GROUPS: INTERACTION AND PERFORMANCE

Joseph E. McGrath
University of Illinois, Urbana

Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632

Library of Congress Cataloging in Publication Data

McGrath, Joseph Edward
Groups, interaction and performance.

Bibliography: p.
Includes index.
1. Social groups. 2. Social interaction.
3. Performance. I. Title.
HM131.M377 1983 302.3 82-24016
ISBN 0-13-365700-0

Editorial/production supervision: Dee Amir Josephson
Cover design: 20/20 Services, Inc.
Manufacturing buyer: Ron Chapman

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-365700-0

Prentice-Hall International, Inc., *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Prentice-Hall of Southeast Asia Pte. Ltd., *Singapore*
Whitehall Books Limited, *Wellington, New Zealand*

A CONCEPTUAL FRAMEWORK FOR THE STUDY OF GROUPS

There are many different perspectives from which one can view a group, and many ambiguities already noted in defining groups and their membership. For such a complex and ambiguous set of concepts, it is often useful to adopt a frame of reference, a map, that models or lays out systematically the various parts of the topic as a research problem. This section offers such a conceptual model for the study of groups (see Figure 1-1).

The point of such a model is to lay out the underlying logic of the problem in a way that can serve as a guiding framework for exploring the problem in its various aspects. For a complex problem, you cannot study everything at once, you cannot think about everything at the same time. This kind of model lets us take the total problem apart, so we can think about and examine evidence about a manageable chunk of it, and then be able to fit the parts back together again. Furthermore, such a framework tells us what batches of things to look at—what sets of variables are likely to be important—and at the same time offers a logic for deciding what sets of relations among these variables are likely to be important to consider.

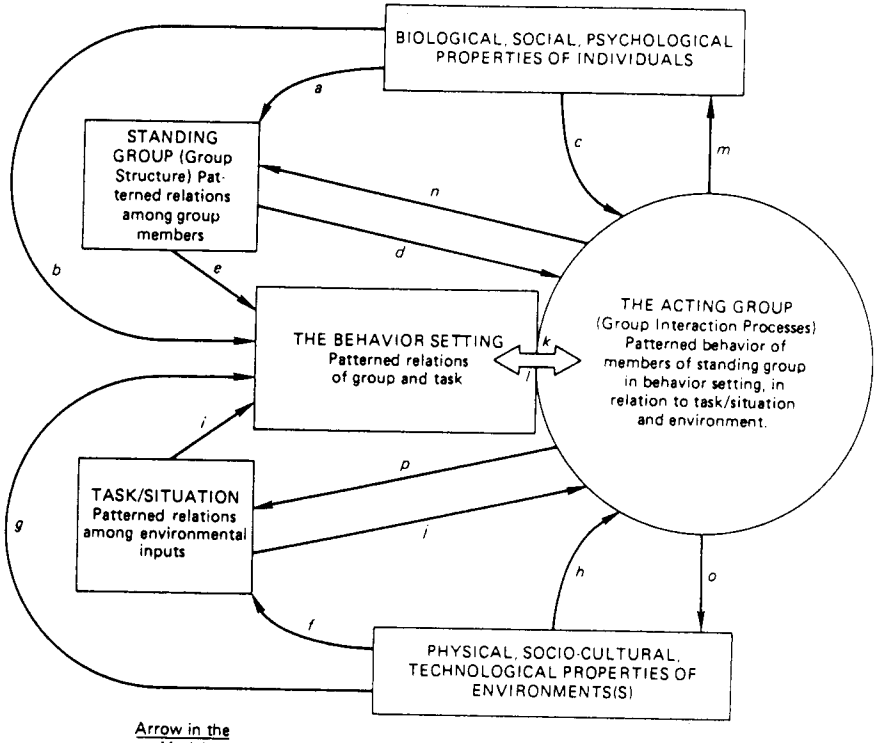
Note that this is intended to be a model of the problem (i.e., studying groups systematically), rather than a theory or model of groups. Such models are sometimes called "metatheories." They reflect a way of looking at the problem that encompasses a whole family of possible substantive theories. But they do not specify any one particular theory. Here, we are talking about *classes* of properties or variables, and the logical relations between those classes. But there is no specification of specific sets of relations between specific sets of variables—as there would be in a substantive theory.

Main Classes of Variables

The central feature, the "essence," of a group lies in the interaction of its members—the *behaving together*, in some recognized relation to one another, of two or more people who also have some past and/or future relation to each other. So *group interaction process* is the centerpiece of the model.

Certain things go into that group process. For one thing, there are participants, or group members. They come to a group interaction with all their "properties" (traits, characteristics, beliefs, habits, etc.). A member may be strong, or extroverted, or wise, or old, or female, or bellicose, or clumsy, or many other things. *Some* of these properties of members may affect group interaction. So, if one wants to understand and perhaps predict aspects of group interaction process, one must take these group member properties into account.

These participants make up the group being considered, and one can think about the pattern of relations among group members, *prior to* any group interaction process, as another batch of potentially important properties or



Arrow in the Model	Relations Implied
a	Member composition of focal group.
b and *c*	Extramember effects of individuals on Behavior Setting and on Group Interaction Process.
d and *e*	Effects of Standing Group (group structure) on Behavior Setting and on Group Interaction Process.
f	Environmental factors as they structure the Task/Situation.
g and *h*	Extratask effects of Environment on Behavior Setting and on Group Interaction Process.
i and *j*	Effects of Task/Situation on Behavior Setting and on Group Interaction Process.
k and *l*	The dynamic relation between the Behavior Setting and Group Interaction Process.
m and *n*	Effects of Group Interaction Process on members, and on the Standing Group.
o and *p*	Effects of Group Interaction Process on Task/Situation, and on the Environment.

FIGURE 1-1 A Conceptual Framework for the Study of Groups

variables. Do group members like each other? Do they have differential influence on each other (for example, does one person exercise more leadership or dominance than the others)? How many members are there and how long have they belonged to this group? Group members are related to each other in many ways; a lot of those relations affect how they behave in relation to one another when they interact. These patterns of relations among members—aspects of *group structure*—also must be taken into account if one wants to understand and predict group interaction process.

Group interaction takes place somewhere, in some *environment*. It may involve a group of workers doing their jobs in an assembly plant; a set of executives holding a conference in a company meeting room; a County Planning Board having its monthly meeting; a family eating dinner on a Wednesday evening in April; a football team getting a dressing room talk between halves of a game; a group of kids playing with some old tires in a dump; two couples at a night-club; an airplane crew flying from Texas to Toronto; a Broadway company rehearsing in a theater. In all of these cases, the group interaction is taking place in an environment that includes both physical and social aspects. Many of these can make a difference in how members behave, hence can alter group interaction process.

Group interaction not only takes place somewhere, it involves the group *doing something*. One very important aspect of all of those settings just enumerated is the "task." Any group interaction (actually, any intact portion of such an interaction) can be characterized in terms of the task(s) that the group (or its members) is trying to carry out: giving (and receiving) a lecture or a sermon or a play; processing steel; assembling an auto; choosing a new vice president; deciding on a zoning variance; preparing a budget justification; arbitrating a grievance; enjoying dinner; having a good time at the nightclub, on the backpacking trip, or in the dump. The task, as you can see from those examples, involves informally assumed goals (e.g., having a good time) as well as assigned jobs (e.g., assembling an auto). What the group is doing, or trying to do, as well as where this is taking place, affects group interaction process in many ways. So, the task situation represents another class of "factors" one must take into account if one wishes to understand and predict group interaction process.

These major classes of inputs—properties of group members; properties of the standing group (group structure); properties of the task/situation; and properties of the surrounding environment—set the conditions under which group interaction takes place. Furthermore, the effects of these four sets of properties, singly and in combination, are forces that shape the group interaction process.

The group interaction process itself is both the result of these shaping forces and the source of some additional forces. While group interaction is greatly affected by those sets of input variables—properties of members, of the group, of the task, and of the environment—it is also patterned, in part, by forces internal to (or indigenous to) the interaction process itself. The latter part of this chapter delves further into the internal forces of group interaction process.

Furthermore, the interaction process and its results represent sources (forces) that potentially lead to changes in those very input conditions: changes in the members themselves; changes in the group structure, or the patterns of relations among members; and changes in the relation of the group to its tasks and to its environment. So, these sets of outputs (or outcomes, or consequences) of group interaction process are parallel to the input classes and, in fact, represent changes in those input variables.

These classes of factors, or "panels" of potentially important variables, are related to one another in relatively complex ways. These panels, and the relations among them, are diagrammed in Figure 1-1. The parts of that model are discussed next.

A Model of Effects by and on Groups

The conceptual framework for study of groups starts with two givens: individual people, who are the members of the group in question (what will be referred to, at times, as the focal group, for clarity of reference); and the environment in which those people are embedded. So we begin with two panels of potentially relevant properties: properties of the group members as individuals; and properties of the physical, socio-cultural and technological environment(s). The former panel includes biographical and demographic characteristics (age, gender, etc.); personality dispositions; beliefs, attitudes and values; moods, feelings, states of mind; and drives, needs, motives, goals and expectations. The latter, environmental, panel includes conditions of the general physical environment (noise, heat, lighting, etc.) and of the social environment (inter-group conflict, loyalty, alienation, etc.).

Both of these panels of variables are huge, perhaps even infinite. So it is necessary to be very selective in terms of what properties are to be included in a study. Such selectivity is one of the functions of theory, as noted earlier. That is, theory functions as a guide to the investigator in selecting variables for study that are thought to be germane to the problem.

When people become interrelated, as when they are members of a group, they develop patterned relationships among themselves—patterned in terms of status, of power, of affection, and of many other aspects. These patterned relationships among group members constitute a group structure. There are many such patterns, such group structures—as many as there are variables or properties on which members can be connected to one another. These include, *at least:* structures defined in terms of composition of members; structure defined in terms of division of labor on tasks; communication structures; power structures and interpersonal relations structures. In the model, the *collection* of all these structures is called the *standing group* (to distinguish it from the *acting group*).

Environmental properties, too, are patterned; and one particular portion is of special importance in the present discussion. That important part is the set of environmental demands/constraints/opportunities that combine to form a

particular task and situation. Environmental properties "play into" more than one task/situation, of course, and even more than one at the same time, just as group members "belong to" more than one group, and even more than one at the same time. So, for clarity, we probably should designate our referent as the focal task/situation, recognizing that the environment abounds with "tasks."

We can consider the juxtaposition of the standing group and the task as the Behavior Setting. The term, behavior setting, is borrowed from the work of Roger Barker and his colleagues (Barker, 1965; Barker & Wright, 1955). But the reader should be warned that I am changing the use of that term in one important respect. When Barker talks of the behavior setting, he is dealing with individuals behaving in environments, or individuals behaving in task/situations; but Barker does not use concepts of group, group structure, or group process at all. Barker sees individuals, and their behavior, as related to one another primarily through the demands of the situation.

In the model, the behavior setting represents a pattern—a fit—between the group as a structured entity (the standing group) and the task/situation as a structured set of requirements/demands/opportunities/possibilities/constraints. Notice, too, that the framework has both properties of individuals and properties of the environment "playing into" the behavior setting directly, as well as indirectly through the group and the task. This is equivalent to saying that, while a particular concert (behavior setting and group interaction process) is to be viewed as *mainly* a juxtaposition of a particular orchestra (standing group) with a particular set of musical compositions (task/situation), properties of the orchestra members (M) and of the concert hall, the city, and perhaps the time of year (E), can also have effects on the results.

All of these form the "inputs" for what I am calling group interaction process (GIP), or the *acting group*. GIP refers to the *processes* that take place when group members actually interact, in behavior settings that carry task structures and environmental effects. Such activity can be described in terms of many processes, including (at least) general structural properties such as level and rate of interaction, distribution of participation, extent of member involvement, and so forth, all of which might be labeled *morphological properties;* the flow of work; the flow of information or communications; the flow of influence; and the flow of interpersonal affect. The *acting group* is the term used in this book for the collection of all of these interactive processes. In a sense, the behavior setting refers to the time-place-thing-person complex that serves as the site for the behavior of the acting group. The acting group and the behavior setting are the "action" and "state" sides of the same coin. In Barker's terms, the behavior setting is "circumjacent to" the group interaction process. This is represented in Figure 1-1 by showing the behavior-setting-to-group-interaction-process relation, and the reciprocal relation, as a double arrow, K and L.

The group interaction process feeds back into, and has effects on, all the panels of input variables out of which it has sprung. Individuals are often changed (for example, their attitudes are influenced) as a result of being members of an acting group. Group interaction can change the structure of the standing group; for example, it can change the pattern of attraction among members. Group interaction sometimes results in effects on the environment;

and it quite often results in a shift in the relation of the focal group to its task/situation. Such changes are usually dealt with in terms of task performance effectiveness or task productivity.

All of these effects (the eleven input arrows, *a* to *k*, and the five feedback arrows, *l* to *p*, in Figure 1-1) are important in principle, and are worthy of study. But many of them have been more thoroughly studied than others, and some of them are of more theoretical or practical significance than others. So the organization of later parts of this book will reflect selective treatment of some of these classes of relations more thoroughly than others. One basis for the selection of particular sets of relations for special attention is my particular conception of the interaction process and what it entails. That conceptualization will be presented next.

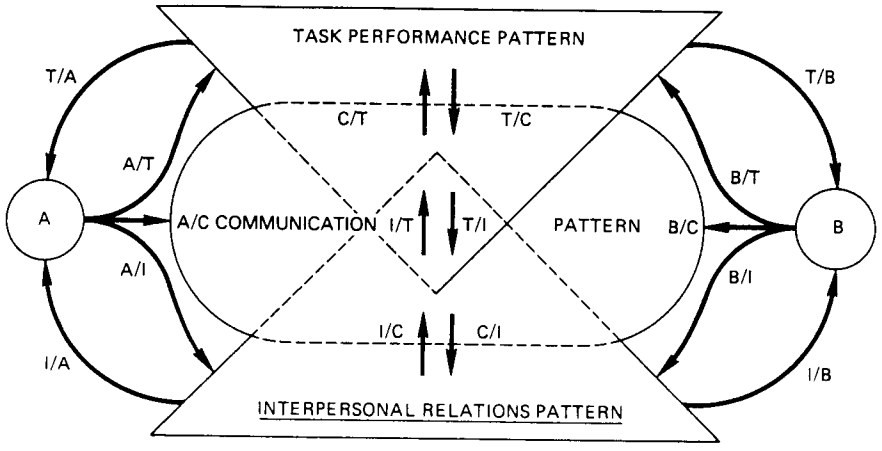
A MICRO-VIEW OF THE INTERACTION PROCESS

When two or more people interact—that is, when they do something together—a rather complex set of processes take place. That interaction can be viewed in terms of three stages or modes. First a behavior by one member (A), verbal or otherwise, can be regarded as a *communication* from A to others (B, C, and so on). A series of such behaviors, by a set of interacting persons, can be regarded as the *communication process*. The *form* or structure of such a series of interactive behaviors or communications entails such factors as the communication channels and modalities used, the distribution of acts among persons and over time. That form or structure can be regarded as a *communication pattern*.

Each such behavior also can be considered with respect to its *content*. In principle, every interactive behavior can be regarded as having both a *task component* and an *interpersonal component*. The task oriented aspects of the participants' activities can be viewed as the *task or action process*, which results in a *task performance pattern*. The interpersonal oriented aspects of those activities can be viewed as the *attraction or acquaintance process*, which results in an *interpersonal relationship pattern*.

The third stage of the interaction process has to do with its *impact*. The three patterns resulting from the interaction (the communication pattern, the task performance pattern and the interpersonal relationship pattern) in turn have effects on one another and on the participants. Such effects constitute the *influence process*, which involves the outcomes or consequences of the interaction for the participants, for their relationships to one another, for their task performance and for their subsequent communications. These relations are shown in Figure 1-2 and listed in Table 1-1.

This trimodal perspective, along with the overall conceptual framework presented earlier, provides the basis for the organization of much of the rest of the book. Part II, (chapters 5 through 11) is devoted to the *task content* of group interaction; that is, the task performance process. In terms of the overall



A, B	Group members	Outcome	
Form		T/A, T/B	Effect of task performance on members
A/C, B/C	Communication process	I/A, I/B	Effect of interpersonal relations on members
Content		C/T, T/C	Effect of communication pattern and task performance on one another
A/T, B/T	Task Activity (Action Process)	C/I, I/C	Effect of communication pattern and interpersonal relations on one another.
A/I, B/I	Interpersonal Activity (Attraction Process)	T/I, I/T	Effect of task performance and interpersonal relations on one another

FIGURE 1-2 Interaction as a Three-Stage Process

conceptual framework, those chapters deal with "arrows" i, j, and p. (See Figure 1-1). By treating the task performance material earlier in sequence than the logic of the models would imply, I want to give that material special prominence. Much of the continued interest in groups over the years has focussed on groups as potential vehicles for improving task performance. Following those chapters on task performance, Part III deals with groups as systems for structuring interaction. Chapters 12 and 13 deal with the form or pattern of interaction and with the communication process. Chapters 14, 15 and 16 deal with the interpersonal content of interaction; that is, with the *acquaintance or attraction* process. Chapters 17 and 18 deal with outcomes or consequences of interaction; that is, with the *influence process.*

Before these presentations, though, the remaining chapters of Part I (chapters 2, 3, and 4) provide some background needed to make the detailed analyses of later parts of the book understandable. First, in chapter 2, there is a brief outline of past research on groups, of trends in that research, and of the role that theory (and its absence) has played in past group research. Chapter 3 presents a discussion of some general features of research methods in the social

TABLE 1-1 Interaction as a Three-Stage Process

PROCESS	*Communication Process*	*Action Process*	*Attraction Process*	*Influence Process*
ASPECTS OF INTERACTION	*Form of Interaction* Modalities Participation patterns Temporal patterns	*Content of Interaction* Task component: Generate Choose Negotiate Execute	Interpersonal component: Affect Control	*Consequences of Interaction* Impact of communication, task and interpersonal patterns on participants A & B and on each other
RESULT	*Communication Pattern*	*Task Performance Pattern*	*Interpersonal Relations Pattern*	*Pattern of Change in:* Participants; Communication; Task; Interpersonal relationships

and behavioral sciences and how those features both enable and constrain research on groups. Chapter 4 takes the concern with method one step further, providing a discussion of the various kinds of social units that have been used in the study of groups and classifies them in terms of their relations to the definition of groups given earlier in this chapter.

STRATEGIC LEVEL ISSUES: CHOOSING A SETTING FOR A STUDY

Research evidence, in social and behavioral sciences, always involves *somebody doing something, in some situation.* When we get such evidence, we can, therefore, "reference" it on three aspects or facets: Whose behavior is it about (which Actors)? What behaviors is it about (which Behaviors)? What situations is it about (which Contexts)?

When you gather a batch of research evidence, you are always trying to maximize three things:

1. The *generalizability* of the evidence *over populations* of actors (A).
2. The *precision* of measurement of the *behaviors* (and precision of control over extraneous facets or variables that are not being studied) (B).
3. The *realism* of the situation or *context* (in relation to the contexts to which you want your evidence to refer) (C).

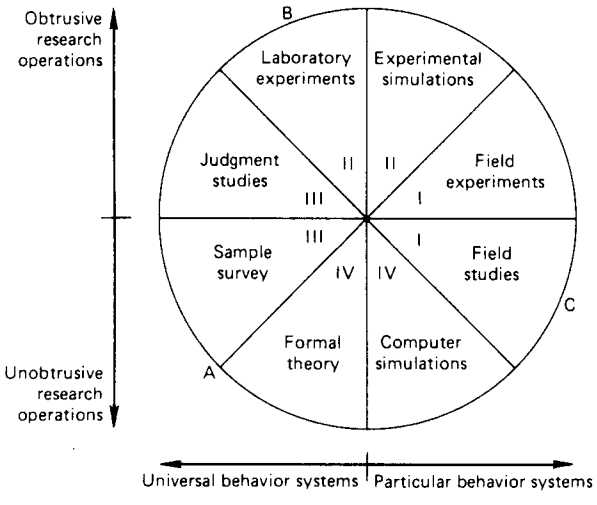
While you always want to maximize A, B, and C simultaneously, *you cannot.* This is one fundamental dilemma of research methods. The very things you can do to increase one of these reduces one or both of the other two. For example, the things you do to increase precision of measurement of behavior and control

of related variables (B) necessarily intrude upon the situation and reduce its "naturalness," or realism (that is, reduce C). Conversely, the things you can do to keep high realism of context (C) will reduce the generality of the populations to which your results can be applied (A) or the precision of the information you generate (B), or both.

The nature of this strategic dilemma is made clearer in Figure 3-1, which shows a set of eight alternative research strategies or settings in relation to one another. That figure shows where among the strategies each of three desired features—generalizability over populations (A), precision in control and measurement of behavior (B), and realism of context (C)—is at its maximum. It also shows, though, that strategies that maximize one of these are far from the maximum point for the other two. The spatial relations in Figure 3-1 emphasize the dilemma just discussed: the very things that help increase one of the desired features—A, B, and C—also reduce the other two. *It is not possible to maximize, simultaneously, all three.* Any one research strategy is limited in what it can do; and research done by any one strategy is flawed—although different strategies have different flaws.

The strategies listed in Figure 3-1 are in four pairs. Some are familiar ones. Field studies refer to efforts to make direct observations of "natural,"

FIGURE 3-1 Research Strategies



I. Settings in natural systems.
II. Contrived and created settings.
III. Behavior not setting dependent.
IV. No observation of behavior required.
A. Point of maximum concern with generality over actors.
B. Point of maximum concern with precision measurement of behavior.
C. Point of maximum concern with system character of context.

ongoing systems (in the present context that means existing groups), while intruding on and disturbing those systems as little as possible. Laboratory experiments are attempts to create the "essence" of some general class of systems (for the present case, groups) in a context in which the researcher can control all (or at least very many) of the extraneous features of the situation, in order to be able to maximize the essential features with precision. The two strategies in between refer to mixtures or compromises. Field experiments are field studies with one major intervention, the deliberate manipulation of some feature whose effects are to be studied. An experimental simulation is a laboratory study in which an effort is made to create a system that is like some class of naturally occurring systems (such as what are called mock juries later in this book), but which are artificial in that they are created by the researcher for study, and people perform in them for research purposes rather than for purposes stemming from their own lives.

Sample surveys are efforts to get information from a broad (and well devised) sample of actors, usually in the form of verbal responses to a relatively small set of questions. Judgment studies are efforts to get responses (usually from a very small and somewhat casually selected sample of "judges") about a systematically patterned and precisely calibrated set of stimuli. Surveys gain much generalizability over populations (A), but give up a lot in precision of measurement (B) to do so. Judgment studies have less generalizability over actors (A), but retain considerable precision of measurement (B). Both surveys and judgment studies try to deemphasize context—actually, to uncouple the behavior (judgment) from the context in which it is done. Thus, both are very low on realism of context (C).

The fourth pair of strategies are theoretical, not empirical. The term formal theory is used here to mean general theory. Such theories are high on generalizability over populations (A) because they attempt to be general; they are not very high on realism of context (C) because by being general they do not deal very concretely with any one context; and they are very low on precision of measurement of behavior (B), because, since they are theoretical rather than empiricial, they in fact involve no behaviors. The strategy called computer simulation refers to attempts to *model* a specific real life system or class of systems. Such effects are also theoretical rather than empirical; hence they are low on B because they do not involve behavior. In comparison to formal theories, computer simulations are higher in C, because they are system-specific; but they thereby lose in A, because they are limited to populations indigenous to that class of systems.

To sum up: Field studies gain realism (C) at the price of low generalizability (A) and lack of precision (B). Laboratory experiments maximize precision of measurement and control of variables (B), at the price of lack of realism (C) and low generalizability (A). Surveys have high generalizability (A) but get it by giving up much realism (C) and much precision (B). Formal theories get generalizability (A) by giving up some realism (C) and much precision (B). The other four strategies are combinations located in between those four just discussed; they have the intermediate gains and losses implied by their positions in the "strategy circle" of Figure 3-1.

Doing research is *not* to be regarded as trying to find the right strategy.

There is no right one. Indeed, they are all "wrong" in the sense that each is inherently limited, flawed. But they are all potentially useful. In considering any set of evidence, one should take into account what strategies were used in obtaining various parts of it, hence the strengths and limitations of that evidence at the strategic level.

DESIGN LEVEL ISSUES: WHAT WILL YOU COMPARE AND WHAT WILL YOU LEARN?

Any study needs a plan for what data will be gathered, how that data will be aggregated and partitioned, and what comparisons will be made within it. Such a study plan is often called a *research design*. As is evident from the preceding discussion, choice of one or another of the various strategies will limit the kinds of designs you can use. But there are also some general features of study designs, and it is those features that are to be discussed here.

Correlation versus Comparison

All research questions can be boiled down to variations of a few basic question forms. One is the *baserate* question: How often (at what rate, or what proportion of the time) does X occur? That is a purely descriptive matter, but is often a very crucial underpinning of other information. A second general form of question is the *relational* question: Are X and Y related? Do they occur together? That question has two major forms. In the correlational form, it is: Is there systematic *covariation* in the value (or amount or degree) of X and the value of Y? For example, does age covary with happiness? A high correlation between X and Y means that when X occurs at a high value, Y is also likely to occur at a high value; and when X is at a low value, Y is also likely to be at a low value. In the example from above, this would mean that older people were, by and large, happier than younger ones. The correlation between X and Y could equally well be high and *negative,* if high values of X went with low values of Y and vice versa. If that were the case for the example, then younger people would be, by and large, happier. There is little or no correlation between X and Y if knowing X doesn't help predict the value of Y. In the example, that would mean that older and younger people both vary in happiness, with some of each having high levels and some of each having less.

Given the example chosen here, of age and happiness, it certainly might occur to the reader that the highest level of happiness might occur, systematically, at some time other than in extreme old age or extreme youth. For example, happiness might increase up to age fifty, then decline. That would describe a nonlinear correlation (and, technically, a nonmonotonic one). There are statistical tools to test for such nonlinearity, although social scientists far too often do not use them when the evidence to be examined might well require them. But as the shape of the relation becomes more complicated—for example, if happiness decreased from young child to adolescent, then increased to age fifty, then decreased, but flattened out after sixty-five—our statistical tools become more cumbersome to use and many of them become less adequate to the task of assessing such complex forms of relation.

Much research in the social and behavioral sciences makes use of correlations, linear and nonlinear, that involve two, three, or more variables. Such a correlational approach requires being able to measure the presence or values of X, and of Y, for a series of "cases" that vary on X and on Y. It can tell you whether X and Y go together; but it *cannot* help you decide whether X is a cause of Y, or vice versa, or neither.

Another form of the relational question is the *comparison* or *difference* question. The difference question involves asking, essentially, whether Y is present (or at a high value) under conditions where X is present (or at a high value), and absent (or low) when X is absent (or low). For example: Do groups perform tasks better (Y) when members like each other (X) than when they do not (X' or "not-X")? You could approach this question in either of two ways. You could go around collecting measures of "liking" until you had found a bunch of groups high on it and another bunch of groups low on it (and perhaps a bunch at intermediate levels), and then compare their average performance scores. That would be, in effect, just a messy version of the correlational approach. The other approach would be to set up some groups with members who do like each other and set up some other groups whose members do not like each other; then to give both sets of groups some common tasks to perform; and then to see if the average task performance (Y) of the "high liking" groups (X) is higher than the average task performance of the "low liking" groups (X'). For the comparison to be most useful, you would need to make sure that the two sets of groups were the same, or comparable, on all the other factors that might affect task performance—such as difficulty of the task, availability of task materials, quality of working conditions, task-related abilities, experience and training of members, and the like. You might render the groups comparable on some of these factors by *controlling* them at a single *constant* value for all groups of both sets. For example, you probably would want to have all groups in both conditions do exactly the same tasks. For some other variables, such as intelligence or abilities of members, that you could not hold at a constant value for all cases, you might want to *match* the groups, on the average, between the two conditions. You might even want to manipulate a second or third variable in addition to group liking—perhaps group size, for example. But you can only manipulate, match, and control a limited number of variables in any one study. You have to do something else about all the rest of the rather large set of potentially relevant factors.

That something else is called *randomization,* or random assignment of cases to conditions. Randomization means use of a random assignment procedure to allocate cases (groups) to conditions (high liking versus low liking, or, if you were also manipulating a second variable such as size, high-liking-large-groups versus high-liking-small-groups versus low-liking-large-groups versus low-liking-small-groups), so that any given case is equally likely to be in any of the conditions.

To do what has been called a "true experiment" (see Campbell & Stanley, 1966), you *must* have randomization of cases to conditions. If you do, then you

strengthen the credibility of your information about high X going with high Y (and low X with low Y); and, since *you* caused X to be high in one set of groups and low in the other, it is at least plausible that X is a cause of Y. If instead of doing such a true experiment, you had just let things vary, measured X and Y, and correlated them, then X might have caused Y, or Y might have caused X, or both X and Y might have been caused by something else that you didn't pay attention to.

You can see that true experiments are potentially powerful techniques for *learning about causal relations among variables.* But, as in all aspects of research methodology, you buy this high power at a high price in two ways: (a) a reduction in the *scope* of your study, insofar as you hold variables constant, and insofar as you make your experimental variables (the X's) occur only at a couple of levels (high or low liking, or three-person versus six-person groups, for example) so that the results of that study will be thereby limited in generalizability; and (b) a reduction in realism of context, inasmuch as your activities (rather than "nature") have created the groups, designed the tasks, and elicited behavior that served your purposes, not the group members' purposes. It has been said that such an experiment lets you learn a lot about very little, whereas a correlational study may let you learn very little about a lot.

Forms of Validity

A study needs to have high validity in regard to four different types of validity questions (see Cook & Campbell, 1979). One, to which we have been attending in the preceding description of the "true experiment," is called *internal validity.* That has to do with the degree to which results let you infer about causal relations. A second form of validity has been called *statistical conclusion validity.* That refers to the confidence with which you can say that there is a *real difference* (in Y scores) between X cases and X' cases. Internal validity deals with a logical question, how to rule out alternative explanations (such as, that Y caused X or that both X and Y stemmed from unmeasured factor Z). But statistical conclusion validity is a statistical question, usually posed in some variation of the following form: How likely is it that the difference in average Y values, between the X batch of cases and the X' batch of cases, could have occurred by *chance?* If the probabilty of such a chance occurrence is less than 1 in 100 (written $p < .01$), or sometimes if it is less than 1 in 20 ($p < .05$), the researcher may conclude that results cannot be attributed only to chance. Usually, such results are said to be "significant" at the .01 or the .05 level.

When results are significant, the researcher may conclude that the hypothesis that only chance was operating *does not* account for the results; but he or she *may not* logically conclude that the hypothesis of interest ("X causes Y") *does* account for them. It is only if the researcher can eliminate most other plausible rival hypotheses (e.g., that Y causes X; that Y is caused by factor Z that also differed between groups, etc.), by the logic of his or her study design, that he or she can continue to entertain the X-causes-Y hypothesis as a plausible—but by no means certain—explanation for the results.

A study also needs to have clearly defined theoretical concepts and conceptual relations, and clearly specified mappings (or translations) of those concepts into empirical operations. This is called *construct validity.* Finally, the

researcher needs to have some basis for estimating how the obtained results would hold up if the hypothesis were tested on other populations of actors, using other measures of the same variables, in other situations and on other occasions in this same situation. Such estimates of generalizability refer to what is called *external validity.*

It will probably be apparent that the devices used to increase internal validity and statistical conclusion validity—the techniques used to gain precision—will threaten the external validity of that particular set of data. But the relation is not a symmetrical one. One should *not* leap to the conclusion that the converse is true. Things that aid external validity (e.g., large and varied samples) may either hinder or help internal validity or have no effect on it. Moreover, it is certainly *not* the case that things that *decrease* internal validity (e.g., not using randomization, or not using experimental manipulation) will somehow increase external validity. If you don't know what you found out in your study (i.e., if your study is low in internal validity or in statistical conclusion validity or in construct validity) then you cannot really determine whether or not, or how broadly, you can generalize it (i.e., what external validity it has)—but it doesn't matter anyhow. If you do know what you found out (i.e., if your study has high internal, statistical and construct validity), then it is important to try to determine how robust and general (i.e., how externally valid) those findings are likely to be.

There is much more to be said about study design, about difference versus correlation studies, about forms of validity, and about ways of dealing with plausible hypotheses that are alternatives to the hypothesis being tested—far more than can be said here. (For further reading on these questions, see Campbell & Stanley, 1966; Cook & Campbell, 1979; Runkel & McGrath, 1972). But perhaps what has been said serves to make several important points:

1. Results depend on methods.
2. All methods have limitations, hence any one set of results is limited, flawed.
3. It is not possible to maximize all desirable features of method in any one study; trade-offs and dilemmas are involved.
4. Each study—each set of results—must be interpreted in relation to other sets of evidence bearing on the same questions.

Some of these same points were made in regard to strategic issues, and some will apply, again, in the discussion of issues at the operational level that now follows.