty with these two alternatives is that the reachability score will be biased upwardly. For a distance matrix with more than 25% infinite elements, the upward bias could be quite substantial. However, until an empirical comparison is made between these two alternatives and $\Sigma f(D)$, no one alternative is clearly superior.

[4]Empirically, the mean may not be the best aggregation if the scores do not approximate a normal distribution. However, using the mean for message temperature and message transmission resistance is consistent with the measures of network density, network reachability, anchor centrality, and anchor reachability which are also mean measures.

CHAPTER IV

METHODS

To test the theory of interactive information diffusion, i.e., equation (6), a non-experimental field study was conducted. The same message was diffused in multiple networks. For the first test of the theory, it was decided to hold constant the amount of message temperature and the amount of message transmission resistance while allowing the other four exogenous variables to vary. Holding the two message variables constant allows for better parameter estimates (i.e., smaller standard errors) for the other four exogenous variables with the same sample size. Allowing all six exogenous variables to vary would require a very large sample of networks and would be a very complex study to conduct and analyze. Thus, the study's major focus is to test the functional form of the theory and estimate the parameters of the amount of network density, length of network reachability, amount of anchor centrality, and length of anchor reachability.

Groups/Subjects

To test equation (6), groups of dormitory residents from the Michigan State University campus were contacted. The criteria for selecting groups were: (a) maximum group size of 100, (b) anticipated cooperation with the researcher, (c) sufficient probable variability of the two network and the two anchor variables for parameter estimation, and (d) topic interest commonality among the groups so that the same message could be reasonably used for all of the groups.

The dormitory living groups on the MSU campus appeared to meet these four criteria. Each dormitory group was composed of a resident assistant (RA) and approximately 50 residents (hereafter referred to as an RA-group). RA-group members live on the same wing and floor of a dormitory.

The dormitory system at MSU is one of the largest in the country. Students are housed in several different dorm structures. The floor structure of the dorms are of three basic types: (a) straight hallways, (b) single jointed hallways (45° and 90° joints) where only half of the hallway can be seen at a time, and (c) curved hallways where only one-fourth to one-third of the hallway can be seen at a time. In addition, while most dorm floors are divided into suites with a bath between two rooms, several other dorms have a community bath for each floor/wing. Because the different hallway structures and bath arrangements could reasonably affect the communication patterns on the floors, it was felt that RA-groups from different dorms would have a high probability of exhibiting variable network structures (i.e., network density and network reachability).

All but one dorm on MSU's campus houses undergraduate students. Some topic interest commonality was expected for undergraduate dorm residents as a whole. It was felt that dorm residents would cooperate in the study if: (a) the message to be diffused was of sufficient interest to dorm residents, (b) the procedures were made as simple and as effortless to carry out as possible, and (c) participation time in the study was kept to no more than 10 minutes. Therefore, because RA-groups appeared to meet all of the above criteria, they were selected as the sample units of analysis in the study.

For all groups, the RA was chosen to be the first person to receive the message (anchor node). This was done for two reasons. To conduct a "survey" on each dorm floor, the permission and cooperation of the RA was needed. Second, RAs are frequently the initial sources of information which circulates in the dorms and, therefore, it was felt that the dissemination of the message would more likely follow the usual pattern of flow.

Permission to conduct a survey in the dorms was obtained by contacting the Director of University Housing Programs. The Director was informed of the nature of the study and the topic of the message to be diffused. An offer was made by the researcher to conduct a training workshop on communication networking for participating dorm directors and RAs. Because of the offer, the Director of University Housing Programs deferred the decision to his six Area Dorm Directors. One Area Director gave permission to contact her respective five dorm directors (again, this was because of the training workshop offer). All five dorm directors gave their permission to contact RAs. During the meeting with each dorm director, the researcher left copies of a two-page handout for all RAs in the dorm. The handout contained a very brief description of the study with a list of advantages to participating in the study (see Appendix A).

A week after meeting with the five dorm directors, RAs were contacted by telephone and asked if they would be willing to participate. RAs were informed about the message topic. It was emphasized that participation by an RA did not obligate other floor residents to participate. An RA's participation allowed other residents the opportunity to receive the information and instructions. Residents could decide for themselves whether to participate by completing a short questionnaire and passing the message on to other floor residents. If the RA agreed to participate, a meeting time was set for the researcher to deliver the materials to the RA. Out of 48 RAs contacted, five refused to participate. A summary description of the 43 RA groups from the five dorms is presented in Table 3. Ten groups with three or fewer participants in the study were dropped from further analysis.

Message

The criteria for selecting the single message were: (a) the message should be highly interesting to the majority of undergraduate dorm residents, and (b) there would be little resistance or reluctance by undergraduate dorm residents to pass the message on to others. A message which meets these two criteria should maximize undergraduate dorm residents' willingness to participate in the study.

The message selected was a list and schedule of the winter term RHA (Resident Hall Association) movies. During the week and on

TABLE 3. Summary Description of Dorms and Participating RA-Groups

Dorm	Design (hallway)	Bath	Male RA-groups (dropped)	Female RA-groups (dropped)	Coed RA-groups (dropped)	Total # of RA-groups (dropped)	RA-groups with 4 or more respondents
Akers	curved	suite	5 (5)	7 (0)	0	12 (5)	7
Holmes	straight	suite	2 (0)	5 (1)	2 (0)	9 (1)	8
Hubbard	straignt	suite	2 (0)	5 (1)	0	7 (1)	6
McDonel	45^{o} jointed	suite	4 (0)	4 (1)	2 (0)	10 (1)	9
Shaw	90^{o} jointed	community	2 (2)	3 (0)	0	5 (2)	3
TOTAL			8	21	4	43	33

Note. An RA-group was dropped from the sample and analyses if there was not a minimum of four participants in the group.

weekends there are one to three movies shown at any one time on campus. They are free to dorm residents. The RHA movie schedule for each term is usually publicized at the beginning of the term. In the middle of fall term, the researcher contacted the RHA Movies Office and was given an advance schedule of movies for winter term. Because the RHA movies are free to dorm residents and the winter schedule contained several movies which were new releases and currently showing in the community theaters, it was felt that most undergraduate dorm residents would find the advance information quite interesting (see Appendix A for the schedule of movies). It was also anticipated that there would be very little resistance to pass the schedule of movies on to others.

Pilot Test

Before conducting the study, a pilot test was conducted to test the procedures and questionnaire. One RA group in one of the dorms was used for the pilot test. A message concerning a Sunday evening pot-luck dinner with a guest speaker in the floor's lounge (meals are not served in the dorms on Sunday evenings) was disseminated.

As a result of this test, several procedures for the study were changed. First, requesting participants to place the completed short questionnaire in the campus mail slot had problems. It was discovered that undergraduate college students are not aware of the campus mail system and/or do not know where the campus mail slot is in their dorm. Using campus mail also meant that participants had to remember to deposit the questionnaire upon leaving the main entrance of their dorm. As an alternative, it was decided to have participants return the questionnaire to a large envelope stapled on the dorm floor's bulletin board located just inside the entrance of each floor.

A change was also made in the layout of the message and questionnaire which were both photoreduced and placed on the same single sheet of paper along with the cover letter and instructions. For the study, the layout was changed so that the single sheet could be cut or torn in half. This would allow participants to keep the half with the message (schedule of winter term RHA movies) and return the half with the completed questionnaire to the researcher.

Based on the pilot test, it was decided that the cover letter (half of one side of the single sheet) needed to be shorter and more interesting to encourage undergraduate student participation. Also, the cover letter needed to be more eye-catching. As a result, the cover letter was shortened so that it did not require photoreduction. In addition, a news bulletin-type heading was added and two newspaper movie ads for two of the movies were placed at the top of the cover letter.

Materials

The materials for each RA group consisted of two large envelopes (10" x 13") and approximately 125 copies of the single sheet containing the message and questionnaire. One envelope (labeled "RETURN" and the RA's name) was placed on the RA group's bulletin board for returning completed questionnaires. The second envelope (labeled with the RA's name only) contained about 125 copies of the message-questionnaire and was taped to the RA's room door.

The message-questionnaire was a single sheet containing the cover letter and instructions on one side and the message and questionnaire (eight questions) on the reverse side (see Appendix A). The cover letter with the two newspaper movie ads at the top was left the original size. Both the instructions to participants and the questionnaire were photoreduced to make each one fit on one-half of one side of the single sheet. The message was placed on half of one side of the sheet. The message consisted of a photoreduced schedule of the winter term RHA movies with seven full-sized newspaper movie ads along the top, left side, and bottom of the half side. The seven ads were for seven of the movies listed in the schedule.

Procedures

As previously mentioned, the RA was selected as the diffusion starting point (anchor node) for each RA-group. The researcher met with each of the 43 RAs in their dorm rooms to give them the message and to explain the instructions. In addition, each RA was given a separate cover letter which emphasized that the study was not a test or evaluation of the RA's communication behavior and that the study's focus was on the dorm floor's overall information-exchange pattern. Because RAs frequently disseminate information to all floor members, it was emphasized that RAs were to follow the instructions on the message-questionnaire sheet. The instructions were to give a copy of the message to only those floor residents whom they would normally pass on informal information to.

RAs were also told that if a floor resident inquired about the envelope on their door or asked for a sheet, the RA could give the person a copy. While RAs are frequent disseminators of information, they are also considered opinion leaders and asked for information. Whether the RA or a floor resident initiated the interaction was considered to be of minor importance since the message would be considered as moving from the RA to the resident in either case. The two anchor variables (anchor centrality and anchor reachability) are based on the directional movement of the message.

The data and time the message was given to each RA was recorded and used as the start time (t_0) for the diffusion process in each RA-group.

The written instructions to participants were as follows:

A. In person, pass a copy of the Winter RHA Films on to only those people on your floor (same RA as you) that you would normally tell this information to.

B. Extra copies of the Winter RHA Films are available in a large envelope on your RA's door, labeled with your RA's name only. Take only the number of copies you need to give to others.

C. Within the next 24 hours (if possible), in person hand copies of the Winter RHA film schedule to the other people you would normally "tell" this information to. (Remember: only people who have the same RA as you.)

D. If someone tells you they have already received the winter film schedule, say to them: "Take it again and answer the questions. Then place it in the large envelope on the BULLETIN BOARD marked RETURN."

E. After answering the questions (on the other side of this sheet), CUT this sheet in half along the dotted line. Then, place THIS HALF with the questions in the large envelope on the BULLETIN BOARD for your floor. The envelope is labeled with "RETURN" and your RA's name. (The remaining half with the Winter RHA Films is yours to keep.)

F. IF you have received this sheet before, answer the questions and place this sheet in the large envelope on the BULLETIN BOARD for your floor.

*NOTE: You may be the receiver of this film schedule and questions many times, but remember you only distribute the film schedule once to other people.

Both the envelopes on the RAs' doors and floor bulletin boards were checked daily to ensure that a sufficient number of message-questionnaires were available and to gather completed questionnaires. During the data-gathering process, one problem arose. First, late on Friday and Saturday nights, several of the envelopes on the bulletin boards in one of the dorms were removed. However, on the following days, the envelopes were replaced by the researcher. Because very limited diffusion activity occurred in the other floors during the two weekends of the study, it was felt that very few, if any, completed questionnaires were lost because of missing "return" enveloped.

Timing

To maximize the opportunity for interactive information diffusion to occur among dorm residents, timing the start of the message's diffusion was important. First, it was recognized that many dorm residents leave the campus each weekend during which little diffusion activity could be expected to occur. Thus, the message was distributed to RAs only on Monday through Wednesday on two consecutive weeks. This allowed several days for the message to circulate on a dorm floor before residents left for the weekend.

The second time consideration was the term schedule. At the beginning of fall term, the majority of dorm residents are new first-year students and their floor communication networks are not yet developed. During the first part of fall term, RAs schedule various activities designed to acquaint residents with one another. By fall midterm, it was expected that floor networks would be fairly well established for the purpose of the study.

The term schedule also required three additional time considerations: (a) midterm week, (b) the four-day Thanksgiving break, and (c) finals week. While information does circulate during midterm examinations week, passing the message about winter term RHA movies would likely have received a very low priority by dorm residents. The four-day Thanksgiving break was 2½ weeks after midterm week and one week before finals week. For the Thanksgiving break, the university expects all dorm residents to leave campus. Thus, no diffusion activity would be possible during this time. In addition, little diffusion activity could be expected during the week before and the week of final examinations. Thus, the optimum time during fall term to introduce and track the message was the 2½-week period between midterm week and Thanksgiving break. Based on the term schedule and weekend considerations, the message was distributed to RAs on Monday through Wednesday of the first and second weeks following midterms. Fifteen RAs received the message two weeks prior to the

Thanksgiving break while the remaining 28 RAs received it 1 to 1½ weeks prior to the break.

Measurement Scales

The general measurement model for each variable was presented in the previous chapter. Therefore, only the measurements specific to the study are presented here.

Logistic b

The logistic b is estimated for each network based on the cumulative proportion of knowers over time. Time was measured as the number of minutes it took the message to spread from the RA (t_0) to respective floor members (t_i). The date and time the message was given to the RA was recorded by the researcher. Subsequent receivers of the message were asked to indicate on the questionnaire the date and time they received the message.

Network Links

Dorm residents who receive the message were asked to name the person who gave it to them. This question established the presence of a pathway and the direction that the message took. To measure the strength of the communication link, participants were asked:

> HOW MANY HOURS do you communicate with the person named in QUESTION #3 [sender] IN A TYPICAL WEEK (roughly estimate to the nearest hour)?

Thus, the directional pathway is based on a single event, i.e., the message, while the path (link) strength is based on a pre-existing very general class of events.

Network Variables

Network density and the network reachability were computed according to equations (8) and (11), respectively.

Anchor Variables

Anchor centrality and anchor reachability were computed according to equations (12) and (13), respectively.

Message Temperature

For the two message variables (message temperature and message transmission resistance), respondents were instructed to:

> Evaluate the RHA film news by using the following scale or "yardstick." You may use any positive number (including 0 and numbers larger than 100) that best represents your feelings.

The scale and an example were:

> If 0 (zero) is NOT AT ALL, and 100 is AVERAGE, then:
>
> EXAMPLE: if your interest in the news of RHA films is twice as much as average, you would write "200" in the space provided.

To measure message temperature, the participants were asked:

> How EXCITED, ENTHUSIASTIC, or INTERESTED are you in the news of the RHA movies for Winter term?

Message Transmission Resistance

Message transmission resistance used the above scale and participants were asked:

> How HESTITANT (for any reason) are you to pass the information on to someone you know might be interested in the news?

In addition, participants were asked their name, dorm, RA's name, how many terms they had lived in the dorm, and whether or not it was the first time they had received the message.

Design Constraints

The study's design imposes five constraints which need to be addressed. First, for the purpose of this study a network is defined as the residents of a dorm floor who receive the message and return the questionnaire. A dormitory as a whole could be viewed as a single network. However, it was determined from interviews with dorm directors and several RAs that there was very little communication interaction between floors in most of the dorms. The exception is communication between brother and sister wings on the same floor in only one of the five dorms. Thus, because communication among dorm residents is relatively restricted to each floor and only five dorms were made available to the researcher, each network is restricted to a single dorm floor of residents. To ensure this boundary for all dorm floors, each message receiver was instructed to pass the message on to only same floor residents (i.e., residents who have the same RA as the sender). Because each RA-group (dorm floor) consists of 50 residents, the maximum size that any network can be is 50.

Although not a constraint per se, it is relevant to mention here that the often used group or network membership definition (e.g., Richards, 1975, 1976) was relaxed. When a network is based on communication about a pre-existing class of events, a group or network member is a node which has at least two linkages in the network. To tract the diffusion of a single message in multiple networks, the networks are based on a single event, i.e., transmission of the Winter Term RHA movie schedule. Any person in an RA-group who received the message (and returned a questionnaire) is considered a network member. Thus, for networks based on a single event, the network membership definition is relaxed to include people with only one linkage.

The second design constraint is that the exogenous variables are constrained to be fixed for each network during the message's diffusion. That is, the exogenous variables are not allowed to vary for an individual network during the diffusion process. The design calls for a single measurement on each variable for each network.

The third design constraint restricts the two message variables (message temperature and message transmission resistance) to be relatively constant. The diffusion of the same message in a relatively homogenous sample (i.e., undergraduate dorm resident groups from the same dorm complex) was intentionally designed to hold the two message variables relatively constant. If the two message variables can be considered as constants (to be determined in the data analyses), then only four instead of six variable parameters need to be estimated from the relatively small sample (N = 33 networks). Eliminating two parameters should decrease the standard errors for the other four parameters and increase the likelihood of significant results.

If the two message variables are considered constant, then the intercept in equation (6) becomes

$$\ln\gamma_{0_1} = \ln\,(\gamma_0\xi_5\xi_6), \tag{14}$$

where γ_{0_1} is a new intercept composed of the original constant and the constant effects of the two message variables.[1]

The attempt to hold the two message variables constant leads to the fourth design constraint of single indicator measurement models for the two message variables. Since the study's design restricts the variability of the two message variables (i.e., diffusion of a single message), multiple indicators of the two variables would likely result in multiple constants. Although the same problem exists for the single indicators, they are included in the set of measurements for the study to provide some information about the message.

Given the study's design and the sample selected (i.e., undergraduate dorm residents), using single, rather than multiple, indicators for the two message variables is more practical. For the administration of the study, it was considered necessary to keep the questionnaire length to one-half of one side of a single sheet. Even with photoreduction, this did not allow for multiple questions on the same exogenous variable. In addition, the single sheet with the questionnaire was designed to maximize dorm residents' visual impressions to encourage participation (e.g., "It <u>looks</u> like it won't take much of my time.").

The fifth design constraint is that the study is based on an opportunity rather than a random sample. Conducting a field study on the diffusion of a single message in multiple networks is not readily condusive to random sampling. In addition, MSU's policy on "surveying" dorm residents prohibits random sampling of dormitory RA-groups. There are two potential problems with opportunity sampling. The first potential problem is the increased possibility of multicollinearity among the exogenous variables when, in fact, they are uncorrelated in the population. However, this possibility may also occur with a small random sample. The second problem is that the results from an opportunity sample are not as generalizable as the results from a random sample. However, the results from the opportunity sample used in this study will provide initial information about the theory's functional form and the utility of the selected exogenous variables.

Analyses to be Conducted

The data analyses will consist of three major stages: (a) estimation of the logistic b parameter for the message diffusion in the networks, (b) computation of each network's density and reachability scores, each anchor's centrality and reachability scores, and the mean message temperature and mean message transmission resistance scores for each network, and (c) testing the theory by using the results from (a) and (b) as the data to test the theory, i.e., equation (6).

The third stage of the data analyses (test of the theory) will consist of testing the functional form of the theory, parameter estimation for the selected appropriate functional form, and a residual analysis to confirm the functional form selected. The functional form of the theory as presented in equation (6) is expected to be logarithmic (ln). This will be tested relative to other functional forms by the log-maximum likelihood criterion (L_{max}) and chi-square test. In other words, it will not be assumed that the residuals resulting from linear regressions on the linearized form of equation (6) will be homoscedastic and normally distributed.

The logistic b parameter (dependent variable) can theoretically vary from zero to infinity (i.e., a single bounded continuous scale). Therefore, single bend transformations can be appropriately used to rescale the variable to meet the assumptions of the general linear model. Also, the two message variables (message temperature and message transmission resistance) are measured by a single bounded scale and, therefore, can be appropriately subjected to single bend transformations. The amounts of network density and anchor centrality are measured by a double bounded scale (i.e., link strength is bounded at zero and 168). However, the two scales can be considered to approximate a single bounded scale since only a very limited few observed values approach the upper limit of the scales. Since network reachability and anchor reachability are computed from the inverse of link strength, they also may be considered to approximate a single bounded scale subject to single bend transformations. The set of exogenous variables will be subjected to single bend transformations to test for additivity.

The set of single bend transformations to be used on all of the variables will be the Box-Cox power family defined as (Box & Cox, 1969; Huang & Moon, 1978):

$$\underline{Y}^* = \begin{cases} (Y_i^\lambda - 1)/\lambda & \lambda \neq 0 \\ \ln Y_i & \lambda = 0 \quad i = 1, 2, \ldots N \end{cases} \tag{15}$$

and

$$\underline{X}_k^* = \begin{cases} (X_{ki}^\mu - 1/\mu & \mu \neq 0 \\ \ln X_{ki} & \mu = 0 \text{ for any } k,\ i=1, 2, \ldots N \end{cases} \tag{16}$$

where $\underline{Y}^*$ is the vector of the independent variable transformed according to equation (15), $\underline{X}^*$ is the matrix of independent variables transformed according to euqation (16), and λ and μ are transformation parameters to be determined. The criterion for selecting the optimum transformation parameter (λ and μ) will be the transformation resulting in the largest log-maximum likelihood (L_{max}) value. The L_{max} for each transformation is computed as (Box & Cox, 1969; Huang & Moon, 1978):

$$L_{max}(\lambda,\mu) = \text{constant} - \frac{N}{2}\ln\sigma^2(\lambda,\mu) + (\lambda-1)\Sigma\ln Y_i \tag{17}$$

where $\hat{\sigma}^2(\lambda,\mu)$ is the variance of estimates.

Alternative appropriate functional forms (i.e., functional forms other than the one resulting in the optimum L_{max}) can be determied by constructing a confidence region around the optimum estimates $(\hat{\lambda},\hat{\mu})$. An approximate 95% confidence region can be constructed around the estimates $(\hat{\lambda},\hat{\mu})$ because:

$$2[L_{max}(\hat{\lambda},\hat{\mu}) - L_{max}(\lambda,\mu)] \tag{18}$$

is approximately distributed as χ^2 with two degress of freedom (Huang & Mood, 1978). Thus,

$$L_{max}(\hat{\lambda},\hat{\mu}) - L_{max}(\lambda,\mu) \leq \tfrac{1}{2}\chi^2_{2},\ \alpha \tag{19}$$

may be used to test a joint hypothesis on λ and μ of a specific functional form such as the untransformed form (i.e., H_0: $\lambda = 1$ and $\mu = 1$) or in the case of the interactive diffusion of information, the logarithmic form (i.e., H_0: $\lambda = 0$ and $\mu = 0$).

To test the logarithmic functional form of the theory and to search for the otpimum functional form, the B-C computer program will be used. The B-C program: (a) performs a set of single bend transformation (i.e., Box-Cox family) on both the dependent and set of independent variables, (b) computes the L_{max} for each transformation, and (c) computes OLS regression results for each transformed equation (i.e., γ_i and $\overline{R}^2$) (Huang & Moon, 1978).[2]

After selecting the appropriate functional form, a residual analysis will be conducted to confirm the selected functional form.

The residual analysis will consist of examining the skewness and kurtosis of the transformed dependent variable (Y^*) (to examine the normality assumption) and examining a residual scatterplot to evaluate the assumptions of both normality and homoscedasticity of Y^*.

FOOTNOTES--CHAPTER IV

[1]Equation (6) is:

$$\ln\eta = \ln\gamma_o + \gamma_1\ln\xi_1 - \gamma_2\ln\xi_2 + \gamma_3\ln\xi_3 - \gamma_4\ln\xi_4 + \gamma_5\ln\xi_5$$

$$- \gamma_6\ln\xi_6 + \ln \zeta.$$

If ξ_5 and ξ_6 are constants, then equation (6) becomes:

$$\ln\eta = \ln\gamma_{o_1} + \gamma_1\ln\xi_1 - \gamma_2\ln\xi_2 + \gamma_3\ln\xi_3 - \gamma_4\ln\xi_4 + \ln \zeta,$$

where

$$\ln\gamma_{o_1} = \ln\gamma_o + \ln\xi_5 + \ln\xi_6,$$

or

$$\ln\gamma_{o_1} = \ln (\gamma_o\xi_5\xi_6).$$

[2]The CDC Cyber 750 version of the B-C computer program is available from the author.

CHAPTER V

RESULTS

The units of analysis for testing the theory of interactive diffusion of information are pre-existing groups of dormitory residents. However, the variable values for each network must first be computed from the data gathered on each network member. That is, all of the variables in the theory are derived measures from individual data. Hence, the first results are the variable scores for each network and the descriptive statistics for the set of variable values. Also included in the descriptive statistics section are the reliability and validity assessments of the two message measurement scales. The second set of results is the test of the theory's logarithmic functional form. Descriptive statistics for the logarithmic (ln) transformed variables are also provided. The third set of results consists of a residual analysis to confirm the appropriateness of the selected functional form. The fourth set of results is the parameter estimates for the exogenous variables in the theory.

Descriptive Statistics

The logistic b parameter (dependent variable) is estimated for each network by linear regression on equation (3). The results for each network are presented in Table B1 in Appendix B. Since the logistic equation is the cumulative proportion of knowers by time, the sample size (T) is the number of different time points and not the network size.[1] T in Table B1 ranges from 3 to 21. While three different time points would be insufficient to describe a double bend curve (i.e., four points are necessary), the linearized form of the logistic equation (i.e., a straight line) requires at least two points. Therefore, while a T of three is exceedingly small, it is sufficient for estimating parameters of a linear equation (i.e., linear in parameters). Time was measured in minutes with t_0 (the time the RA (anchor) received the message) assigned a value of zero.

The logistic b in Table B1 is a summary rate of how fast the message diffused in each network. The correlations (r) in Table B1 are a measure of how well the diffusion data for each network (cumulative proportion of knowers vs. time) fit the logistic S-shaped curve. The correlations ranged from .526 to .994 with 24 of 33 networks having an r significant at $p \leq .05$. Given the very small T (number of time points) for all of the sample networks, it was expected that some of the correlations would not be significant at $p \leq .05$.[2]

To determine if the standard error (S.E.) of the logistic b decreases with an increasing T, the logistic b and its S.E. were ordered according to the size of T in Table 4. The ratio of the logistic b to its S.E. (i.e., *t* value) was then computed. It is apparent that for a sample size range of three to 21, the S.E. of b does not systematiclly decrease with a larger T.[3] The smaller the

TABLE 4. Logistic $\hat{b}$ and S.E. of $\hat{b}$ as T Increases

Network	T	Logistic $\hat{b}$	S.E. of $\hat{b}$	$\hat{b}$/SE Ratio
16	3	.005544	.001109	5.00
29	3	.034598	.002367	14.62
17	4	.077288	.090065	.86
09	4	.001639	.000339	4.83
31	4	.002499	.002861	.87
08	5	.001536	.000157	9.78
10	5	.001209	.000094	12.86
18	5	.000949	.000577	1.64
22	5	.029596	.002830	10.46
23	5	.002646	.002115	1.25
26	5	.005338	.000553	9.65
33	5	.004732	.000498	9.50
02	6	.000566	.000438	1.29
21	6	.001303	.000160	8.14
03	7	.003466	.001312	2.64
05	7	.020881	.013216	1.58
24	7	.003056	.000518	6.87
27	7	.003057	.001795	1.70
30	7	.000934	.000124	7.53
32	7	.002544	.000973	2.61
12	8	.010292	.003857	2.67
13	8	.004800	.000929	5.17
14	8	.003791	.000537	7.06
20	8	.001349	.000491	2.75
25	8	.007658	.000398	19.24
06	9	.002853	.000306	9.32
28	10	.000711	.000370	1.92
19	11	.000438	.000049	8.94
01	12	.000331	.000040	8.28
04	12	.002996	.000451	6.64
07	12	.003542	.000805	4.40
11	15	.000377	.000066	5.71
15	21	.000195	.000049	3.98

Note. T is the number of different time points.

standard error relative to the parameter's estimated value, the more consistent the estimate is (Hamushek & Jackson, 1977). The standard error of the parameter is the standard deviation of the distribution of that parameter. To evaluate the consistency ("goodness") of the logistic $\hat{b}$ estimate, the criterion of b greater than two times its standard error is used. Thus, estimates of logistic b which are not larger than two times its S.E. (eight cases) were eliminated from a latter set of analyses and are discussed below.

The descriptive statistics of the participant sample as a whole (i.e., aggregate of individual participants) for the variables measured directly by the questionnaire are presented in Table 5. Except for the amount of link strength, the values of the variables are based only on the first time participants completed the questionnaire. For participants who completed multiple questionnaires (i.e., received the message more than once), only three individuals reported different values for message temperature and message transmission resistance on subequent questionnaires. Lag time is the time between when an RA received the message (t_0) and the time when respective floor members received the message (t_i), i.e., $t_i - t_0$. In Table 5 lag time is presented in both minutes and hours. The mean time the message took in reaching network members was 38.04 hours or about one and one-half days. Lag time ranges from 0 (time the RAs received the message) to 485 hours or about 20 days. The mean link strength (i.e., the number of hours the receiver communicates with the sender in a typical week) is 15.44 hours and ranges from 0 to 160

TABLE 5. Descriptive Statistics of the Participant Sample

Variables	$\overline{X}$ (N)	S.D.	Skewness	Kurtosis	Low - High (Range)
Lag time in minutes	2300.70 (272)	4616.72	3.73	16.00	0 to 29155 (29155)
Lag time in hours	38.04 (272)	76.92	3.73	16.00	0 to 485 (485)
Link Strength (Hours Communcate/week)	15.44 (292)	27.25	2.86	8.73	0 to 160 (160)
Message Temperature	168.30 (267)	146.20	4.19	20.44	0 to 988[a] (988)
Message Transmission Resistance	43.49 (266)	98.54	6.90	62.59	0 to 988[a] (988)

Note. N, total number of participants, is based on only the first time participants completed the questionnaire, except for the amount of link strength.
[a]Values of 1,000 and greater were coded 988.

hours. The upper limit on the link strength scale is 168 hours (24 hours times 7 days). Participants who reported 40 hours or more also wrote on the questionnaire that the sender is a roommate. The distributions of lag time and link strength are positively shewed and more peaked than a normal distribution.

Since message temperature and message transmission resistance are new measurement scales, their reliability and validity need to be assessed. Because the study's design holds the two message variables relatively constant, test-retest measures was not explicitly made. However, 38 of the 292 participants completed the questionnaire two or more times during the diffusion process. While the time interval between t_1 and t_2 varies from several minutes to two days, a conservative measure of reliability can be computed for those participants. The reliability coefficient is .9944 ($p \leq .05$) for the amount of message temperature and .9948 ($p \leq .05$) for the amount of message transmission resistance. Based on the limited data of unequal time intervals, the two message measurement scales appear to be relatively stable.[4]

One way to assess validity is to determine if the variable behaves systematically relative to other variables. The correlation between message temperature and message transmission resistance is -.251 while the two message variables have extremely low to zero correlations with the other four exogenous variables. This is not surprising since the study's design constrained the two message variables to be constants. However, the message was selected on the basis of being highly interesting to dorm residents while minimizing any hesitancy to pass on the message. The amount of message temperature has a mean of 168.30 which indicates that participants consider the information about winter term RHA movies a little more than one and one-half times more interesting than average. This indicates that the message was appropriately perceived as having a relatively high amount of message temperature.

The mean for the amount of message transmission resistance (43.49) is about half as much as average. This indicates that resistance to pass on the message is relatively low. The mean of message transmission resistance may have been even lower had the question not immediately followed the message temperature (interest) question on the questionnaire. As it is, some participants may have misread the question or mentally reversed the measurement scale. Because some large values did occur, the mean and the range of message transmission resistance are higher than was expected. If the message had been on a sensitive topic, then large values would have been reasonable and expected. Based on the message criteria for selection and the means for the two message variables, there is some tentative basis for the validity of the two message measurement scales. For the participant sample as a whole, both message variables are positively skewed with very peaked distributions.

The sociograms of the single message's diffusion are presented in Figures C1 through C33 in Appendix C. In the figures, the dots with numbers represent nodes (network members) with node 01 being the RA, the directed lines represent the direction of the message's flow or directed communication link, and the numbers on the directed link represent the link's strength (i.e., hours per week communicating). The sociograms are vigraphs, i.e., value directed graphs. Of the 33 vigraphs, 29 are radial or tree-branching in appearance and anchored on the RA (node 01) while four of the vigraphs are relatively interlocking (Figures C6, C7, C11, and C17). The RAs in all four interlocking vigraphs are women with two of them residing in the same dorm. Examining Table 6 to be presented, there does not appear to be any further commonality among the four interlocking vigraphs. For five vigraphs (Figures C8, C22, C26, C27, and C28), the message did not move past the first-order zone while the message moved into the fourth-order zone of three vigraphs (Figures C5, C6, and C7).

The amount of network density, length of network reachability, amount of anchor centrality, and length of anchor reachability were computed from the vigraphs and are presented in Table B2 along with each network's score for the estimated logistic b parameter, amount of message temperature, and amount of message transmission resistance. The N in the table is the network size and is not necessarily the same N used for computing the network's logistic b, message temperature, and message transmission resistance. The T for the logistic b was the number of different time points (see Table B1) and there are several missing values for message temperature and message transmission resistance.

Descriptive statistics of the seven untransformed variables are presented in Table 6. The sample size (N) for the variables is 33

TABLE 6. Descriptive Statistics of the Untransformed Variables for a Sample Size of 33 Networks

Variables	$\overline{X}$	S.D.	Skewness	Kurtosis	Low - High (Range)
Logistic $\hat{b}$	.007	.015	3.758	15.810	.00195 to .077288 (.076)
Network Density	2.410	1.779	.864	-.145	.3667 to 6.9667 (6.600)
Network Reachability	.463	.215	.403	-.214	.1289 to 1.0214 (.892)
Anchor Centrality	4.147	3.581	1.863	3.804	.6667 to 16.5556 (15.889)
Anchor Reachability	.544	.278	.958	.855	.1614 to 1.2905 (1.129)
Message Temperature	163.572	56.667	2.061	6.447	88.20 to 384.50 (296.300)
Message Transmission Resistance	46.087	36.834	1.953	4.357	0.00 to 175.00 (175.00)

networks. The mean for each variable is the mean of each column in Table B2. Based on the skewness and kurtosis, it is apparent that the logistic b (dependent variable) needs to be rescaled (or transformed). To meet the assumptions of the general linear model for estimating and testing the exogenous variable parameters, the residuals ($e = Y-\hat{Y}$) should approach a normal distribution with homoscedastic variance (Hanushek & Jackson, 1977). The heterogeneity of the logistic b is discussed below. Of the six exogenous variables, only the amount of network strength appears to approach a normal distribution while the other five variables are mildly positively skewed. However, the skewness of the exogenous variables does not effect the estimates of the partial slopes (γ_i) nor the test of significance of the standard errors of the partial slopes (Hanushek & Jackson).

To further explore the possibility of heteroscedastic variance in the logistic b variable, bivariate scatterplots were made of the logistic b with each of the six exogenous variables (Figures D1 - D6 in Appendix D). As can be seen in all six scatterplots, there is one logistic b outlier in the upper center or left quadrant. Based on the positive skewness and the outlier, it is apparent that the logistic b parameter variable is nonnormally distributed with heteroscedastic variance. Subsequent analyses test whether or not the logarithmic (ln) transformation specified by the theory is sufficient to correct the nonnormality and heterosecdasticity of the logistic b. In addition, later analyses also explore the effect of removing the outlier.

The correlations among the seven variables are presented in Table 7. All of the bivariate relationships of the logistic b with each of the exogenous variables are in the expected directions. However, it should be noted that the absolute correlation values of message temperature and message transmission resistance with the logistic b are very low. This is not surprising given the diffusion of a single message and the criterion of topic interest commonality for the sample selection. Among the exogenous variables, there appears to be a high, although not perfect, degree of multicollinearity between network reachability and anchor reachability (r = .916). The high correlation between these variables is a result of using $\Sigma f(D)$ rather than $n(n-1)$ as the denominator in computing network reachability. For four of the vigraphs, $\Sigma f(D)$ is the same value as $(n-1)$ which is the denominator for computing anchor reachability, i.e., four vigraphs resulted in network reachability scores identical to their respective anchor reachability scores. Many of the other vigraphs have network reachability scores very close to their anchor reachability scores because the message diffused very little beyond the first-order zone.

Test of the Theory

The test of the theory includes: (a) a search for the optimum transformation, (b) chi-square test to determine if the logarithmic transformation specified by the theory is significantly different from the optimum transformation, (c) separate removal of each exogenous variable from the analysis along with the removal of several combinations of variables (d) a partial parallel set of analyses in which the logistic b outlier is removed, and (e) a partial parallel

TABLE 7. Correlations of the Untransformed Variables

Variables	1	2	3	4	5	6	7
Logistic $\hat{b}$	1.0				(symmetrical)		
Network Density	.209	1.0					
Network Reachability	-.286	-.331	1.0				
Anchor Centrality	.207	.197	-.452	1.0			
Anchor Reachability	-.203	-.214	.916	-.509	1.0		
Message Temperature	.015	.019	-.079	.125	-.084	1.0	
Message Transmission	-.115	-.109	-.109	-.052	-.122	-.251	1.0

Note. N=33 groups

set of analyses in which eight cases which had a ratio of the logistic b to its S.E. less than two are removed.

Table 8 presents a summary of the results from OLS regression using the Box-Cox family of transformations. Using all six exogenous variables (i.e., none removed) in the equation produces the optimum transformation of $\hat{\lambda} = -.10$ for the logistic b and $\hat{\mu} = 1.35$ for the set of exogenous variables with an optimum L_{max} $(\hat{\lambda},\hat{\mu})$ of 187.869. ($(\hat{\lambda},\hat{\mu})$ refers to the optimum transformation parameters resulting in the highest L_{max}.) The approximate 95% confidence interval (C.I.) was computed with a chi-square of 2 degrees of freedom (χ_2^2):

$$L_{max}(\hat{\lambda},\hat{\mu}) - L_{max}(\lambda,\mu) \leq \tfrac{1}{2}\chi_2^2(1 - \alpha), \quad (19)$$

with $\alpha \leq .05$,

$$187.869 - L_{max}(\lambda,\mu) \leq 2.996 \quad (20)$$

$$184.873 \leq L_{max}(\lambda,\mu) \quad (21)$$

which results in the L_{max} interval of 184.873 to 187.969.[5]

The L_{max} value for $\lambda = 1$ and $\mu = 1$ (i.e., regression with untransformed variables) is 137.955 which falls well outside the 95% C.I. This was expected based on the theory and the preliminary analyses which indicate nonnormality and heteroscedasticity in the logistic b dependent variable. Since the $L_{max}(1,1)$ is a substantial distance from the lower boundary of the 95% C.I., it is eliminated from further analyses.

TABLE 8. Results from Box-Cox Transformation Regressions with a Sample Size of 33 Networks

Variables Removed from the Regression	Optimum $\hat{\lambda},\hat{\mu}$	Optimum L_{max}	L_{max} for ln,1	L_{max} for ln,ln	$\bar{R}^2$ for ln,ln (F)	R^2 for ln,ln[b]	R for ln,ln
None[a]	-.10,1.35	187.869	187.316	186.984	.0949 (1.56)	.2646	.5144
Network Density	-.10,1.45	188.404	187.847	187.363	.1154 (1.84)	.2536	.5036
Network Reachability	-.10,.05	186.930	186.118	186.504	.0681 (1.47)	.2137	.4623
Anchor Centrality	-.10,.15	187.345	186.712	187.040	.0980 (1.70)	.2389	.4888
Anchor Reachability	-.10,1.70	187.595	186.986	187.094	.1010 (1.72)	.2415	.4914
Message Temperature	-.10,1.60	187.796	187.273	186.977	.0945 (1.67)	.2360	.4858
Message Transmission Resistance	-.10,1.35	188.491	187.911	187.111	.1018 (1.73)	.2422	.4921
Message Temperature Message Transmission Resistance	-.10,1.45	188.346	187.870	187.289	.1114 (2.00)	.2225	.4717
Anchor Reachability Message Temperature Message Transmission Resistance	-.10,1.75	188.078	187.515	187.036	.0977 (2.16)	.1823	.4270
Network Density Message Temperature Message Transmission Resistance	-.10,1.50	188.819	188.339	187.537	.1247 (2.52)	.2067	.4547

Note. λ and μ were changed at .05 intervals in the Box-Cox runs.
[a]For $\lambda=1$ and $\mu=1$, $L_{max}=137.955$ which was not in the 95% confidence interval.
[b]Results are not significant at $\underline{p} \leq .05$.

The theoretically specified log-log (ln,ln) transformation falls within the 95% C.I. (L_{max}(ln,ln) = 186.984) as did the log transformation of the dependent variable only (L_{max}(ln,1) = 187.316, i.e., regression of the logarithmic transformation of the logistic b with the untransformed set of six exogenous variables. While the (ln, ln) transformation is theoretically relevant, the (ln,1) transformation is empirically relevant based on the positively skewed distribution of the logistic b and the outlier identified in the bivariate scatterplots. Because (ln,1) transformation with all six exogenous variables in the equation resulted in a higher L_{max} value than the L_{max}(ln,ln) and both transformations are within the 95% C.I., L_{max}(ln,1) results are included in Table 8 for comparison. Since $\bar{R}^2$, R^2, and R are not valid criteria for the selection or test of an appropriate functional form (Anscombe, 1974; Bauer, 1981), they are presented for only the theoretical (ln,ln) transformation as additional information on the removal of exogenous variables from the equation.[6]

The results of the separate removal of each exogenous variable are presented in Table 8. Both the L_{max}(ln,1) and L_{max}(ln,ln) for each equation are within the respective 95% C.I., i.e., between $L_{max}(\hat{\lambda},\hat{\mu})$ and ($L_{max}(\hat{\lambda},\hat{\mu})$ - 2.996).

Rather than using the change in R^2 to assess the effects of adding or deleting an exogenous variable from an equation, the change in $L_{max}(\lambda,\mu)$ is used. As demonstrated by Bauer (1981), $L_{max}(\hat{\lambda},\hat{\mu})$ is a more stable and, therefore, a more appropriate criterion than R^2 to use in selecting transformation parameters for functional forms. Since a functional form consists of both the transformation parameters and the addition or deletion of variables from an equation, the change in $L_{max}(\lambda,\mu)$ values are used to evaluate the effects of removing variables from the equation.

Removing one exogenous variable at a time, the removal of network density results in the highest $L_{max}(\ln,\ln)$ (i.e., 187.363), while the removal of message transmission resistance results in the highest $L_{max}(\hat{\lambda},\hat{\mu})$ (188.491) and highest $L_{max}(\ln,1)$ (187.911). The single removal of the other exogenous variables had varying minimal positive and negative effects on $L_{max}(\hat{\lambda},\hat{\mu})$, $L_{max}(\ln,1)$ and $L_{max}(\ln,\ln)$.

Since a single message was diffused in the multiple networks, the effect of the joint removal of message temperature and message transmission resistance is examined. The L_{max} for $(\hat{\lambda},\hat{\mu})$, (ln,1), and (ln,ln) increased compared to the respective L_{max} for all six variables included (i.e., none removed). However, the differences between L_{max} $(\hat{\lambda},\hat{\mu})$ for the removal of message temperature and message transmission resistance and $L_{max}(\hat{\lambda},\hat{\mu})$ for the equation with all six exogenous variables included is less than 2.996. Therefore, the two message variables can be considered constants (i.e., set to 1) and excluded from the equation.

The joint removal of anchor reachability, message temperature, and message transmission resistance was conducted. Anchor reachability is jointly removed with the two message variables since anchor reachability and network reachability are highly correlated (r = .91) and the removal of anchor reachability resulted in a higher L_{max} than network reachability. The $L_{max}(\hat{\lambda},\hat{\mu})$, $L_{max}(\ln,1)$ and $L_{max}(\ln,\ln)$ for the joint removal of the three variables are higher than respective L_{max} values for the inclusion of all six exogenous variables. However, the joint removal results in a lower L_{max} than the single removal of network density.

Among the single removal of variables, the removal of network density results in the highest L_{max}. Thus, the joint removal of network density, message temperature, and message transmission resistance was conducted. The joint removal of these three variables results in the highest obtained $L_{max}(\hat{\lambda},\hat{\mu})$, $L_{max}(\ln,1)$ and $L_{max}(\ln,\ln)$.

The highest $L_{max}(\hat{\lambda},\hat{\mu})$ in Table 8 is 188.819 ($\hat{\lambda}$ = -.10, μ = 1.50) for the joint removal of network density, message temperature, and message transmission resistance. In constructing the 95% C.I., the two degrees of freedom are based on the two transformation parameters $(\hat{\lambda},\hat{\mu})$ to be estimated rather than on the number of partial slopes (γ_i). Thus, the $\frac{1}{2}\chi_2^2(1 - \alpha)$, i.e., 2.996, remains the same for each C.I. Therefore, each $L_{max}(\lambda,\mu)$ can be evaluated against the 95% C.I. around (-.10,1.50) which is referred to as the overall 95% C.I. For the $L_{max}(-.10,1.50)$, the approximate 95% C.I. is:

$$L_{max}(-10,1.50) - L_{max}(\lambda,\mu) \leq \tfrac{1}{2}\chi_2^2(1 - \alpha) \qquad (22)$$

$$188.819 - L_{max}(\lambda,\mu) \leq 2.996 \qquad (23)$$

$$185.823 \leq L_{max}(\lambda,\mu) \tag{24}$$

which results in the L_{max} interval of 185.823 to 188.819.

All of the L_{max} values in Table 8 fall within this overall 95% C.I. Hence, all three sets of transformations on the various combinations of exogenous variables are empirically appropriate. However, since the functional form specified by the theory is (ln,ln), it is the preferred transformation. For (ln,ln), the various combinations of exogenous variables all are within the overall 95% C.I. However, the equation with the removal of network density, message temperature, and message resistance has the highest L_{max}(ln,ln). This equation includes network reachability (ξ_2), anchor centrality (ξ_3), and anchor reachability (ξ_4). The theoretical form of the equation is written:

$$\eta = \gamma_{o_1} \frac{\xi_3^{\gamma_3}}{\xi_2^{\gamma_2}\xi_4^{\gamma_4}} \zeta, \tag{25}$$

where $\gamma_{o_1} = (\gamma_o \xi_1 \xi_5 \xi_6)$ and γ_o is the intercept in equation (5). Linearizing equation (25) by taking natural logarithms (ln) results in:

$$\ln\eta = \ln\gamma_{o_1} - \gamma_2 \ln\xi_2 + \gamma_3 \ln\xi_3 - \gamma_4 \ln\xi_4 + \ln\zeta. \tag{26}$$

The partial slopes and their S.E.s are presented for both equations (6), i.e., the equation with all six exogenous variables, and (26) in a later table.

Looking down the column labeled "optimum $\hat{\lambda},\hat{\mu}$" in Table 8, it can be seen that the removal of different exogenous variables effects $\hat{\mu}$ for the set of exogenous variables and not $\hat{\lambda}$ for the logistic b (endogenous variable). The $\hat{\lambda}$ for all combinations of exogenous variables is -.10 while $\hat{\mu}$ ranges from .05 to 1.75.

Table 9 presents the descriptive statistics for the logarithmic (ln) variables. Comparing the skewness and kurtosis statistics in Table 9 with Table 6, transforming the logistic b by ln removes most of the nonnormality in the untransformed variable. Except for anchor centrality and message temperature, the ln transformation over-corrects the positive skewness in the untransformed exogenous variables. However, as mentioned previously, the nonnormality of the exogenous variables is not relevant to the appropriate functional form for linear analysis. The descriptive statistics for the residuals (at the bottom of the table) are discussed in a later section on residual analysis.

Because a logistic b outlier was discovered in the untransformed bivariate scatterplots, the effect of removing the case with the outlier is examined.[7] The descriptive statistics of the untransformed variable with the removal of the outlier case (i.e., N = 32) are presented in Table 10. Comparing Table 10 with Table 6 (i.e., untransformed variables with N = 33), the removal of the outlier

TABLE 9. Descriptive Statistics of the Logarithmic (ln) Variables with a Sample Size of 33 Networks

Variables	$\overline{X}$	S.D.	Skewness	Kurtosis
Logistic $\hat{b}$	-5.928	1.365	.425	.298
Network Density	.575	.843	-.305	-.877
Network Reachability	-.892	.525	-.542	-.514
Anchor Centrality	1.118	.789	.189	-.560
Anchor Reachability	-.738	.527	-.216	-.448
Message Temperature	5.050	.301	.737	1.367
Message Transmission Resistance	3.201	2.094	-3.362	11.330
Residuals from Equation (6)	0.0	1.171	.681	1.105
Residuals from Equation (26)	0.0	1.216	.368	1.235

TABLE 10. Descriptive Statistics of the Untransformed Variables with a Sample Size of 32 Networks

Variables	$\overline{X}$	S.D.	Skewness	Kurtosis
Logistic $\hat{b}$	.005	.008	2.786	7.387
Network Density	2.345	1.768	.973	.127
Network Reachability	.472	.212	.386	-.141
Anchor Centrality	4.070	3.611	1.937	4.003
Anchor Reachability	.551	.279	.918	.809
Message Temperature	162.694	57.345	2.102	6.484
Message Transmission Resistance	47.528	36.467	2.024	4.491

decreases both the skewness (from 3.758 to 2.786) and the kurtosis (from 15.810 to 7.387) of the logistic b.

The transformation analysis is repeated with a sample size of 32 networks (i.e., the outlier case removed) for equation (6) (all six exogenous variables, i.e., none removed) and for equation (26) (network density, message temperature, and message transmission resistance removed). The results are presented in Table 11. Removing the outlier case raises $(\hat{\lambda},\hat{\mu})$ for both equations. The $L_{max}(\hat{\lambda},\hat{\mu})$ for both equations with N = 32 increased from the $L_{max}(\hat{\lambda},\hat{\mu})$ for the same equations with N = 33. However, $L_{max}(\ln,1)$ and $L_{max}(\ln,\ln)$ with N = 32 decreased. The L_{max} values in Table 11 are within the overall 95% C.I. and, therefore, there is not significant improvement in removing the outlier case.

The descriptive statistics of the logarithmic transformed variables with N = 32 are presented in Table 12. Compared to N = 33 (Table 9), the ln transformation with N = 32 results in a more normal distribution of the logistic b (i.e., from a skewness of .425 to .170 and from a kurtosis of .298 to .025).

As discussed in the previous section (see Table 5), eight of the ratios of the logistic b to its S.E. are less than two. The corresponding eight cases were removed and a parallel set of analyses to N = 32 was conducted for N = 25. For a sample size of 25, the descriptive statistics of the untransformed variables are presented in Table 13. Because the logistic b outlier case is also one of the eight removed cases, the skewness and kurtosis statistics for the

TABLE 11. Results from Box-Cox Transformation Regressions with a Sample Size of 32 Networks

Variables Removed from the Regression	Optimum $\hat{\lambda},\hat{\mu}$	Optimum L_{max}	L_{max} for ln,1	L_{max} for ln,ln	$\bar{R}^2$ for ln,ln (F)	R^2 for ln,ln	R for ln,ln
None	-.05,2.40	187.947	186.980	185.566	-.0322 (.84)	.1676	.4094
Network Density Message Temperature Message Transmission Resistance	0,3.0	188.921	187.770	186.307	.0145 (1.15)	.1099	.3315

Note. R^2s are not significant at $p \leq .05$.

TABLE 12. Descriptive Statistics of the Transformed (ln) Variables with a Sample Size of 32 Networks

Variables	$\overline{X}$	S.D.	Skewness	Kurtosis
Logistic $\hat{b}$	-6.033	1.244	.170	.025
Network Density	.546	.840	-.254	-.861
Network Reachability	-.865	.510	-.615	-.246
Anchor Centrality	1.094	.789	.261	-.474
Anchor Reachability	-.723	.528	-.287	-.369
Message Temperature	5.044	.303	.800	1.444
Message Transmission Resistance	3.444	1.581	-4.436	23.134

TABLE 13. Descriptive Statistics of the Untransformed Variables with a Sample Size of 25 Networks

Variables	$\overline{X}$	S.D.	Skewness	Kurtosis
Logistic $\hat{b}$	.005	.008	2.899	8.027
Network Density	2.193	1.793	1.317	1.056
Network Reachability	.446	.187	.135	-.894
Anchor Centrality	3.740	3.021	1.521	1.941
Anchor Reachability	.530	.258	1.006	1.724
Message Temperature	169.411	62.845	1.792	4.847
Message Transmission Resistance	49.133	40.817	1.787	3.029

logistic b with N = 25 are very similar to those in Table 10 (untransformed variables with N = 32).

The transformation results with N = 25 are presented in Table 14. All of the resulting L_{max} values in Table 14 are substantially less than the L_{max} values for N = 33 and N = 32 (i.e., Tables 8 and 11, respectively) and well below the lower boundary of the overall 95% C.I. However, the ln transformation for the logistic b with N = 25 most nearly approximates a normal distribution (see Table 15).

Based on a 95% C.I. for the $L_{max}(\hat{\lambda},\hat{\mu})$ values, the logarithmic (ln,ln) form of the theory cannot be rejected. While the elimination of the one logistic b outlier results in a slightly higher L_{max}(ln,ln) value, it is not statistically significant (i.e., less than a 2.996 increase in L_{max}). Therefore, all 33 networks are retained for the estimation of the partial slopes (γ_i) and regression coefficients ($\bar{R}^2$, R^2, and R). The choice between the equation with all six exogenous variables, equation (6), or the one with network reachability, anchor centrality, and anchor reachability, equation (26), will be made after the residual analysis and presentation of the partial slope estimates with their S.E.s. The following section presents the residual analysis for equations 6 and 26.

Residual Analysis

A residual analysis was conducted for the logarithmic (ln,ln) equation with all six exogenous variables, equation (6), and the logarithmic (ln,ln) equation with only three of the exogenous variables, equation (26), i.e., after network density, message

TABLE 14. Results from Box-Cox Transformation Regressions with a Sample Size of 25 Networks

Variables Removed from the Regression	Optimum λ,μ	Optimum L_{max}	L_{max} for ln,1	L_{max} for ln,ln	$\bar{R}^2$ for ln,ln (F)	R^2 for ln,ln[b]	R for ln,ln
None	-.05,-.40	143.190	142.284	142.981	-.0676 (.75)	.1993	.4464
Network Density Message Temperature Message Transmission Resistance	0,5.0[a]	143.916[a]	142.964	142.572	-.1032 (.25)	.0347	.1863

[a]After numerous runs, the search for the optimum transformation was discontinued at this point. λ appeared to stablize at 0 while increases of .05 in μ continued to result in very small increases in L_{max} $(\hat{\lambda},\hat{\mu})$.

[b]Results are not significant at $p \leq .05$.

TABLE 15. Descriptive Statistics of the Logarithmic (ln) Variables with a Sample Size of 25 Networks

Variables	$\overline{X}$	S.D.	Skewness	Kurtosis
Logistic $\hat{b}$	-5.991	1.270	.003	.170
Network Density	.472	.825	.015	-.842
Network Reachability	-.909	.487	-.750	.060
Anchor Centrality	1.036	.771	.131	-.671
Anchor Reachability	-.753	.511	-.379	-.091
Message Temperature	5.077	.331	.537	.802
Message Transmission Resistance	3.376	1.783	-3.952	18.221

temperature, and message transmission resistance are removed. The residual analysis for each equation consists of a scatterplot of the resulting residuals or errors (i.e., $\hat{e} = Y - \hat{Y}$) and the description of residuals (i.e., skewness and kurtosis).

The scatterplot of residuals (i.e., residuals vs. $\hat{Y}$) for equation (6) (all six exogenous variables) is presented in Figure E1 in Appendix E. The residuals appear to be mildly heteroscedastic, i.e., as Y increases, the variability appears to decrease. Since the sample size is quite small (N = 33), some slight patterning would be expected even if the residuals were random in the population.

Figure E2 presents the residual scatterplot for equation (26) (three exogenous variables of network reachability, anchor centrality, and anchor reachability). Virtually no pattern of the residuals appear to be present.

The descriptive statistics for the residuals from both equations are presented at the bottom of Table 9. The residuals from equation (6) have a slightly more positive skewness than the residuals from equation (26). However, the residuals from equation (26) have a slightly higher kurtosis value than the residuals from equation (6). Based on just the skewness statistic, both sets of residuals approximate a normal distribution. The peakedness (kurtosis) of both residual distributions may be the result of the small sample size since the distribution of the logarithmic logistic b (dependent variable at the top of Table 9) does closely approximate a normal distribution (i.e., skewness = .425 and kurtosis = .298). Based on the residual analyses, no clear empirical distinction exists between equations (6) and (26).

The unstandardized parameter estimates (γ_i) and their S.E.s for equations (6) and (26) are presented in Table 16. While the S.E.s of the partial slopes in both equations are relatively high (i.e., ratio of γ_i/S.E. < 2), this may be the effect of a very small sample size (N = 33) and relatively low variability in the exogenous variables. However, based on the $\hat{\gamma}_i$/S.E. ratio, the parameter estimates for network reachability and anchor reachability are better for equation (26) than equation (6) (i.e., the smaller the S.E. relative to $\hat{\gamma}_i$, the larger the $\hat{\gamma}_i$/S.E. ratio).

For equation (6), the signs of the coefficients for anchor reachability (ξ_4) and message temperature (ξ_5) are reversed from what is expected according to the theory. Recall the theoretical multiplicative equation:

$$\eta = \gamma_0 \frac{\xi_1^{\gamma_1} \xi_3^{\gamma_3} \xi_5^{\gamma_5}}{\xi_2^{\gamma_2} \xi_4^{\gamma_4} \xi_6^{\gamma_6}} \zeta. \qquad (5)$$

Based on the results of equation (6), the multiplicative equation is:

$$\eta = \frac{\xi_1^{\gamma_1} \xi_3^{\gamma_3} \xi_4^{\gamma_4}}{\gamma_0 \xi_2^{\gamma_2} \xi_5^{\gamma_5} \xi_6^{\gamma_6}} \zeta, \qquad (27)$$

where γ_0 and ξ_5 are in the denominator and ξ_4 is in the numerator.

TABLE 16. Regression Results of the Two Logarithmic Equations

Variables	Equation 6 $\hat{\gamma}$ (S.E.)	Equation 6 $\hat{\gamma}_i$/S.E.	Equation 26 $\hat{\gamma}$ (S.E.)	Equation 26 $\hat{\gamma}_i$/S.E.
Network Density (ξ_1)	.202 (.325)	.622	---------	
Network Reachability (ξ_2)	-1.774 (1.323)	1.341	-2.395 (1.203)	1.991
Anchor Centrality (ξ_3)	.366 (.384)	.953	.335 (.372)	.901
Anchor Reachability (ξ_4)	1.264 (1.396)	.905	1.735 (1.313)	1.321
Message Temperature (ξ_5)	-.802 (.797)	1.006	---------	
Message Transmission Resistance (ξ_6)	-.109 (.123)	.886	---------	
Constant (γ_0)	-2.702 (4.164)	.649	-7.157 (.509)	14.061
R^2	.26		.21	
(F)	(1.56)		(2.52)	
L_{max}(ln,ln)	186.984		187.837	

Note. Parameter estimates are unstandardized (N=33). The exogenous variable parameters are not significant at $p \leq .05$.

For equation (26) (network density (ξ_1) message temperature (ξ_5), and message transmission resistance (ξ_6) have been removed), the multiplicative equation is:

$$\eta = \frac{\xi_3^{\gamma_3} \xi_4^{\gamma_4}}{\gamma_{0_1} \xi_2^{\gamma_2}} \zeta. \qquad (28)$$

The sign reversals for anchor reachability (ξ_4) and message temperature (ξ_5) are discussed in the next chapter.

Summary

In summary, the results of the data analyses indicate that:

1. The logarithmic (ln) logistic b approximates a normal distribution.

2. The logarithmic (ln) transformation of both the logistic b and the set of exogenous variables is within the 95% C.I. for $L_{max}(\hat{\lambda},\hat{\mu})$ while the untransformed variables (1,1) were not in the 95% C.I.

3. The elimination of the one logistic b outlier results in little improvement and the elimination of the eight cases with a ratio of logistic b/S.E. < 2 results in insignificantly lower L_{max} values.

4. The residual analysis confirms the choice of the logarithmic transformation parameters, i.e., the logarithmic residuals have approximately homoscedastic variance and are approximately normally distributed.

5. Only small differences exist between equation (6) with all six exogenous variables and equation (26) with only network reachability, anchor centrality, and anchor reachability.

6. The signs of influence for anchor reachability and message temperature are reversed from what is predicted by the theory.

The results are discussed in the next chapter as they relate to the theory, design of the study, and literature.

FOOTNOTES--CHAPTER V

[1]To minimize the possibility of aliasing (Arundale, 1979), minutes rather than hours were used as the time unit. An alternative approach to estimating the logistic b is to use the same time span for all 33 networks. Using this approach, one would first determine what the maximum diffusion time was among the 33 networks, i.e., the longest time the message took to reach any of the participants. Then the maximum diffusion time is used as the time span for all 33 networks in estimating the b parameter. For example, if the longest diffusion time is 100 hours and the time interval is one hour, then the number of time points (T) for each network is 101 (100 plus t_o). This approach would have the effect of standardizing the sample size (time points) for estimating the logistic b parameter.

[2]The probability level (p) is based on the assumption of non-autocorrelation which generally cannot be met when using cumulative values. However, r and p are presented to give some indication of how well the data fit the logistic curve. Also, the presence of autocorrelation in single-equation models implies only inefficiency in the parameter estimates (Hanushek & Jackson, 1977). Thus, autocorrelation is a problem, but not a serious one for the estimation of the logistic b parameter.

[3]The correlation between $\hat{b}$/S.E. Ratio (i.e., t value of $\hat{b}$) and the network size (N) is -.06.

[4]The means and standard deviations of message temperature and message transmission resistance for the test-retest reliability are:

Variable	$\overline{X}_{t_1}$ (S.D.)	$\overline{X}_{t_2}$ (S.D.)
Message temperature	185.74 (154.98)	182.58 (155.90)
Message Transmission Resistance	23.68 (41.48)	24.55 (41.19)

Note: N is 38 and the time intervals ranged from 0 to 51.5 hours with a mean time interval of 6.57 hours (S.D. = 11.19 hours).

[5]Additional computer runs were not made to determine the joint transformation of (λ,μ) for corresponding L_{max} intervals due to computer costs.

[6]R^2 and R were computed from $\overline{R}^2$ in the B-C (Box-Cox) computer results. The adjusted R^2 ($\overline{R}^2$) is defined as:

$$\overline{R}^2 = R^2 - \left(\frac{k-1}{N-k}\right)(1 - R^2),$$

where N is the number of cases and k is the number of coefficients estimated (including the constant when unstandardized partial slopes are estimated) (SPSS-60000 Supplement, p. 4-16). The Nie et al. (1975, p. 358) SPSS manual incorrectly indicates k is equal to the number of independent variables only.

[7]The case containing the outlier rather than just the outlying value of the logistic b was removed because the B-C program cannot handle missing values.

CHAPTER VI

DISCUSSION

This chapter is divided into four major sections: (a) discussion of results, (b) critique and suggestions for future research, (c) outcomes of research and theory application, and (d) an overall summary. The discussion of the results in Chapter V are related to the theory presented in Chapter III. The second section is a critique of the research with suggestions for future research. Based on the results discussion and critique, the third section discusses the outcomes of the research and applications by practitioners. The fourth section is an overall summary of Chapters I through VI.

Discussion of Results

The theory of the rate of interactive information diffusion specifies that the log-log equation (6) should be the appropriate functional form. Based on the 95% C.I. constructed for each optimum L_{max} and the overall 95% C.I., the log-log form failed to be rejected. As a check of this result, a residual analysis was conducted. For the small sample size (N = 33), the logarithmic (ln) residual scatterplots closely approximate randomness, i.e., no apparent residual pattern. Based on the skewness and kurtosis statistics, both the logarithmic residuals and the logarithmic logistc b (dependent variable) closely approximate a normal distribution.

The untransformed equation (1,1) is rejected. However the logarithmic transformation of the dependent variable only (ln,1) also failed to be rejected. Although the (ln,1) form more often resulted in a slightly higher L_{max} value than (ln,ln), the (ln,ln) form is the preferred transformation based on the theory. Thus, the results indicate that the functional form of the theory is multiplicative with a log-normal error term.

Except for the two sign reversals (anchor reachability and message temperature), the results are all in the expected direction. The lack of significant R^2s and relatively high standard errors for the Box-Cox regressions may be mainly due to the small sample size of networks (N = 33). Even the small sizes of the networks (N ranged from four to 25) contributed to some insignificant R^2s in fitting the logistic curve to the diffusion data. While sample size may account for the lack of significant results, the magnitude of the R^2s from the Box-Cox regressions are relatively small.

Accounting for only 26% of the variance in the logistic b with all six exogenous variables is not considered very high. The relatively small R^2s may indicate inappropriate variables to describe the networks and the anchors. This issue is discussed in further detail in the next section.

In addition to testing the transformation parameters, the effects of eliminating variables singly and jointly was also examined. Equation (26), which included network reachability, anchor centrality, and anchor reachability, resulted in the highest L_{max} value for the (ln,ln) transformations. That is, the two message variables and network density can be considered as constants.

The logarithmic form of equation (6), which includes all six exogenous variables, is:

$$\ln\eta = -\ln\gamma_o + \gamma_1\ln\xi_1 - \gamma_2\ln\xi_2 + \gamma_3\ln\xi_3 + \gamma_4\ln\xi_4 - \gamma_5\ln\xi_5 - \gamma_6\ln\xi_6 + \ln\zeta. \qquad (29)$$

The multiplicative form for Equation (6) is:

$$\eta = \frac{\xi_1^{\gamma_1}\xi_3^{\gamma_3}\xi_4^{\gamma_4}}{\gamma_o\xi_2^{\gamma_2}\xi_5^{\gamma_5}\xi_6^{\gamma_6}}\,\zeta. \qquad (29)$$

Notice that the signs for anchor reachability (ξ_4) and message temperature (ξ_5) are reversed from the theory, i.e., equation (6) and (5), respectively. This is discussed below.

The logarithmic form of equation (26) is:

$$\ln\eta = -\ln\gamma_{o_1} - \gamma_2\ln\xi_2 + \gamma_3\ln\xi_3 + \gamma_4\ln\xi_4 + \ln\zeta. \qquad (30)$$

The multiplicative form of equation (26) is:

$$\eta = \frac{\xi_3^{\gamma_3}\xi_4^{\gamma_4}}{\gamma_{o_1}\xi_2^{\gamma_2}}\,\zeta \qquad (31)$$

where γ_{o_1} is a different intercept than γ_o in equation (29).

Although network density (ξ_1) and the two message variables (ξ_5 and ξ_6) are removed, anchor reachability (ξ_4) still has a positive rather than the expected negative influence on the logistic b (i.e., growth rate of the accumulated proportion of knowers). For equation (26), γ_{o_1} is a combination of intercept and the effects of the removed variables:

$$\ln\gamma_{o_1} = \ln\ (\gamma_o \xi_1^{\gamma_1} \xi_5^{-\gamma_5} \xi_6^{-\gamma_6}), \qquad (32)$$

where γ_o is the intercept in equation (6).

Comparing equations (6) and (26), equation B is slightly preferred based on: (a) higher L_{max} value, (b) two out of three higher $\hat{\gamma}_i$/S.E. ratios, (c) the small variability of network density, (d) considering message temperature and message transmission resistance as constants, and (e) slightly better residual analysis results. While equation (26) is not clearly superior to equation (6), it is preferred for this study. The preference rests largely on being able to consider the two message variables as constants which is consistent with the study's design constraints. The dormitory RA-groups were selected in the hope that they would have common topic interest for the diffusion of a single message, i.e., the two message variables could be considered constants.

The sign reversals for anchor reachability and message temperature warrant discussion. Four possible reasons for a parameter sign reversal are: (a) a near-zero bivariate correlation between the

endogenous and exogenous variables, (b) multicollinearity among exogenous variables, (c) analysis of a sub-group of the population, and (d) misspecification of the relationship, i.e., the reverse sign may be correct.

Message temperature was expected to be positively, rather than negatively, related to the logistic b. A possible indication of why the sign reversed can be found in the simple correlations of the untransformed variables (Table 7). The simple correlation sign for message temperature with the logistic b is positive, as expected. However, the correlation magnitude is so small ($r = .05$) as to approach zero, i.e., message temperature is constant. Thus, there is virtually no relationship between message temperature and the logistic b. Bivariate regression with the two variables would produce an almost horizontal regression line. When such a very small positive bivariate relationship exists, it would be very easy for the sign to reverse in the regression analysis with other exogenous variables. The near-zero bivariate relationship suggests that message temperature be treated as a constant in equation (26).

The sign reversal for anchor reachability poses a different problem. Again, the simple correlation sign is negative, as expected. However, based on the magnitude ($r = -.203$) one could not expect the relationship to reverse itself in the presence of other exogenous variables. However, the correlation between anchor reachability and network reachability is extremely high ($r = .91$). A high degree of multicollinearity between the two variables could account for anchor reachability's sign reversal in the regressions. However, when network reachability is removed from the regression equation, the sign of anchor reachability still remains positive. Thus, multi-collinearity does not appear to be the reason for anchor reachability's parameter sign to be positive rather than negative.

Another possible reason for the sign reversal is the relatively homogeneous RA-group sample which would classify as a sub-group of the population of all possible groups. Harary and Batell (1981) demonstrate that results obtained from an analysis of a sub-group can be in marked contrast to the results obtained from the data as a whole. They state that "failure to stratify with respect to important variables . . . can result in conclusions exactly contrary to the true nature of the system under investigation" (p. 36). Thus, if other types of groups (e.g., business organizations, social organizations, etc.) had been included in the sample, anchor reachability may have resulted in the expected negative sign.

On the other hand, it may well be that anchor reachability is positively related to the logistic b. For large networks composed of several to many cliques, the optimum anchor node may be the liaison person. Liaisons have weaker links than clique/group members and are important to the flow of information between cliques (e.g., Freidkin, 1980; Granovetter, 1973; Liu & Duff, 1972). Because of the weaker links, liaisons would have higher reachability scores. If most of the RAs (anchors) in the sample are liaisons rather than clique/group members, their reachability scores would be positively related to the logistic b. Thus, a distinction needs to be made between an anchor who is a liaison and an anchor who is a clique/group member.

Since there was no mass media transmission of the message prior to or during the study, the intercept is considered as only a scaling factor. Therefore, the sign of the intercept (γ_0) for equation (6) is not theoretically relevant to this study.

A large proportion of the networks based on a single event show a radial pattern. When a network radiates out from one person (i.e., anchor), Harary et al. (1965) refer to them as "a tree from a point graph," while Rogers and Kincaid (1981) refer to them as "radial personal networks." As suggested by Rogers and Kincaid, and Richards (1976), information diffusion networks based on a single event have less interconnectedness than networks based on a pre-existing class of events. Since previous interactive information diffusion research has not analyzed single message flow networks, no comparison is possible. The networks constructed by Coleman et al. (1957) and Rogers and Kincaid are based on influence networks for a pre-existing class of events for innovation adoption.

In summary, the results indicate that:

1. The exogenous variables may be treated as multiplicative with a log-normal error term.
2. The untransformed linear form (1,1) of the theory is rejected, i.e., results in heteroscedastic and non-normally distributed residuals.
3. As network density (ξ_1) (i.e., direct communication activity) increases, the logistic b (η) (i.e., the growth rate of accumulated knowers) increases. However, for

the RA-subgroup sample, network density is best considered as a constant.

4. As network reachability (ξ_2) (i.e., communication distance) increases, the logistic b (η) decreases.
5. As anchor centrality (ξ_3) (i.e., anchor's out-going direct communication activity) increases, the logistic b increases.
6. As anchor reachability (ξ_4) (i.e., anchor's out-going communication distance) increases, the logistic b increases. This is contrary to the theory.
7. As message temperature (ξ_5) (i.e., excitement, enthusiasm, and/or interest in the information) increases, the logistic b (η) decreases. However, message temperature is best considered as a constant because of design constraints.
8. As message transmission resistance (ξ_6) (i.e., reluctance to pass on the information) increases, the logistic b (η) decreases. Because of the study's design constraint (i.e., a single message diffused in multiple homogeneous networks), message transmission resistance is best considered as a constant.

Critique and Suggestions for Future Research

This section is a critique of the study with recommendations for future research. This is the first interactive information diffusion study to measure network _and_ message variables and to track the message's spread in multiple networks. Criticisms are classified into five categories: (a) groups/subjects, (b) message, (c) diffusion data gathering, (d) network data, and (e) variables. The weaknesses in the study suggest some of the next research studies. At the end of the section, a summary of the recommendations is provided.

Groups/Subjects

The size of individual networks creates a sampling problem. Fitting the logistic equation to each network's accumulated proportion of knowers resulted in 23 out of 33 significant r^2s ($p \leq .05$). All 10 insignificant r^2s are for networks with seven or fewer time points. One way to increase the number of time points is to increase the network size. Based on the 50-member RA-groups, it was hoped that the networks would range up to about 50 members for better statistical results in fitting the logistic equation. However, even with the relatively low network sizes (4 to 25), 23 sets of accumulated knowers resulted in significant r^2s of .452 or greater.

To hold the two message variables constant for the diffusion of a single message in multiple networks, groups were selected for topic commonality. While this was effective for considering the two message variables as constants, it is possible that the groups are too homogeneous. As discussed in the previous section, analysis of a homogeneous subgroup (and opportunity sample) may account for the sign reversal of the length of anchor reachability. Also, it may be one reason for the low variability in the amount of network density.

Therefore, future research should sample from a variety of group types, e.g., community groups and business organizations.

One of the major criticims of this study is the small sample size of 33 networks. The small sample size is a major factor contributing to the lack of significant results. If the sample size were 60 or more, the R^2s for equations (6) and (26) would be significant at $p \leq .05$.[1] Thus, for better statistical results, a sample size of at least 60 networks should be used in future research.

Message

For the diffusion of a single message, it was desirable to hold the two message variables constant. However, this does not allow for an adequate test of the message variables (i.e., amount of message temperature and amount of message transmission resistance). The ideal study would be the diffusion of different sets of messages in different sets of groups. Each message set should consist of four or more messages and each group set should be 64 (i.e., a minimum of two levels for six variables, 2^6) or more groups. If one or more of the six variables in the theory is held constant, then the number of groups in a set could be reduced. A group set could consist of relatively the same type of groups while differing across sets. While this type of study would provide the best information for testing the theory, it would be very complex (and expensive) to administer. However, this study can be viewed as a set of smaller studies, i.e., a program of research.

Another study weakness is the inadequate reliability and validity assessments of the two message measurement scales. Further research which allows the amount of message temperature and the amount of message transmission resistance to vary should provide reliability and validity assessments for the scales. In addition, multiple indicators for each exogenous message variable should be investigated.

Diffusion Data Gathering

To gather the message diffusion data, a short questionnaire and the printed message (single sheet) were circulated in each dorm area. Participants were instructed to return the questionnaire to an envelope on their floor bulletin board and to obtain additional message copies to pass on to others from an envelope on their RA's door. Participants were told they could keep the Winter Term RHA movie schedule. It was hoped that this would provide sufficient incentive or reward to complete and return the questionnaire. However, 10 RA-groups were eliminated because participation was too low (i.e., fewer than four participants) and 20 out of the remaining 33 groups had fewer than 10 participants. It is likely that the message was passed to many more people, but they failed to return the questionnaire. Low participation could be because dorm residents knew the message would be published in the school newspaper at the beginning of winter term. It is also possible that dorm floors have established tight cliques beyond which the message will not diffuse on its own.

Gathering the diffusion (and network) data as people receive the message eliminates most problems of recalling who gave them the message and when. However, the data are still dependent on people's willingness to cooperate. Alternatively, all RA-group members could have been surveyed at the completion of the diffusion process (two or three weeks after the RA received the message). While this procedure may have identified a larger number of knowers, the diffusion data would be less reliable. Thus, there is a trade-off in using either of the two diffusion data gathering procedures. A diffusion data gathering procedure needs to be developed which will maximize identifying knowers and maximize the reliability of the diffusion data.

Network Data

The network data for the study were derived from a single message diffusion event, not on a series of diffusion events. This was done to identify the directional path that the single message took and to minimize participants' time for better cooperation. However, this may also be one reason why the amount of network density has low variability and anchor reachability is positive rather than negative. Basing the network analysis on a single event may also have contributed to the low R^2 for both equations 6 and 26. It would be most interesting in future research to compare network data on both the single event and a pre-existing class of events on the message topic.

For use as a predictive tool, the network and anchor variables should be based on a class of pre-existing events. This would allow a practitioner to survey a network prior to a message's diffusion

to (a) select an optimal anchor or set of anchors and (b) predict the amount of time required for the diffusion. Thus, it is recommended that future research determine the influence of network and anchor variables based on a pre-existing class of events. In addition, it is suggested that the relationship between single event based variables and pre-existing class of events based variables be investigated.

Variables

This part includes a critique of the conceptual and operational definitions of the variables used in the study.

The use of vigraphs (value directed graphs) rather than ordinary or directed graphs for the two network (network density and network reachability) measures was beneficial. The use of vigraphs provided a more precise description of the networks than would have been possible with directed graphs. This allowed for greater differences between networks. Thus, it is recommended that future network based research continue the use of vigraphs.

From the results, network density was a poor predictor of the logistic b. While its low variability may account for this, it is also possible that network density is not a good measure of network structure. Rogers and Kincaid (1981) found that network density was not a good predictor of adoption and eliminated it after their initial analysis. In a very recent study, Friedkin (1981) demonstrated that the relationship between network size and network density based on ordinary graphs is nonlinear and heteroscedastic. Based on his Monte

Carlo simulations, Friedkin concludes that "density is not a generally useful indicator of network structure" (p. 50). This appears to be supported by this study.

The problem with network density appears to reside in the operational definition. For ordinary and directed graphs, density is defined as the ratio of the actual number of links to the total possible number of links in the network (e.g., Barnes, 1969; Friedkin, 1981; Granovetter, 1976; Richards, 1976; Rogers & Kincaid, 1981). When this operational definition is extended to vigraphs, it becomes apparent that this definition is an average network link strength measure. This is very different from the density measure used in the physical world.

Density in the physical world is the ratio of mass to the volume it occupies. Counting the number of unit links or summing the strength of links is equivalent to counting the number of particles of unit mass or summing the mass of each particle. However, the denominator value of n(n-1) is the size or "volume" of the adjacency matrix and not the "volume" in space that a network occupies. Using n(n-1) is actually computing the area of the adjacency matrix and forces all networks into two-dimensional space. Since using n(n-1) results in the mean link strength for the network, it is suggested that this measure be referred to as the mean network strength rather than as network density.

If the concept of density as used in the physical world is desirable for describing the structure of communication networks, then an equivalent measure for networks needs to be developed. One suggestion for a measure of "volume" is to take the trace of the mean corrected distance matrix pre-multiplied by its transpose.[2] The trace of this matrix is the total variance or the sum of squares of the distance matrix. In multidimensional scaling (MDS) and factor analysis, an eigenvalue may be normed to represent the sum of squared projections onto the corresponding eigenvector. Since the trace is also the sum of eigenvalues, the trace or the sum of the eigenvalues can be viewed as a gross measure of network "volume." Thus, an appropriate measure of network density for digraphs might be:

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} L_{ij(adj)}}{\sum_{i=1}^{n} e_i \text{ (dist.)}} \qquad (33)$$

and for vigraphs:

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} S_{ij(adj)}}{\sum_{i=1}^{n} e_{i(dist)}} \qquad (34)$$

where $\sum_{i=1}^{n} e_{i(dist.)}$ is the sum of the eigenvalues resulting from MDS on the distance matrix.

Volume in the physical world is the extension of an object measured on dimension one, times the extension measured on dimension two, times the extension measured on dimension three (i.e., length times width times height).

While MDS (or factor analysis) has been used as a way of analyzing networks (e.g., Barnett, 1979; Bonachich & Domhoff, 1981; Brophy, 1976; Farace & Mabee, 1980), eigenvalues do not appear to have been used as measures of network "volume" for computing network density.

The operational definition of network reachability was extended to vigraphs and was modified to be applicable for incomplete distance matrices. The modification replaced the n(n-1) denominator with $\Sigma f(D)$. A modification is necessary to deal with distances of infinity, i.e., nodes which are not reachable. However, after computing the network reachability and anchor reachability scores, it was discovered that four networks had identical scores for both variables while many others had very similar scores; in other words, there is a very high r^2 between the two variables. This occurred because of the large number of infinite cells in the distance matrix. Thus,the network reachability denominator $\Sigma f(D)$ is the same or similar to the anchor reachability denominator, (n-1). For "a tree from a point" networks, $\Sigma f(D)$ is, therefore, not a satisfactory modification. It is possible, however, that networks based on a pre-existing class of events rather than a single event would not have this problem. In any event, future research should evaluate alternatives to handling incomplete distance matrices.

An alternative to the reachability problems for digraphs and vigraphs is to replace the network reachability variable with some similar variable. One possible variable is the diameter of the network, i.e., the longest geodesic of the network (Bavalas, 1948; Harary et al., 1965). Thus, instead of using the average geodesic, the maximum geodesic is used to describe the distance of the network.

The network diameter has certain intuitive appeal for interactive diffusion. It would indicate the longest distance the message would have to travel from one end to the other in the network. However, for interactive information diffusion which starts at some node (or set of nodes), the major focus is on the distance from the anchor to the farthest point in the network. Thus, it may be reasonable to replace both network and anchor reachability with the maximum geodesic between the anchor and the farthest node.

If the anchor node is the lowest distance-based node in the network, this distance would be the network's radius (e.g., Bavalas, 1948). However, the anchor node may not always be the most central. I will call anchor length a general measure of the maximum geodesic from the anchor to the farthest node. If the anchor is the most central distance based node, then the anchor length will be at a minimum and will be the same as the radius of the network. Like reachability, anchor length can range up to infinity if the anchor is an isolate. It is expected that the greater the anchor length, the slower the accumulated growth rate of knowers in interactive information diffusion. Thus,it is suggested that future research investigate the utility of using anchor length as an alternative to both network reachability and anchor reachability.

Summary of Suggestions for Future Research

The above critique of the study with recommendations for future research has resulted in a lengthy discussion. Therefore, the following is a brief summary of the suggestions for future research:

1. The sample size of networks should be increased for better statistical results.
2. The range of network size should be increased.
3. The sample of networks should represent a variety of group types, i.e., community and business organizations.
4. Multiple messages should be diffused.
5. The reliability and validity of the two message scales (message temperature and message transmission resistance) should be further assessed.
6. Compare network and anchor variables from a single event with the same variables from a class of pre-existing events.
7. Determine the influence of network and anchor variables based on a pre-existing class of events rather than a single event.
8. Continue the use of vigraphs for network based variables.
9. Evaluate alternative operational definitions (measures) of network density.

10. Evaluate alternatives for handling incomplete distance matrices, i.e., infinite distance for network reachability.
11. Investigate the utility of replacing both network reachability and anchor reachability with anchor length, i.e., a new variable based on the maximum geodesic of the anchor to the farthest node.

No one study can possibly do all the work that needs to be done for the development and testing of a new theory. Hence, the above recommendations represent a program of future research for interactive information diffusion.

Outcomes for Research and Theory Application

Outcomes of Research

The major outcome of this research has been to provide a basis for future research in developing and testing a theory of the rate of interactive information diffusion. The clearest results of this study indicate that the multiplicative, rather than the linear, form of the variables is an appropriate functional form, i.e., equation (5). However, depending on which combination of the six exogenous variables is used, only 21 to 26% of the variance in the logistic b parameter is explained.

The results also indicate that network density has questionable utility in explaining the interactive diffusion rate. It was suggested in the previous section that alternative direct communication activity measures be examined.

During the study it became apparent that the possibilities of infinite reachability present a problem in both directed and value directed graphs (vigraphs). To deal with infinite reachability in the study, the n(n-1) denominator in the network reachability formula was replaced by the more general $\Sigma f(D)$. However, because most of the networks are radial, this modification resulted in a very high correlation between network reachability and anchor reachability. For networks based on a single event, $\Sigma f(D)$ was not found to be a very useful method. Thus, for infinite reachability between any two nodes in a network, other alternatives need to be investigated.

The research began by developing and testing a theory which would explain the variations in the rate of spread for different messages in multiple networks. Although there are many weaknesses in the study, these weaknesses provide direction for future research in modifying and testing the theory presented.

Theory Application

A single study testing a new theory cannot possibly answer all of the necessary questions. Additional research is needed before the theory can be considered useful for practitioners. With further research along the lines outlined above, the rate of interactive information diffusion theory has possibilities for developing into a useful tool for information diffusion practitioners. However, even in its fetal development, the core of the theory focuses a practitioner's attention on three fundamental factors of the process: (a) the amount of direct communication activity in a network, (b) the communication distance information must traverse to reach the network members, and (c) the information's characteristics which facilitate or inhibit its spread.

When indicators of the factors can explain at least 80% of the variance in the accumulated growth rate of knowers, the theory will be useful for practitioners. It could then predict the amount of time needed for a specific message to diffuse in a specific network. While it may be difficult to readily change the communication structure of a large group of people, the theory will help guide in selecting the optimum group member (anchor) to be the first receiver of the information.

The theory could also be used as a diagnostic tool for improving the information-exchange pattern for a group of people. Returning to the example presented at the beginning of Chapter I, a person in production scheduling may not be receiving necessary information on time, or not at all, to adequately perform his/her job. The proposed theory would suggest an investigation of the person's communication distance from the originators of the needed information, i.e., the sales people. Strategies specific to the situation could be developed to decrease the person's (or department's) communication distance. Alternatively, if the originator of the information (e.g., sales) is responsible for disseminating information to other parts of the system (e.g., accounting, production scheduling, quality control, and data processing), the theory would suggest increasing the amount of direct communication activity of the originator. In addition, the person responsible for production scheduling may also need increased direct communication activity in his/her own department.

With further development, the theory will also help guide the construction of messages, and provide a way to evaluate alternative messages. For example, the theory focuses the practitioner's attention on constructing messages which will maximize receiver's enthusiasm about the information while minimizing receivers' reluctance to pass it on to others because of the message's topic.

Overall Summary

The term "interactive information diffusion" describes the process of person-to-person spread of information in a social system. For successful diffusion, the overall behavior of the accumulated number of knowers is described by the logistic equation. However, the logistic equation does not explain or predict the variations in the rate of spread for different messages in different networks.

A review of the literature shows that interactive information diffusion is a multi-disciplinary research interest. Both network and message variables have been repeatedly suggested as important to the diffusion process. However, no research has been found which directly relates variables claimed to be important to the rate of interactive information diffusion. Therefore, the purposes of this research were to: (a) review the literature for possible variables, (b) develop a theory (model) which relates the selected variables to the overall behavior of the process (i.e., the logistic equation), and (c) conduct a study to test the theory (model).

The strategy used to develop the theory was to modify the logistic equation. The logistic b parameter is a summary growth rate of the accumulated number of knowers. Since the purpose of the theory is to explain and predict the rate of the overall interactive diffusion process, the logistic b parameter became the dependent variable. Based on a broad area of literature, six variables were selected as exogenous variables to explain and predict variations of the logistic b. The six exogenous variables are: (a) amount of network density, (b) length of network reachability, (c) amount of anchor centrality, (d) length of anchor reachability, (e) amount of message temperature, and (f) amount of message transmission resistance.

Based on the six variables' boundary conditions, the theory specifies a multiplicative functional form of the variables. To test the theory, a non-experimental field study was conducted. A single message was diffused in multiple networks (N = 33). Dormitory floors consisting of a Resident Assistant (RA) and 50 residents each were selected as the target groups. A relatively homogeneous set of groups was selected to maintain topic commonality for the single message to be diffused, i.e., to treat the two message variables as constants. Based on the selected groups and criteria for the messages, the message was an advance Winter Term schedule for RHA movies.

The results of the data analyses indicate that the multiplicative functional form is appropriate. In addition, the results indicate that the two message variables (message temperature and message transmission resistance) can be considered constants in the study.

Contrary to the proposed theory, network density was not found to be an important predictor variable. In addition, anchor reachability was found to have a positive, rather than negative, relationship with the logistic b in the multi-exogenous equation. While possible explanations are offered, future research is needed.

Few studies are without their pit-falls and this one is no exception. However, this need not be a disadvantage if the weaknesses are examined and used to provide direction for future research. Therefore, a critique of the study was presented along with suggestions for future research. The two major suggestions for future studies are to increase the sample size and investigate alternative variables for describing the structure of the network and the location of the anchor.

The major outcome of this research has been to provide a basis and direction for developing and testing a theory of the rate of interactive information diffusion process.

FOOTNOTES--CHAPTER VI

[1]The sample size needed for a significant F at $p \leq .05$ for reported R^2 was computed from:

$$F = \frac{R^2/k}{(1-R^2)/(N-k-1)},$$

where k is the number of parameters to be estimated and N is the sample size (Nie et al., 1975, p. 335).

[2]Suggested by Edward L. Fink in a personal communication, January 1982.

APPENDICES

APPENDIX A

DIFFUSION STUDY HANDOUTS

INFORMATION DIFFUSION STUDY

The proposed project is a field study of the diffusion of the same message through multiple communication networks to examine the effects of two network variables on the speed of information diffusion. The study will consist of person-to-person dissemination of an actual message (½ page). The message topic will be based on the expected interests of the participating groups. A sample of 35-40 groups from Michigan State University residence halls will be used. Each group will consist of one Resident Assistant (RA) and 50 residents.

The questionnaire, instructions, and cover letter (a single sheet) will accompany the message as copies of the message pass from person-to-person. The RA will be the first person to receive the message with instructions to give copies of the message-questionnaire to other floor members s/he would normally tell the information to. Participants will be instructed to obtain the needed number of message-questionnaires from a large envelope on their RA's door labeled with the RA's name only. While a participant may be the receiver of the message-questionnaire many times, they will be instructed to distribute the message only once. Upon completion of the questionnaire, participants will be instructed to fold and staple it and return it to the researcher via campus mail.

NOTE: See the attached page for advantages of the study to dorms.

For any questions, contact:

Connie Bauer
441 Communication Arts Bldg.
Dept. of Communication
355-6666

INFORMATION DIFFUSION STUDY

Advantages of the Study to Dorms:

1. Participation can help development of the "student community" by helping to build information-exchange linkages on the floor.

   a. Participation would encourage and provide an opportunity for students on the floor to interact to develop acquaintanceships beyond the superficial level of "Hi, how are you?"

   b. Passing the message can give the student experience as an information source.

   c. Being known as an information source (i.e., sharing information with others) usually accrues positive attributions.

   d. Participation can help to increase awareness of the importance of information-exchange linkages with other floor members.

   e. Frequent information-exchanges facilitate integration of individuals in a group and facilitate group cohesiveness (i.e., a "sense of community").

2. Summary of research findings:

   a. Can provide an overview picture of the information-exchange network on the floor, for example:
      1. How interconnected floor members are
      2. The average number of steps/links from any one person to any other person on the floor
      3. How long it takes to diffuse information on the topic of study via person-to-person on the floor

   b. Can provide a basis and starting point for an RA to further develop or modify the information-exchange network on her/his floor (e.g., indentify and integrate "isolates," and create or increase bridges between subgroups on the floor, etc.).

   c. Knowing who is the most central person in the information-exchange network on the floor can facilitate future diffusion of information on the floor.

   d. Can increase the awareness and knowledge of the importance of communication networking forboth RAs and floor residents.

For any questions, contact:
Connie Bauer
441 Com. Arts Bldg.
Dept. of Communication
355-6666

INSTRUCTIONS

A. In person, pass a copy of the Winter RHA Films on to only those people on your floor (same RA as you) that you would normally tell this information to.

B. Extra copies of the Winter RHA Films are available in a large envelope on your RA's door, labeled with your RA's name only. Take only the number of copies you need to give to others.

C. Within the next 24 HOURS (if possible), in person hand copies of the Winter RHA Film schedule to the other people you would normally "tell" this information to. (Remember: only people who have the same RA as you.)

D. If someone tells you they have already received the winter film schedule, say to them: "Take it again and answer the questions. Then, place it in the large envelope on the BULLETIN BOARD marked RETURN."

E. After answering the questions (on the other side of this sheet), CUT this sheet in half along the dotted line. Then, place THIS HALF with the questions in the large envelope on the BULLETIN BOARD for your floor. The envelope is labeled with "RETURN" and your RA's name. (The remaining half with the Winter RHA Films is yours to keep.)

F. IF you have received this sheet before, answer the questions and place this sheet in the large envelope on the BULLETIN BOARD for your floor.

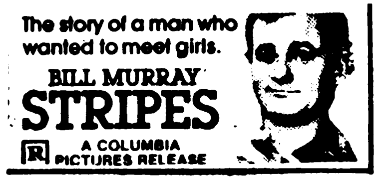
*NOTE: You may be the receiver of this film schedule and questions many times, but remember you only distribute the film schedule once to other people.

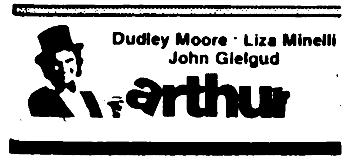
Your responses to the enclosed questions will be kept in confidence. Only myself and my research team will see the completed questions and your name will be immediately converted to an ID number. Your participation is very valuable to us and we appreciate your time and cooperation in this project. However, your participation is strictly voluntary. You may withdraw from this project at any time without penalty.

Thank you for your assistance. If you have any questions, contact:

Connie Bauer
441 Communication Arts Bldg.
Department of Communication
355-6666

CUT





! WINTER RHA FILMS !

You have just received advance news of the RHA film schedule for Winter term. The list of RHA films on the back of this sheet is yours to keep. In return for this advance news, please take a few minutes to answer the questions on the back of this sheet. Then, pass a copy of this sheet on to other floor members. Your cooperation is really appreciated.

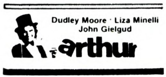
Your RA and dorm director have agreed to let me ask you to participate in this project. The results will help us understand how important information circulates among dorm residents.

If you have any questions, please contact:

Connie Bauer
441 Communication Arts Bldg.
Department of Communication
355-6666 (office)
484-3962 (home)

Please turn this sheet over, read the film schedule and answer the few questions. IF this is NOT the first time you have received this sheet, follow INSTRUCTION F.

WINTER RHA FILMS



Jan 7-10	Caddyshack, Superman II
Jan 12-13	Bonnie and Clyde
Jan 14-17	Kentucky Fried Movie, Harold and Maude
Jan 18-19	The Godfather
Jan 21-24	Stripes, Breaker Morant, Atlantic City
Jan 26-27	High Noon
Jan 28-31	Alien, Thief
Feb 1-2	Heaven's Gate
Feb 2-3	Little Big Man
Feb 4-7	The Rose, Bustin Loose
Feb 9-10	Psycho
Feb 11-14	The Four Seasons, Outland, Sleeper
Feb 15-16	North by Northwest
Feb 18-21	Coal Miner's Daughter, For Your Eyes Only
Feb 25-28	Airplane, Brubaker, American Pop
Mar 4-7	Arthur, Annie Hall
Mar 11-19	Fame Concert for Bangladesh, King of Hearts

CUT

QUESTIONS

Please answer the following questions:

1. Your dorm's name:________________ Your floor #:________

   Your RA's name:________________ Your name:________________

2. How many terms have you lived in this dorm? _______terms

3. Who gave this sheet to you? Name:________________________
   first last

4. HOW MANY HOURS do you communicate with the person named in Question #3 IN A TYPICAL WEEK (roughly estimate to the nearest hour)? ________hours

5. When did you read the list of films? Date: ____ ___ Time:________ AM PM
   month day (circle one)

6. Is this the FIRST time you have received this sheet? ___YES ___NO

For Questions #7 and #8: Evaluate the RHA film news by using the following scale or "yardstick." You may use any positive number (including 0 and numbers larger than 100) that best represents your feelings.

If 0 (zero) is NOT AT ALL, and 100 is AVERAGE, then:

EXAMPLE: if your interest in the news of RHA films is twice as much as average, you would write "200" in the space provided.

7. How EXCITED, ENTHUSIASTIC, or INTERESTED are you in the news of the RHA movies for Winter term? ________(you can use numbers larger than 100).

8. How HESITANT (for any reason) are you to pass the information on to someone you know might be interested in the news? ________(you can use numbers larger than 100).

Please TURN OVER for INSTRUCTIONS on passing the schedule of winter films on to other floor members.

MICHIGAN STATE UNIVERSITY

COLLEGE OF COMMUNICATION ARTS AND SCIENCES
DEPARTMENT OF COMMUNICATION

EAST LANSING · MICHIGAN · 48824

November, 1981

Dear RA:

Thank you for expressing your willingness to participate in a study of communication behavior (information diffusion) on your floor. You are among the first people on campus to receive information on the RHA films scheduled for Winter term.

Please note: This study is NOT a test or evaluation of your communication behavior. I am only interested in the overall picture of the communication on your floor.

Please follow the instructions on the white sheet (the message-questionnaire) that accompanies this letter. Give a copy to only those floor residents that you would normally tell informal information to. DO NOT give a copy of the message to all of your floor residents.

If you have any questions, please ask the person delivering this letter or call me.

Cordially,

Connie L. Bauer

Connie L. Bauer
441 Communication Arts Bldg.
Dept. of Communication
355-6666 (office)
484-3962 (home)

APPENDIX B

TABLES OF VARIABLE VALUES

FOR EACH NETWORK

TABLE B1. Logistic Parameters for each Network

Network ID	T	$\hat{a}$ (S.E.)	$\hat{b}$ (S.E.)	r^2 (r)	F significance
01	12	-.355 (.354)	.000331 (.000040	.873 (.934)	.000
02	6	-1.260 (2.636)	.000566 (.000438)	.294 (.542)	.266
03	7	-1.622 (1.484)	.003466 (.001312)	.582 (.763)	.046
04	12	-.808 (.468)	.002996 (.000451)	.815 (.903)	.000
05	7	-3.643 (3.315)	.020881 (.013216)	.333 (.577)	.175
06	9	-2.924 (.535)	.002853 (.000306)	.926 (.962)	.000
07	12	-5.423 (1.492)	.003542 (.000805)	.659 (.812)	.001
08	5	-2.405 (.577)	.001536 (.000157)	.970 (.985)	.002
09	4	-.370 (1.002)	.001639 (.000339)	.921 (.960)	.040
10	5	-1.567 (.396)	.001209 (.000094)	.982 (.991)	.001
11	15	.049 (.418)	.000377 (.000066)	.717 (.847)	.000
12	8	-1.548 (1.353)	.010292 (.003857)	.543 (.737)	.037

TABLE B1 (cont'd)

Network ID	T	$\hat{a}$ (S.E.)	$\hat{b}$ (S.E.)	r^2 (r)	F significance
13	8	-1.930 (.834)	.004800 (.000929)	.816 (.904)	.002
14	8	-.011 (.476)	.003791 (.000537)	.893 (.945)	.000
15	21	-1.652 (.675)	.000195 (.000049)	.452 (.673)	.001
16	3	-.495 (1.122)	.005544 (.001109)	.962 (.981)	.126
17	4	-2.157 (5.237)	.077288 (.090065)	.269 (.519)	.481
18	5	-2.420 (3.037)	.000949 (.000577)	.475 (.689)	.198
19	11	-.523 (.354)	.000438 (.000049)	.900 (.949)	.000
20	8	-1.515 (1.307)	.001349 (.000491)	.557 (.746)	.033
21	6	-.221 (.479)	.001303 (.000160)	.943 (.971)	.001
22	5	-1.408 (.474)	.029596 (.002830)	.973 (.987)	.002
23	5	-1.331 (2.926)	.002646 (.002115)	.343 (.586)	.300
24	7	-2.868 (.891)	.003056 (.000518)	.874 (.935)	.002

TABLE B1 (cont'd)

Network ID	T	$\hat{a}$ (S.E.)	$\hat{b}$ (S.E.)	r^2 (r)	F significance
25	8	-1.583 (.222)	.007658 (.000398)	.984 (.992)	.000
26	5	-.428 (.453)	.005338 (.000553)	.969 (.984)	.002
27	7	-3.525 (3.089)	.003057 (.001795)	.367 (.606)	.149
28	10	-3.903 (2.659)	.000711 (.000370)	.316 (.562)	.091
29	3	-1.543 (.676)	.034598 (.002367)	.989 (.994)	.067
30	7	-.185 (.478)	.000934 (.000124)	.919 (.958)	.001
31	4	-1.384 (4.877)	.002499 (.002861)	.276 (.526)	.474
32	7	-.790 (1.226)	.002544 (.000973)	.578 (.760)	.047
33	5	-.430 (.460)	.004732 (.000498)	.968 (.984)	.002

Note. T is the number of different time points in estimating the logistic parameters.

TABLE B2. Variable Values for Each Network

Network ID	N	Logistic $\hat{b}$	Network Density (S.D.)	Network Reachability (S.D.)	Anchor Centrality (S.D.)	Anchor Reachability (S.D.)	Message Temperature (S.D.)	Message Transmission (S.D.)
01	13	.000331	1.1603 (5.8060)	.2617 (.3290)	5.0833 (5.4349)	.3385 (.3490)	199.67 (252.37)	61.25 (66.20)
02	6	.000566	2.6667 (12.7829	1.0214 (.7273)	1.6000 (3.0496)	1.2314 (.7853)	128.80 (62.18)	42.00 (88.43)
03	7	.003466	1.9048 (8.0541)	.4920 (.4405)	4.5000 (6.0249)	.5971 (.4547)	179.80 (83.31)	20.00 (44.72)
04	15	.002996	1.2381 (6.3155)	.3229 (.3976)	1.5000 (3.1317)	.5262 (.4512)	111.73 (50.61)	33.40 (58.72)
05	10	.020881	4.3444 (17.4409)	.1866 (.4549)	16.5556 (31.4528)	.2080 (.1899)	156.90 (94.59)	25.10 42.43)
06	10	.002853	6.9667 (20.0955)	.2908 (.3947)	1.5556 (4.6667)	.3295 (.4311)	195.50 (116.11)	40.00 (51.64)
07	13	.003542	2.4615 (11.1282)	.7328 (.6180)	1.1667 (2.8551)	1.2905 (.7132)	176.92 (63.30)	30.77 (43.49)
08	5	.001536	1.4000 (3.9921)	.5387 (.5327)	7.0000 (6.9761)	.5387 (.5327)	384.50 (403.02)	12.50 (25.00)
09	8	.001639	.6429 (2.4601)	.3934 (.2142)	1.7143 (1.7043)	.4536 (.1944)	276.86 (316.78)	28.57 (39.34)
10	5	.001209	5.8000 (25.0023)	.7036 (.4452)	1.0000 (.8165)	.8772 (.2515)	183.33 (76.38)	00.00 (00.00)
11	20	.000377	2.0000 (10.3786)	.4617 (2.0421)	2.8421 (3.2191)	.3247 (.1713)	176.90 (201.52)	34.00 (45.93)
12	8	.010292	3.2500 (15.1493)	.3587 (.3093)	12.1429 (26.0284)	.3716 (.3638)	209.36 (163.63)	32.14 (47.25)
13	11	.004800	2.0273 (13.4890)	.4553 (.3882)	1.4000 (1.5055)	.5853 (.3794)	156.67 (99.34)	38.89 (54.65)
14	11	.003791	.6091 (3.0473)	.6096 (.4766)	3.3000 (5.4579)	.6209 (.4681)	164.67 (97.17)	33.33 (35.36)

TABLE B2 (cont'd)

Network ID	N	Logistic $\hat{b}$	Network Density (S.D.)	Network Reachability (S.D.)	Anchor Centrality (S.D.)	Anchor Reachability (S.D.)	Message Temperature (S.D.)	Message Transmission (S.D.)
15	23	.000195	1.3317 (11.1155)	.3933 (1.5767)	1.0000 (2.4137)	.5049 (.4803)	234.30 (246.10)	25.87 (41.63)
16	6	.005544	4.9333 (15.9869)	.1289 (.1544)	10.8000 (10.0349)	.1614 (.1916)	126.67 (25.82)	16.67 (25.82)
17	6	.077288	4.4667 (8.7247)	.1726 (.2408)	6.6000 (10.6911)	.2945 (.3996)	191.67 (49.16)	00.00 (00.00)
18	5	.000949	4.7500 (20.0785)	.6711 (.4695)	1.2500 (1.2583)	.8361 (.3352)	116.67 (28.67)	50.00 (50.00)
19	14	.000438	.4231 (2.0578)	.7023 (.5025)	3.2308 (5.0192)	.7692 (.5209)	160.50 (81.96)	32.25 (70.91)
20	10	.001349	1.7000 (7.0654)	.4917 (.5019)	8.1111 (10.6706)	.5954 (.4983)	176.00 (136.97)	125.80 (307.42)
21	6	.001303	.6667 (2.1867)	.7238 (.4684)	2.0000 (2.8284)	.8486 (.3969)	88.20 (54.09)	50.00 (50.00)
22	5	.029596	.6000 (1.4290)	.4875 (.3660)	3.0000 (1.8257)	.4875 (.3660)	103.75 (41.51)	50.00 (40.83)
23	6	.002646	1.6667 (6.6661)	.6901 (.4916)	3.0000 (5.0498)	.8224 (.4133)	137.50 (54.20)	50.83 (76.84)
24	7	.003056	3.1667 (9.6219)	.1508 (.1213)	5.6667 (8.0416)	.2090 (.1062)	115.00 (90.78)	52.86 (62.37)
25	10	.007658	1.0222 (6.4459)	.4111 (.3313)	3.0000 (3.5000)	.4537 (.3573)	148.33 (71.76)	22.22 (44.10)
26	5	.005338	.9000 (1.9974)	.2524 (.0963)	4.5000 (1.9149)	.2524 (.0963)	125.00 (28.87)	175.00 (95.74)
27	7	.003057	4.5238 (18.2560)	.2571 (.1880)	3.3333 (1.2111)	.3389 (.1319)	168.57 (79.04)	57.14 (83.81)
28	10	.000711	.3667 (1.5024)	.5750 (.4187)	3.6667 (3.3912)	.5750 (.4187)	125.00 (33.33)	42.50 (55.34)
29	4	.034598	3.0833 (10.0676)	.5393 (.5335)	.6667 (1.5000)	.7095 (.5031)	175.00 (35.36)	100.00 (00.00)

TABLE B2 (cont'd)

Network ID	N	Logistic $\hat{b}$	Network Density (S.D.)	Network Reachability (S.D.)	Anchor Centrality (S.D.)	Anchor Reachability (S.D.)	Message Temperature (S.D.)	Message Transmission (S.D.)
30	7	.000934	.3750 (1.3289)	.7620 (.6104)	2.2857 (2.6904)	.8083 (.6574)	112.86 (54.69)	64.29 (47.56)
31	4	.002499	1.9167 (4.2738)	.5458 (.5245)	7.3333 (6.4291)	.3944 (.5245)	137.50 (47.87)	25.00 (50.00)
32	8	.002544	5.0000 (20.3720)	.1950 (.3098)	4.2857 (4.5356)	.1838 (.0857)	143.75 (41.73)	123.50 (349.31)
33	5	.004732	2.1500 (5.6221)	.2875 (.3571)	1.7500 (2.8723)	.4032 (.3988)	110.00 (31.62)	25.00 (50.00)

Note. Some networks have missing values for message temperature and message transmission resistance. N is the size of the network.

APPENDIX C

SOCIOGRAMS

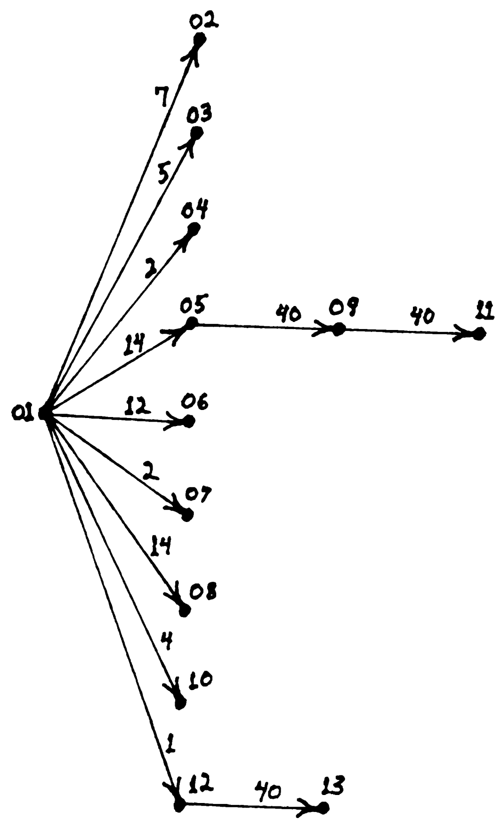


Figure C1. Sociogram of network 01.

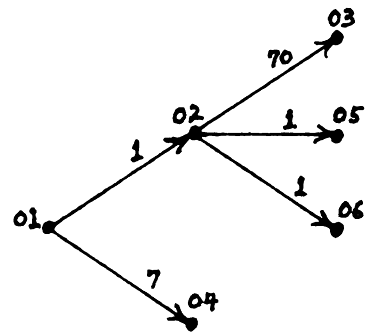


Figure C2. Sociogram of network 02.

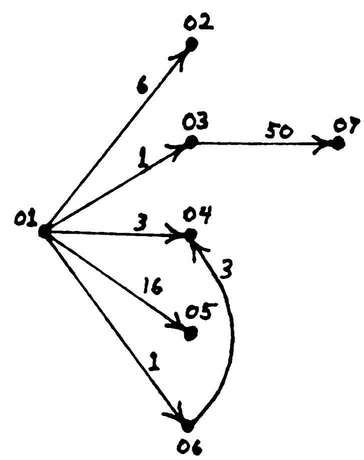


Figure C3. Sociogram of network 03.

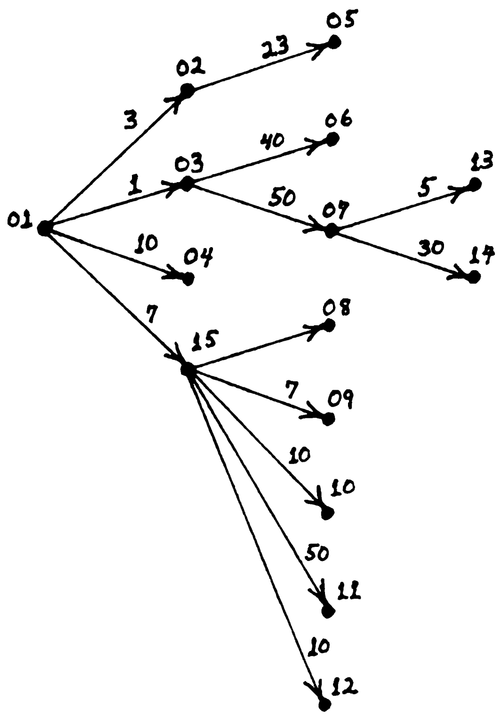


Figure C4. Sociogram of network 04.

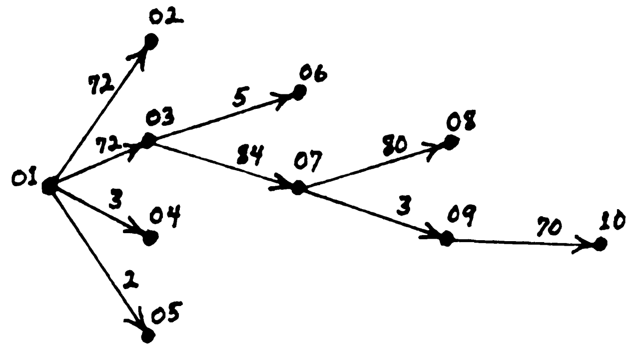


Figure C5. Sociogram of network 05.

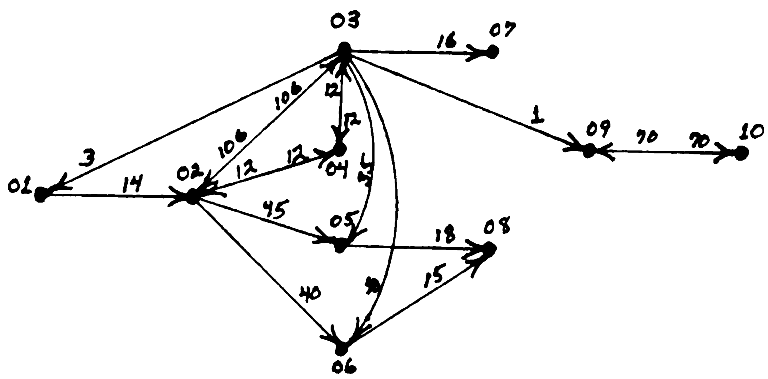


Figure C6. Sociogram of network 06.

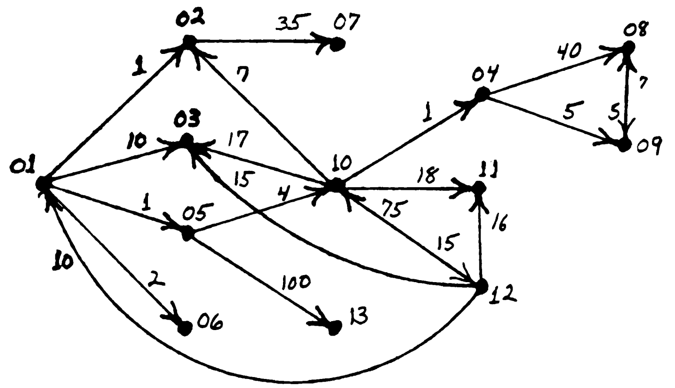


Figure C7. Sociogram of network 07.

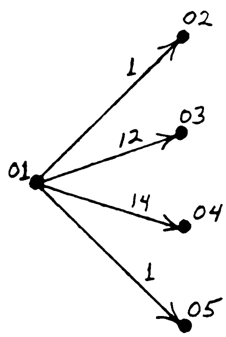


Figure C8. Sociogram of network 08.

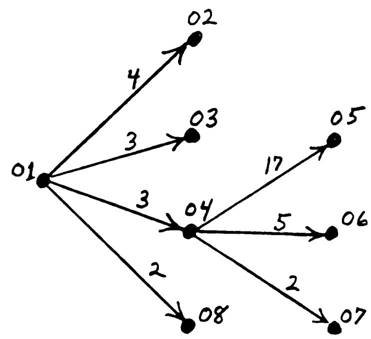


Figure C9. Sociogram of network 09.

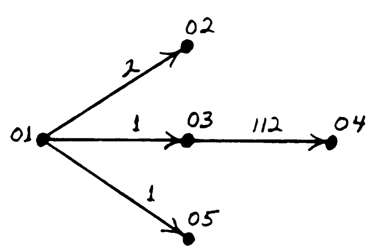


Figure C10. Sociogram of network 10.

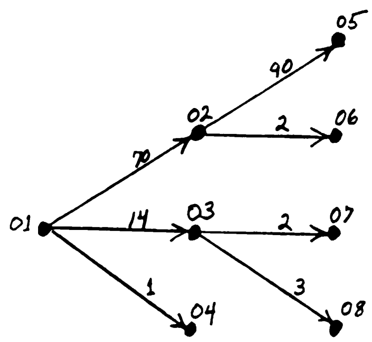
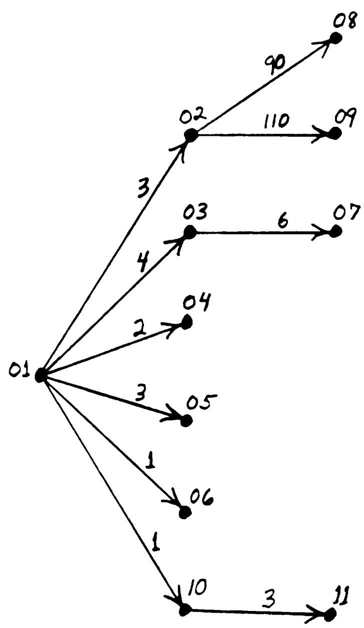


Figure C12. Sociogram of network 12.



C13. Sociogram of network 13.

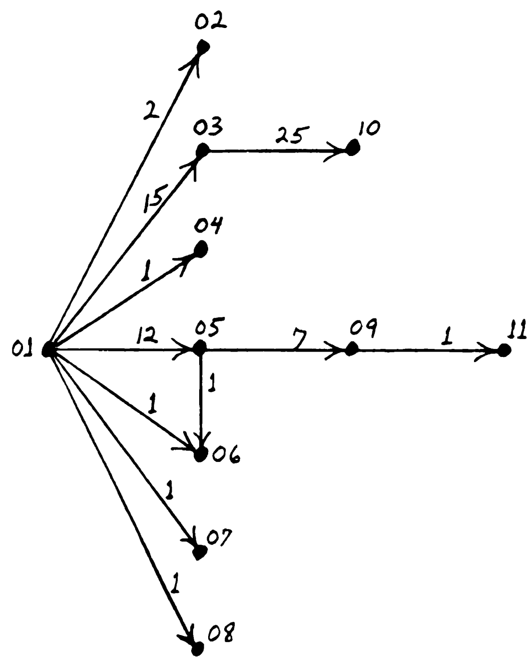


Figure C14. Sociogram of network 14.

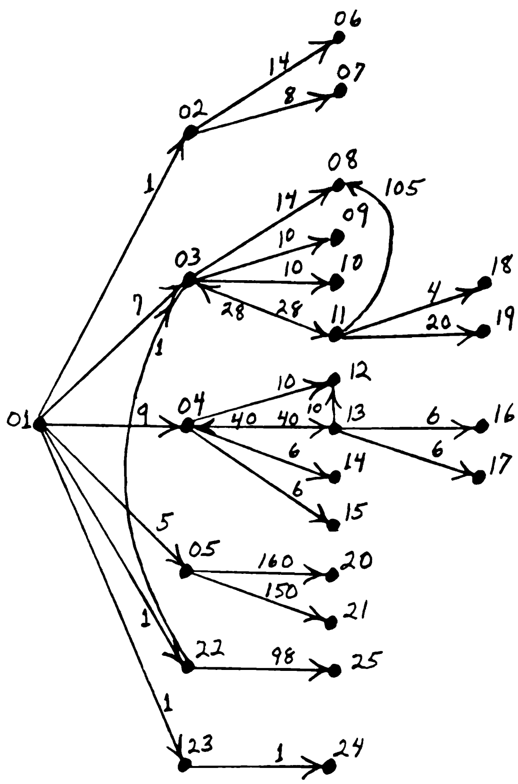


Figure C15. Sociogram of network 15.

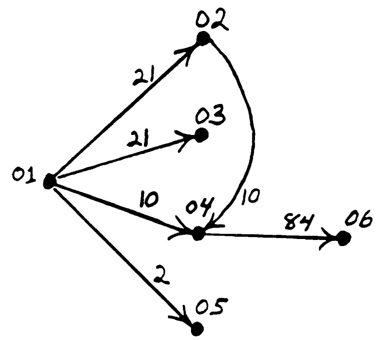


Figure C16. Sociogram of network 16.

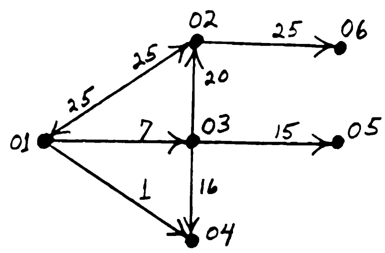


Figure C17. Sociogram of network 17.

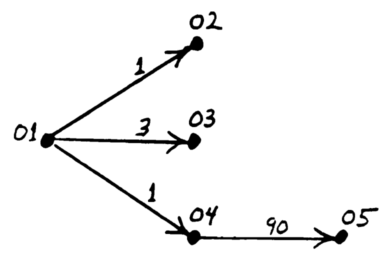


Figure C18. Sociogram of network 18.

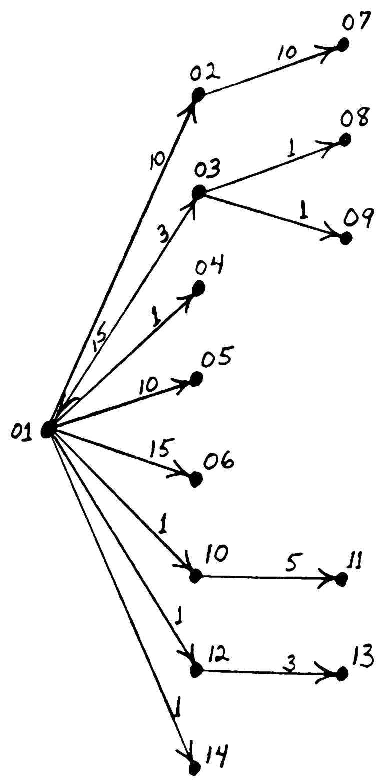


Figure C19. Sociogram of network 19.

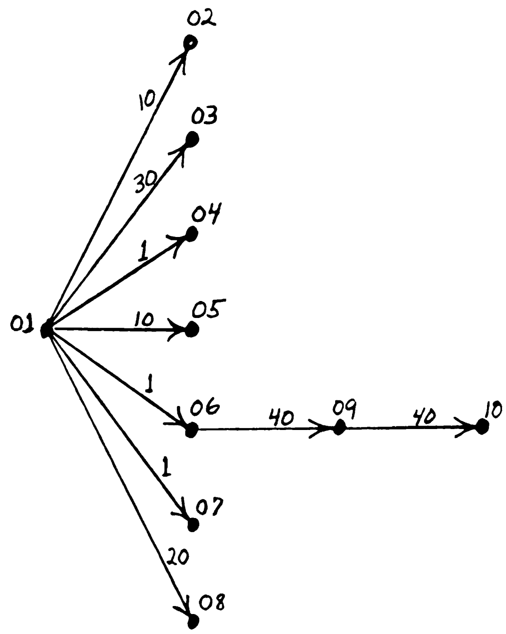


Figure C20. Sociogram of network 20.

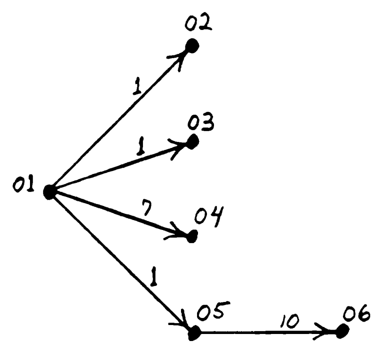


Figure C21. Sociogram of network 21.

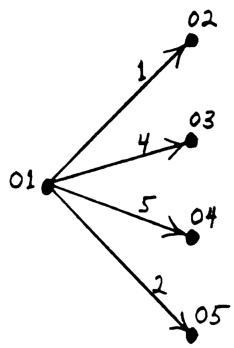


Figure C22. Sociogram of network 22.

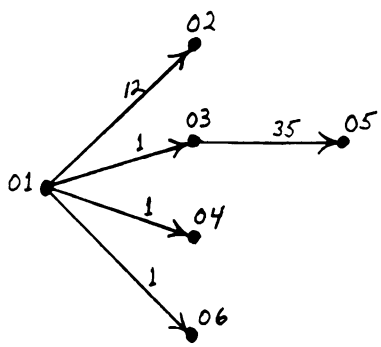


Figure C23. Sociogram of network 23.

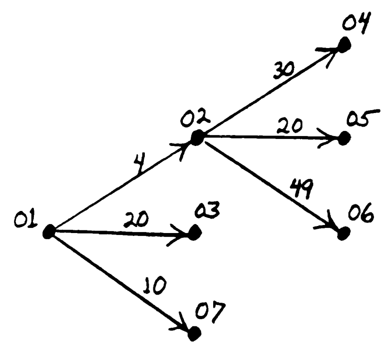


Figure C24. Sociogram of network 24.

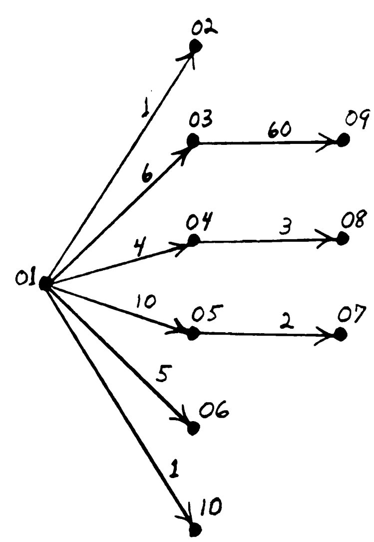


Figure C25. Sociogram of network 25.

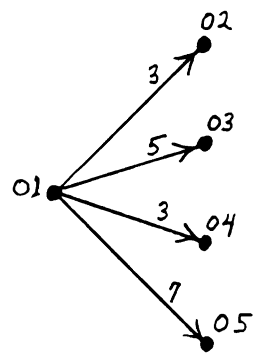


Figure C26. Sociogram of network 26.

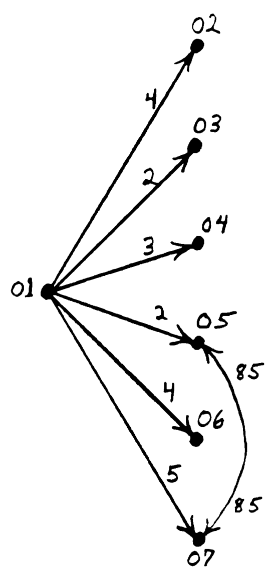


Figure C27. Sociogram of network 27.

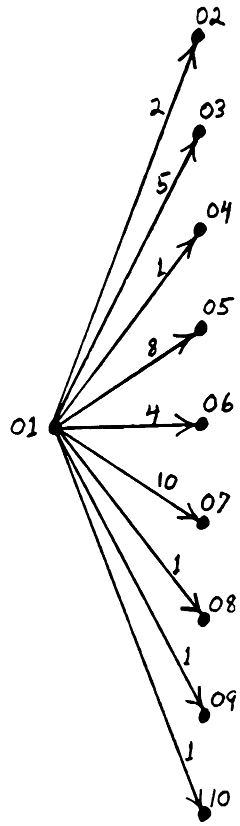


Figure C28. Sociogram of network 28.

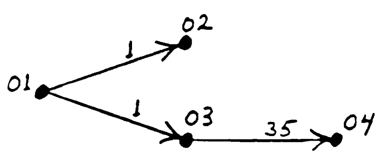


Figure 29. Sociogram of network 29.

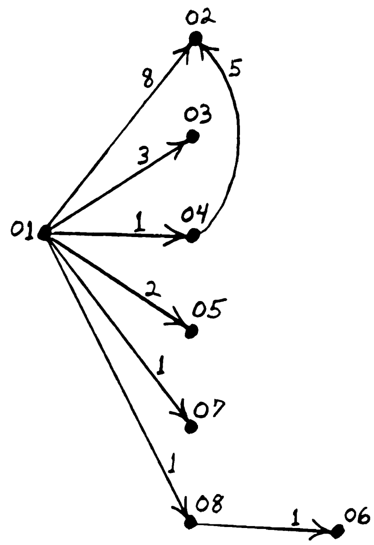


Figure C30. Sociogram of network 30.

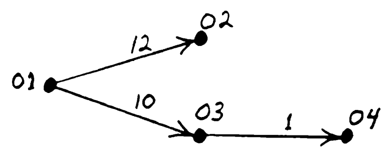


Figure C31. Sociogram of network 31.

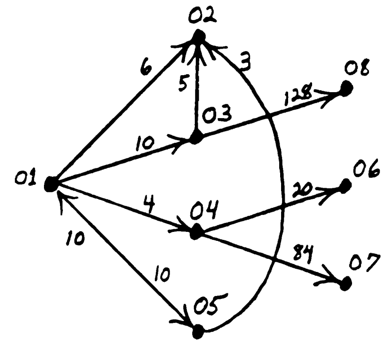


Figure C32. Sociogram of network 32.

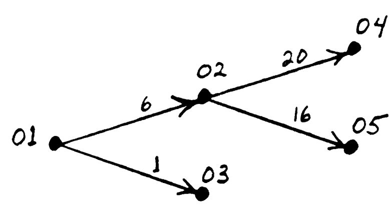


Figure C33. Sociogram of network 33.

APPENDIX D

SCATTERPLOTS OF UNTRANSFORMED VARIABLES

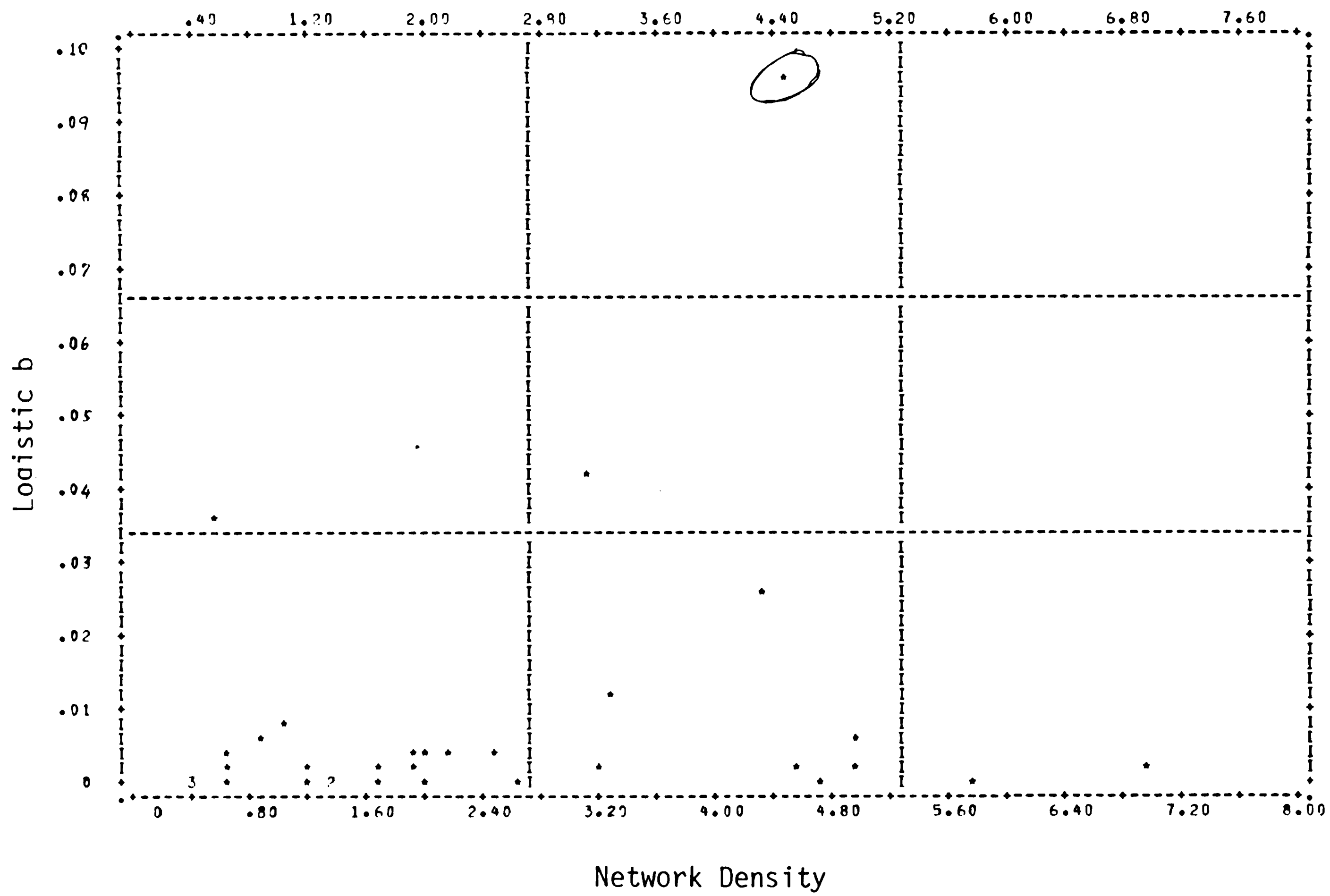


Figure D1. Scatterplot of the logistic b vs. network density (untransformed variables).

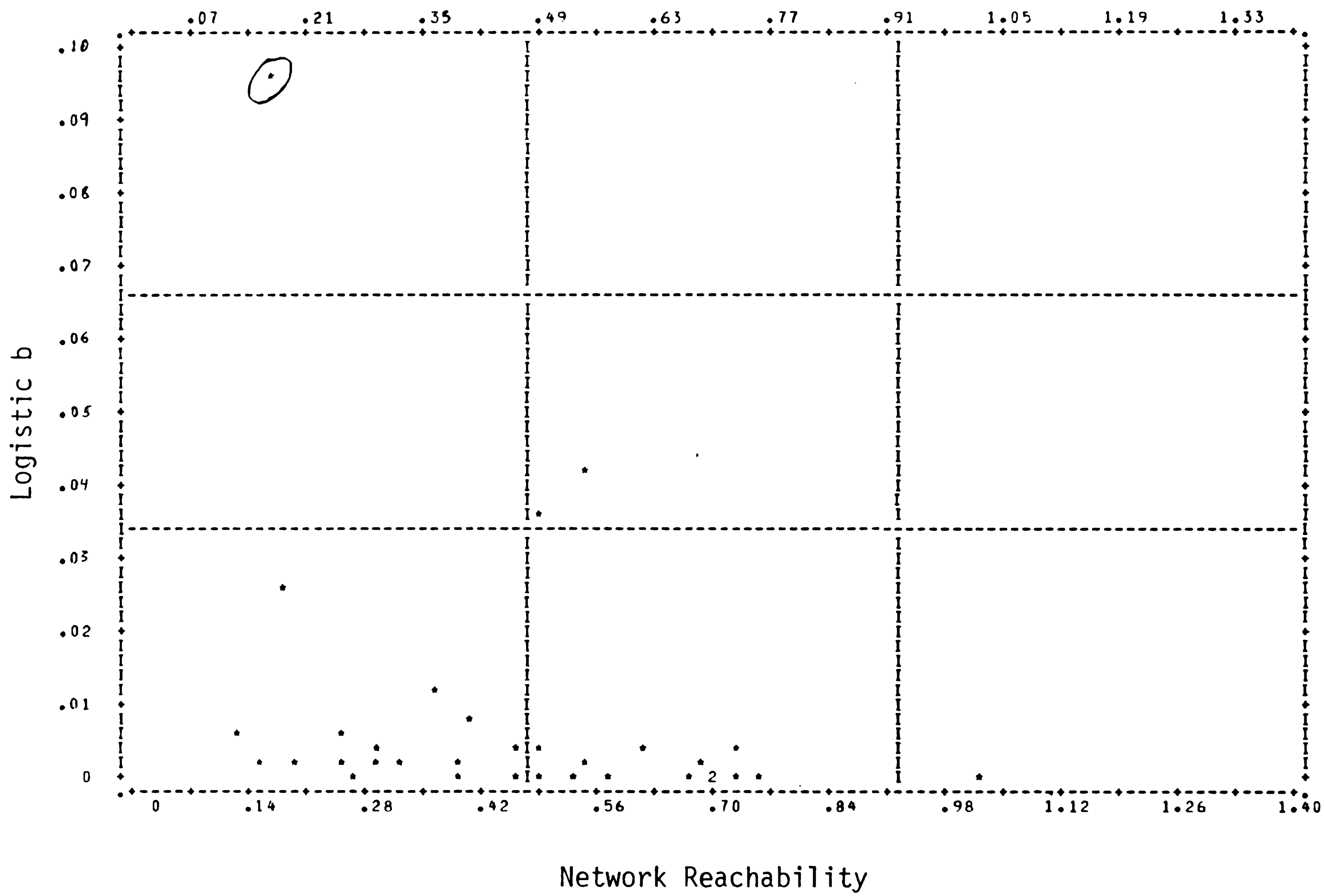


Figure D2. Scatterplot of the logistic b vs. network reachability (untransformed variables).

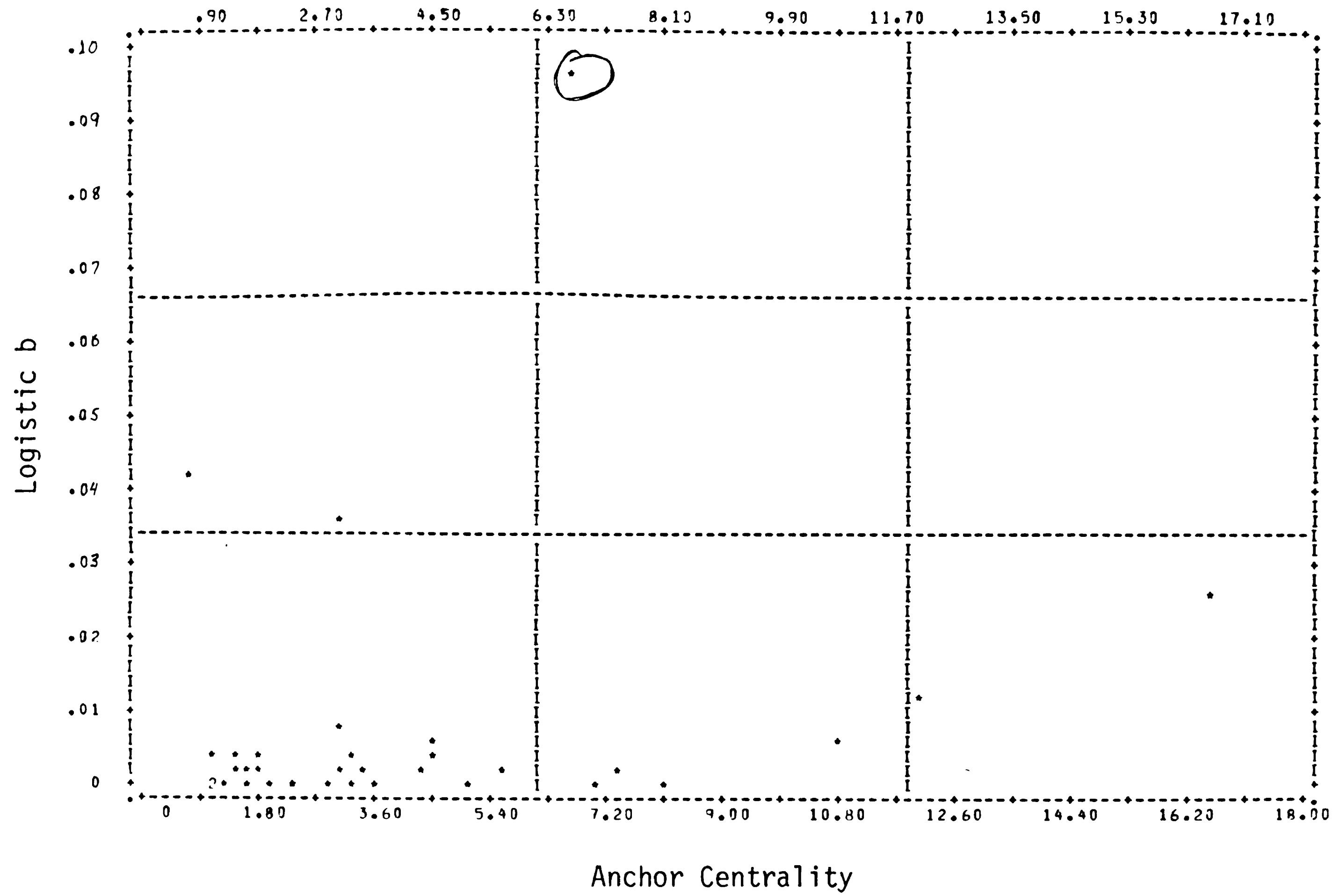


Figure D3. Scatterplot of the logistic b vs. anchor centrality (untransformed variables).

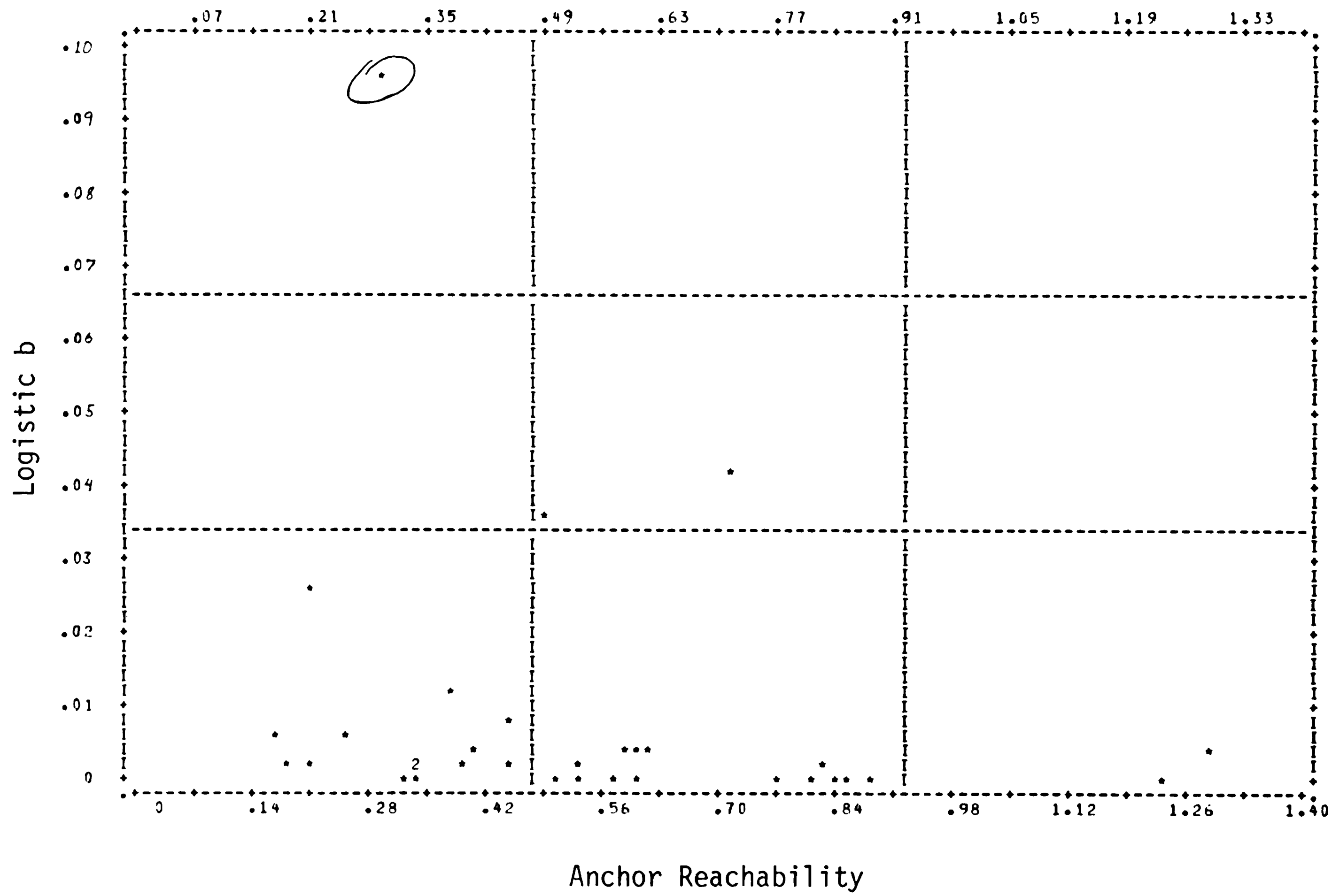


Figure D4. Scatterplot of the logistic b vs. anchor reachability (untransformed variables).

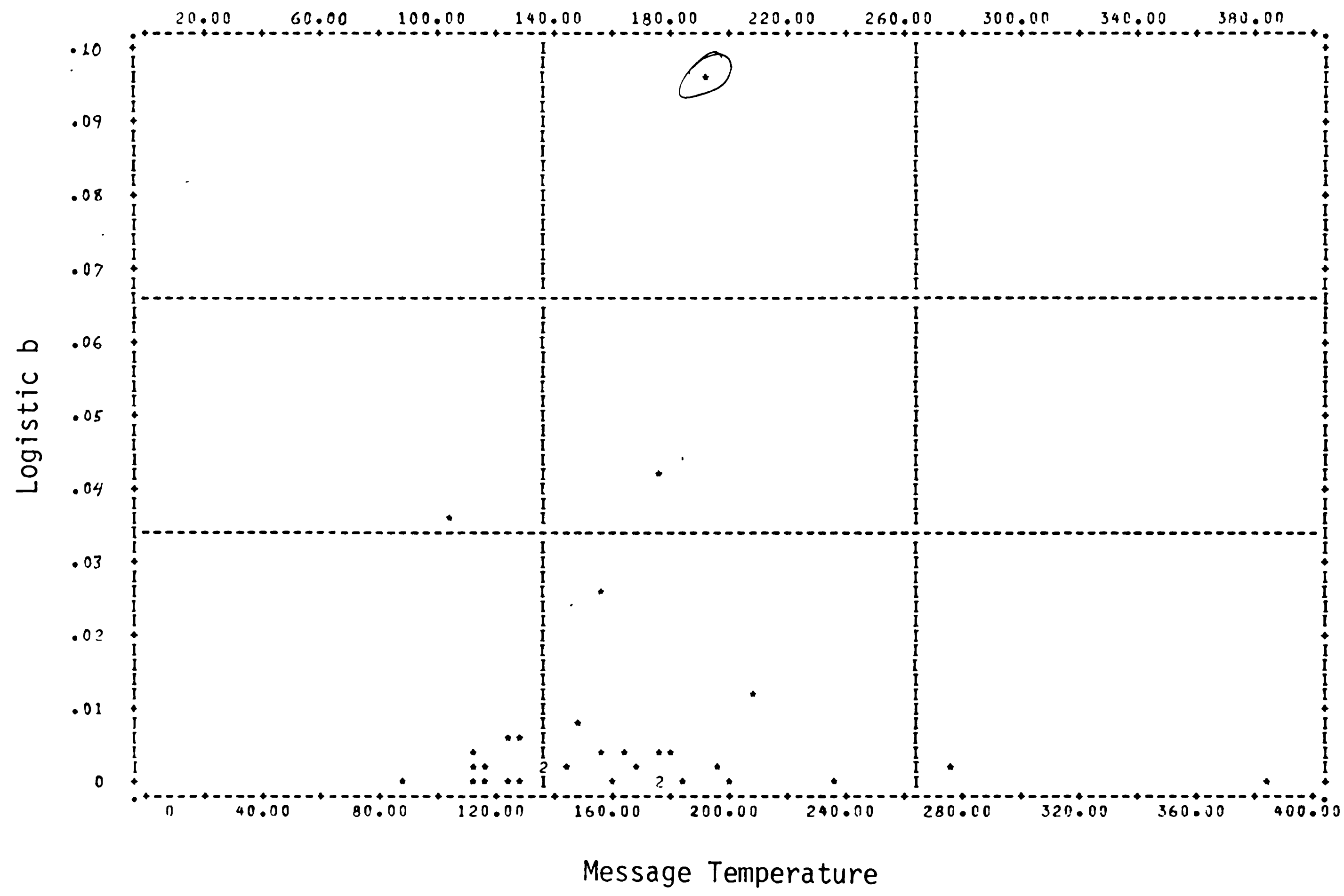


Figure D5. Scatterplot of the logistic b vs. message temperature (untransformed variables).

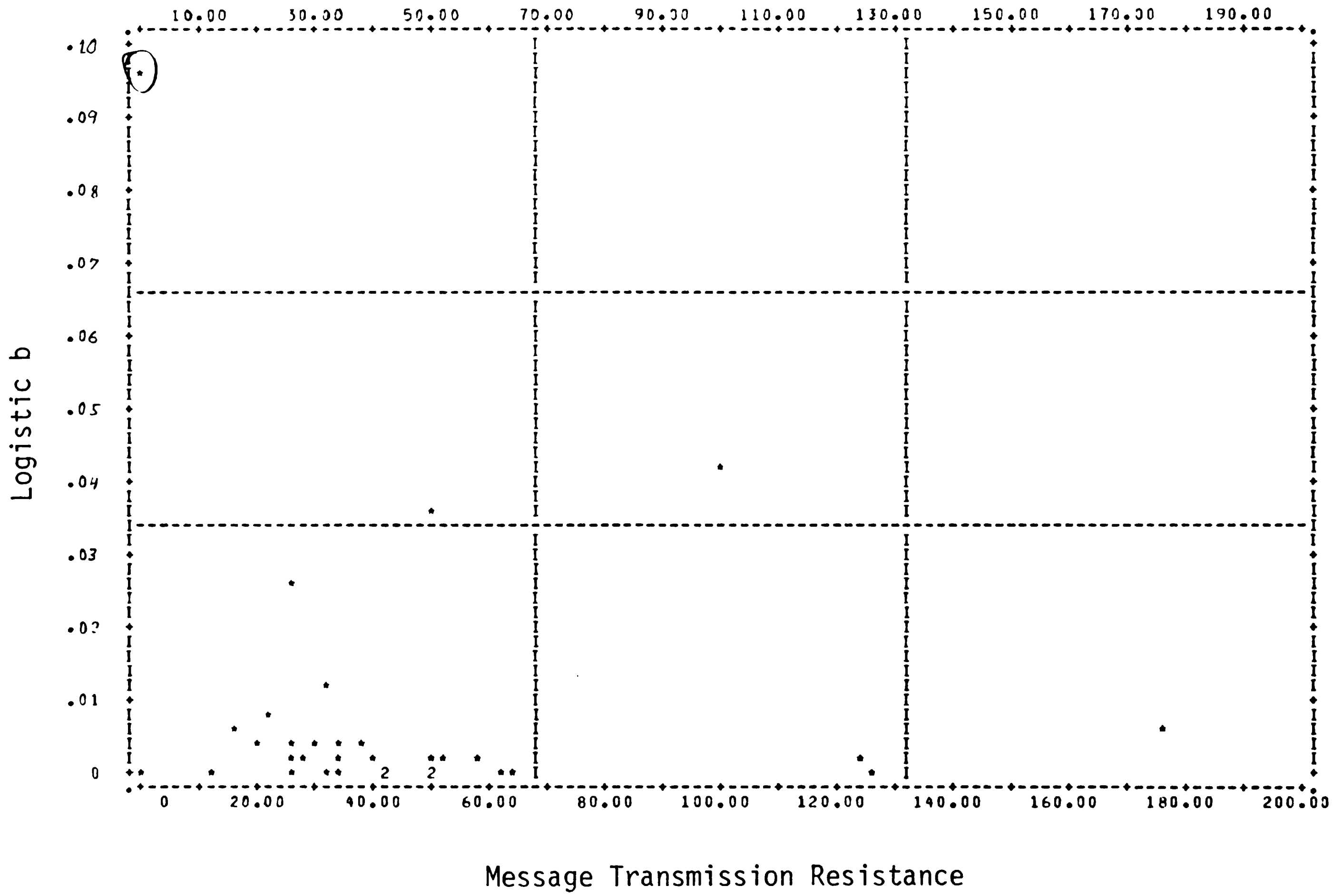


Figure D6. Scatterplot of the logistic b vs. message transmission resistance (untransformed variables).

APPENDIX E

RESIDUAL SCATTERPLOTS

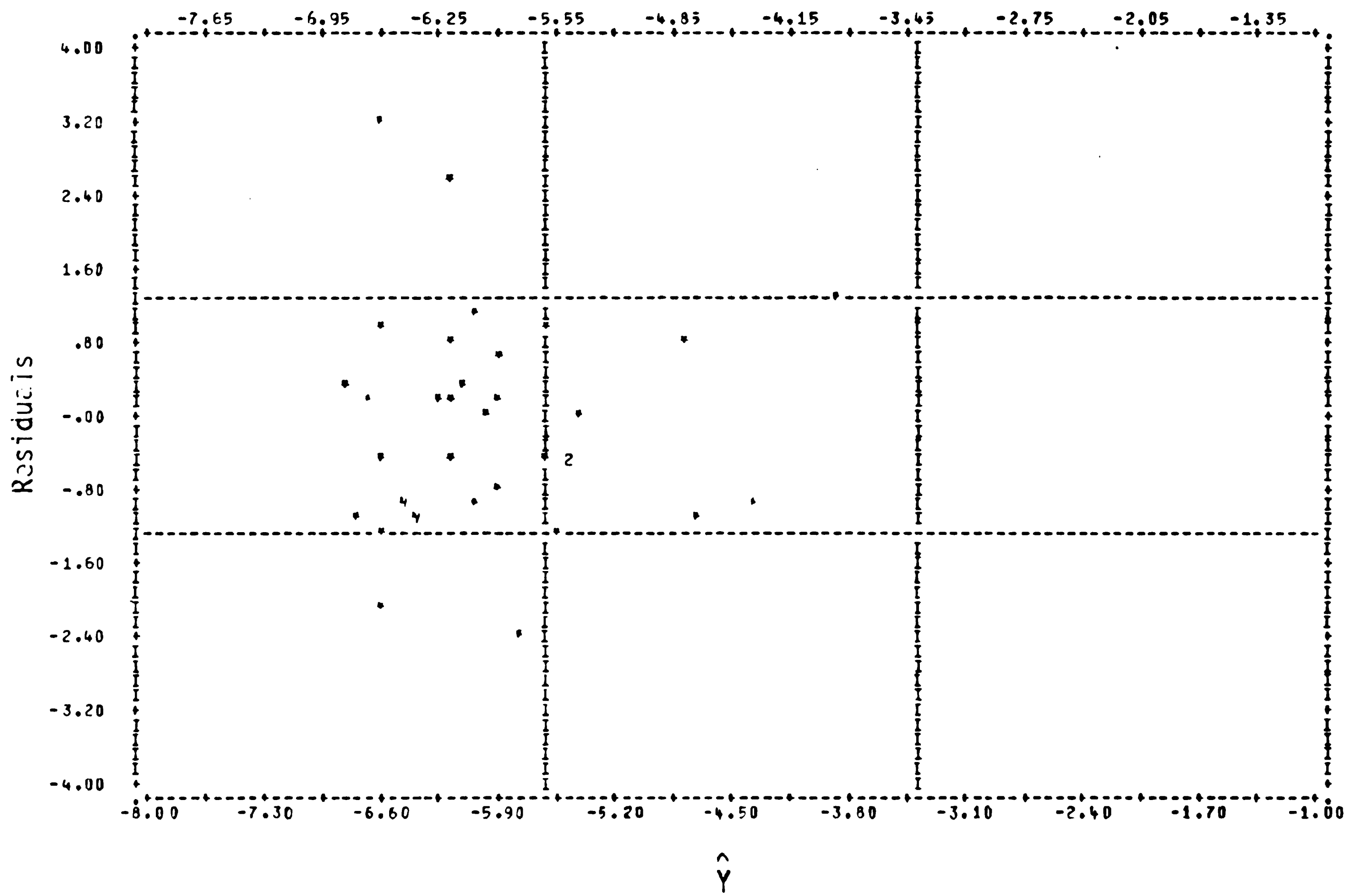


Figure E1. Residual scatterplot for equation 6.

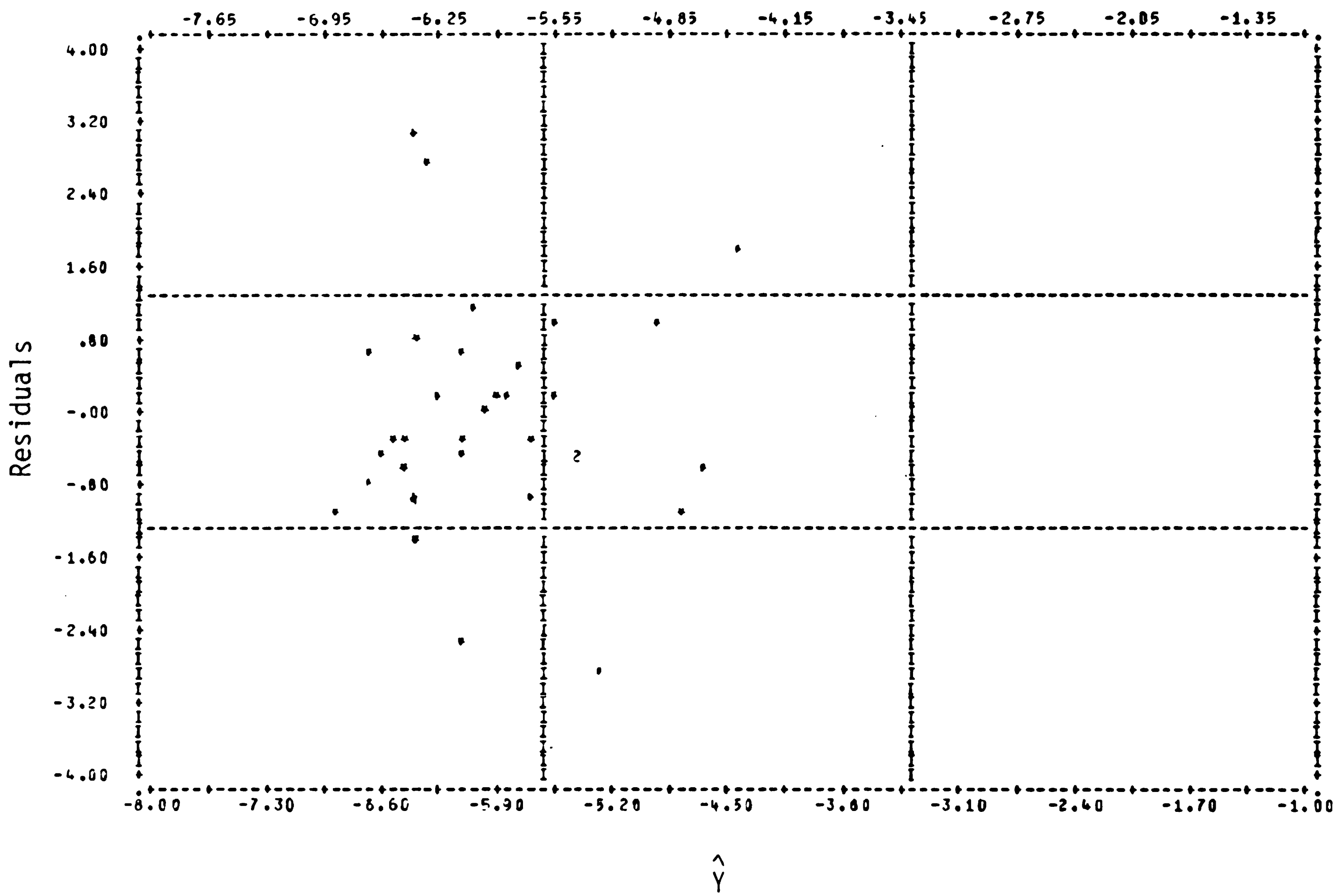


Figure E2. Residual scatterplot for equation 26.

REFERENCES

REFERENCES

Allen, R.K. A comparison of communication behaviors in innovative and non-innovative secondary schools. Unpublished Ph.D. thesis, Department of Communication, Michigan State University, 1970.

Anscombe, F.J. Special problems of statistical analysis of outliers. In D.L. Sills (Ed.), International Encyclopedia of the Social Sciences, Vol. 15. New York: Macmillan, 1968.

Arundale, R.B. Studying change over time: Criteria for sampling from continuous variables. Paper presented at the International Communication Association, Philadelphia, PA., May 1-5, 1979.

Bailey, N.T.J. The mathematical theory of epidemics (1st ed.). London: Griffin, 1957.

Bailey, N.T.J. The mathematical theory of infectious diseases and its applications (2nd ed.). New York: Hafner Press, 1975.

Barnes, J.A. Graph theory and social networks: A technical comment on connectedness and connectivity. Sociology, 1969, 3, 215-232.

Barnett, G.A. Spatial modelling of social networks with applications to the diffusion process: An initial analysis. Paper presented at Workshop on Metric Multidimensional Scaling, International Communication Association, Philadelphia, Pennsylvania, May, 1979.

Bartholomew, D.J. Stochastic models for social processes (2nd ed.). New York: John Wiley & Sons, 1973.

Bartholomew, D.J. Continuous time diffusion models with random deviation of interest. Journal of Mathematical Sociology, 1976, 4, 187-199.

Bauer, C.L. An examination of the utility of Box and Cox transformations for communication research. Unpublished manuscript, Department of Communication, Michigan State University, 1981.

Bavelas, A. A mathematical model for group structure. Applied Anthropology, 1948, 7, 16-30.

Becker, M.H. Sociometric location and innovativeness: Reformulation and extension of the diffusion model. American Sociological Review, 1970, 35, 267-282.

Bonacich, P., & Domhoff, G.W. Latent classes and group membership. Social Networks, 1981, 3, 175-196.

Bott, E. Urban families: Conjugal roles and social networks. Human Relations, 1955, 8, 345-383.

Box, G.E.P., & Cox, D.R. An analysis of transformations. Journal of the Royal Statistical Society, Series B, 1962, 26, 211-143.

Brophy, M. A study of the interrelationship between the social structure and the cognitive belief system or "culture" of a social unit. Unpublished M.A. thesis, Department of Communication, Michigan State University, 1976.

Brown, L.A. Diffusion processes and location: A conceptual framework and bibliography. Philadelphia: Regional Science Research Institute, 1968.

Budd, R.W., MacLean, M.S., Jr., & Barnes, A.M. Regularities in the diffusion of two major news events. Journalism Quarterly, 1966, 43, 221-230.

Cane, V.R. A note on the size of epidemics and the number of people hearing a rumour. Journal of the Royal Statistical Society, Series B, 1966, 28, 487-490.

Chaddha, R.L., & Chitgopekar, S.S. A "generalization" of the logistic curves and long-range forecasts (1966-1991) of residence telephones. Bell Journal of Economics and Management Science, 1971, 2, 542-560.

Coleman, J.S. Introduction to mathematical sociology. New York: Free Press, 1964.

Coleman, J., Katz, E., & Menzel, H. The diffusion of an innovation among physicians. Sociometry 1957, 20, 253-270.

Czepiel, J. Patterns of interorganizational communication and the diffusion of a major technological innovation in a competitive industrial community. Academy of Management Journal, 1975, 18, 6-24.

Daley, D.K., & Kendall, D.G. Epidemic and rumors. Nature, 1964, 204, 118.

Danes, J.E. Communication models of the message-belief change process. In B.D. Ruben (Ed.), Communication Yearbook 2. New Brunswick, N.J.: Transaction Books, 1978.

Davis, J.A. Structural balance, mechanical solidarity, and interpersonal relations. American Journal of Sociology, 1963, 68, 444-463.

Deutschman, P.J., & Danielson, W.A. Diffusion of knowledge of the major news story. Journalism Quarterly, 1960, 37, 345-355.

Dietz, K. Epidemics and rumours: A survey. Journal of the Royal Statistical Society, Series A, 1967, 130, 505-528.

Dodd, S.C. Testing message diffusion in controlled experiments: Charting the distance and time factors in the interactance hypothesis. American Sociological Review, 1953, 18, 410-416.

Dodd, S.C. Diffusion is predictable: Testing probability models for laws of interaction. American Sociological Review, 1955, 20, 392-401.

Dodd, S. Formulas for testing opinions. Public Opinion Quarterly, 1958/59, 22, 537-554.

Dodd, S.C., & McCurtain, M. The logistic diffusion of information through randomly overlapped cliques. Operational Research Quarterly, 1965, 16, 51-63.

Durkheim, E. [The rules of sociological method] (G.E.G. Catlin, Ed. and S.A. Solovay & J.H. Mueller, trans.). New York: Free Press, 1966. (Originally published, 1895.)

Edwards, J.A., & Monge, P.R. The validation of mathematical indices of communication structure. In B.D. Ruben (Ed.), Communication Yearbook I. New Brunswick, N.J.: Transaction Books, 1977.

Erbe, W. Gregariousness, group membership, and the flow of information. American Journal of Sociology, 1962, 67, 502-516.

Farace, R.V., & Mabee, T. Communication network analysis methods. In P.R. Monge & J.N. Cappella (Eds.), Multivariate techniques in human communication research. New York: Academic Press, 1980.

Farace, R.V., Monge, P.R., & Russell, H.M. Communicating and organizing. Reading, Mass.: Addison-Wesley, 1977.

Fathi, A. Diffusion of a 'happy' news event. Journalism Quarterly, 1973, 50, (2), 271-277.

Feldman, J., & Lie, S.S. Some simple mathematical models for adoption processes: The case of the comprehensive school reform in Norway. Quality and Quantity, 1974, 8, 299-325.

Freeman, L.C. A set of measures of centrality based on betweeness. Sociometry, 1977, 40, 35-41.

Freeman, L.C. Centrality in social networks conceptual clarification. Social Networks, 1978/79, 1, 215-239.

Friedkin, N. A test of structural features of Granovetter's strength of weak ties theory. Social Networks, 1980, 2, 441-442.

Friedkin, N.E. The development of structure in random networks: An analysis of the effects of increasing network density on five measures of structure. Social Networks, 1981, 3, 41-52.

Funkhouser, G.R. A probabilistic model for predicting news diffusion. Journalism Quarterly, 1970, 47, 41-45.

Funkhouser, G.R., & McCombs, M.E. Predicting the diffusion of information to mass audiences. Journal of Mathematical Sociology, 1972, 2, 121-130.

Garabedian, P.G., & Dodd, S.C. Clique size as a factor in message diffusion. Sociological Inquiry, 1962, 32, 71-81.

Giancoli, D.C. Physics: Principles with applications. Englewood Cliffs, N.J.: Prentice-Hall, 1980.

Goffman, W. Mathematical approach to the spread of scientific ideas. The history of mast cell research. Nature, 1966, 212, 449-452.

Goffman, W., & Newill, V.A. Generalization of epidemic theory, and application to the transmission of ideas. Nature, 1964, 204, 225-228.

Granovetter, M. The strength of weak ties. American Journal of Sociology, 1973, 78, 1360-1380.

Granovetter, M. Network sampling: Some first steps. American Journal of Sociology, 1976, 81 (6), 1287-1303.

Gray, L.N., & von Broembsen, M.H. On simple stochastic diffusion models. Journal of Mathematical Sociology, 1974, 3, 231-244.

Greenberg, B.S. Diffusion of news of the Kennedy assassination. Public Opinion Quarterly, 1964a, 28, 225-232.

Greenberg, B.S. Person-to-person communication in the diffusion of news events. Journalism Quarterly, 1964b, 41, 489-494.

Greenberg, B.S., Brinton, J.E., & Farr, R.S. Diffusion of news about an anticipated major news event. Journal of Broadcasting, 1965, 9, 129-142.

Griliches, Z. Hybrid corn: An exploration in the economics of technological change. Econometrica, 1957, 25, 501-523.

Guimaraes, L.L. Communication integration in modern and traditional social systems: A comparative analysis across twenty communities of Minas Gerais, Brazil. Unpublished Ph.D. thesis, Department of Communication, Michigan State University, 1972.

Hägerstrand, T. Innovation diffusion as a spatial process. Lund, Sweden: C.W.K. Gleerup, 1953.

Hamblin, R.L., Jacobsen, R.B., & Miller, J.L.L. A mathematical theory of social change. New York: John Wiley & Sons, 1973.

Hammer, M. Social access and the clustering of personal connections. Social Networks, 1980, 2, 305-325.

Hanushek, E.A., & Jackson, J.E. Statistical methods for social scientists. New York: Academic Press, 1977.

Harary, F., & Batell, M.F. What is a system? Social Networks, 1981, 3, 29-40.

Harrary, F., Norman, R.Z., & Cartwright D. Structural models: An introduction to the theory of directed graphs. New York: Wiley & Sons, 1965.

Haroldsen, E.O., & Harvey, K. The diffusion of 'shocking' good news. Journalism Quarterly, 1979, 56, 771-775.

Hill, R.J., & Bonjean, C.M. News diffusion: A test of the regularity hypothesis. Journalism Quarterly, 1964, 41, 336-342.

Huang, C., & Moon, H.C. B-C: A Box-Cox procedure estimation routine user guide. Faculty Series, Division of Agricultural Economics, College of Agriculture, University of Georgia, 1978.

Hyman, H.H., & Sheatsley, P.B. Some reasons why information campaigns fail. Public Opinion Quarterly, 1947, 11, 412-423.

Karmeshu, & Pathria, R.K. Stochastic evolution of a nonlinear model of diffusion of information. Journal of Mathematical Sociology, 1980, 7, 59-71.

Katz, E. The two-step flow of communication: An up-to-date report on an hypothesis. Public Opinion Quarterly, 1957, 21, 61-78.

Katz, E., & Lazarsfeld, P.F. Personal influence: The part played by people in the flow of most communication. Glencoe: The Free Press, 1955.

Katz, E., Levin, M.L., & Hamilton, H. Traditions of research on the diffusion of innovations. American Sociological Review, 1963, 28, 237-252.

Landahl, H.D. On the spread of information with time and distance. Bulletin of Mathematical Biophysics, 1953, 15, 367-381.

Landau, H.G., & Rapoport, A. Contribution to the mathematical theory of contagion and spread of information. I: Spread through a mixed population. Bulletin of Mathematical Biophysics, 1953, 75, 173-183.

Laroche, M. A model of attitude change in groups following a persuasive communication: An attempt at formalizing research findings. Behavioral Science, 1977, 22, 246-257.

Larson, O.N., & Hill, R.J. Mass media and interpersonal communication in the diffusion of a news event. American Sociological Review, 1954, 19 (4), 426-433.

Larson, O.N., & Hill, R.J. Social structure and interpersonal communication. American Journal of Sociology, 1958, 63, 497-505.

Lave, C.A., & March, J.G. An introduction to models in the social sciences. New York: Harper & Row, 1975.

Lazarsfeld, P.F., Berelson, B., & Gaudet, H. The people's choice (2nd ed.). New York: Columbia University Press, 1948.

Lin, N., & Burt, R.S. Differential effects of information channels in the process of innovation diffusion. Social Forces, 1976, 54, 256-274.

Lionberger, H.F., & Hassinger, E. Neighborhoods as a factor in the diffusion of farm information in a northwest Missouri farming community. Rural Sociology, 1954, 19, 377-384.

Liu, W.T., & Duff, R.W. The strength of weak ties. Public Opinion Quarterly, 1972, 48 (1), 361-366.

McLuhan, M. Understanding media: The extensions of man. New York: McGraw-Hill, 1964.

Menzel, H., & Katz, E. Social relations and innovations in the medical profession: The epidemiology of a new drug. Public Opinion Quarterly, 1955, 19, 337-352.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Bent, C. SPSS: Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill, 1975.

Mitchell, J.C. The concept and use of social networks. In J.C. Mitchell (Ed.), Social networks in urban situations. Manchester: University Press, 1969.

Monge, P.R., Edwards, J.A., & Kirste, K.K. The determinants of communication and communication structure in large organizations: A review of research. In B.D. Ruben (Ed.), Communication Yearbook 2. New Brunswick, N.J.: Transaction Books, 1978.

Monin, J.P., Benoyoun, R., & Sert, B. [Initiation to the mathematics of the process of diffusion, contagion and propagation] (M. Brandon, trans.). Paris: Mouton, 1976.

Nisbet, R. The sociology of Emile Durkheim. New York: Oxford University Press, 1974.

Olinick, M. An introduction to mathematical models in the social and life sciences. Reading, Mass.: Addison-Wesley, 1978.

Peay, E.R. Connectedness in a general model for valued networks. Social Networks, 1980, 2, 385-410.

Pemberton, H.E. The curve of cultural diffusion. American Sociological Review, 1936, 1, 547-556.

Rapoport, A. Spread of information through a population with socio-structural bias: I. Assumption of transitivity. Bulletin of Mathematical Biophysics, 1953a, 15, 523-533.

Rapoport, A. Spread of information through a population with socio-structural bias: II. Various models with partial transitivity. Bulletin of Mathematical Biophysics, 1953b, 15, 535-546.

Rapoport, A. The diffusion problem in mass behavior. General Systems, 1956, 1, 48-55.

Rapoport, A. A probabilistic approach to networks. Social Networks, 1979, 2 (1), 1-18.

Rapoport, A., & Rebhun, L. On the mathematical theory of rumour spread. Bulletin of Mathematical Biophysics, 1952, 14, 375-383.

Richards, W.D., Jr. A manual for network analysis (using the NEGOPY network analysis program). Unpublished paper, Institute for Communication research, Stanford University, 1975.

Richards, W.D. A coherent systems methodology for the analysis of human communication systems. Unpublished doctoral dissertation, Stanford University, 1976.

Rogers, E.M. Network analysis of the diffusion of innovations. In P.W. Holland & S. Leinhardt (Eds.), Perspectives on social network research. New York: Academic Press, 1979.

Rogers, E.M., & Bhowmik, P.K. Homophily-heterophily: Relational concepts for communication research. Public Opinion Quarterly, 1971, 34, 523-538.

Rogers, E.M., & Kincaid, D.L. Communication networks: Toward a new paradigm for research. New York: Free Press, 1981.

Rogers, E.M., with Shoemaker, F.F. Communication and innovations: A cross-cultural approach (2nd ed.). New York: Free Press, 1971.

Rosen, S., & Tesser, A. On reluctance to communicate undesirable information: The MUM effect. Sociometry, 1970, 33, 253-263.

Schneider, I.A., & Fett, J.H. Diffusion of mass media messages among Brazilian farmers. Journalism Quarterly, 1978, 53, 494-500.

Shannon, C.E., & Weaver, W. The mathematical theory of communication. Urbana, Ill.: University of Illinois Press, 1949.

Shoemaker, F.F. System variables and educational innovativeness in Thai government secondary schools. Unpublished Ph.D. thesis, Department of Communication, Michigan State University, 1971.

Simmel, G. The sociology of Georg Simmel (K.H. Wolff, Ed. and trans.). New York: Free Press, 1950.

Simmel, G. [The web of group-affiliations] (R. Bendix, trans.). New York: Free Press, 1955. (Originally published, 1922).

Spitzer, S.P. Mass media vs. personal sources of information about the presidential assassination: A comparison of six investigations. Journal of Broadcasting, 1964-1965, 9, 45-50.

Sørensen, Aa. B. Mathematical models in sociology. Annual Review of Sociology, 1978, 4, 345-371.

Taga, Y., & Ish, K. On a stochastic model concerning the pattern of communication. Diffusion of news in a social group. Annuals Institute Statistical Mathematics, 1959, 11, 25-43.

Tarde, G. The laws of imitation. Gloucester, Mass.: Henry Holt, 1903.

Tesser, A., & Rosen, S. The reluctance to transmit bad news. In L. Berkowitz (Ed.), Advances in experimental social psychology. New York: Academic Press, 1975.

von Bertalanffy, L. An outline of general systems theory. British Journal of the Philosophy of Science, 1950, 1 134-165.

Welch, R.E., Jr. The use of magnitude estimation in attitude scaling: Constructing a measure of political dissatisfaction. In D. Nimmo (Ed.), Political communication and public opinion in America. Santa Monica, Calif.: Goodyear, 1978.