

HANDBOOK OF RESEARCH METHODS IN CLINICAL PSYCHOLOGY

SECOND EDITION

Philip C. Kendall
Temple University

James N. Butcher
University of Minnesota

Grayson N. Holmbeck
Loyola University of Chicago

EDITORS



John Wiley & Sons, Inc.

New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

1999

Chapter 1

OVERVIEW OF RESEARCH DESIGN ISSUES IN CLINICAL PSYCHOLOGY

ALAN E. KAZDIN, PH.D.

In clinical psychology, there are special demands placed on researchers due to the broad range of content areas, populations, and settings that serve as the focus of research. Key content areas include the biological, psychological, social, and cultural influences on adjustment and maladjustment and assessment, diagnosis, treatment, and prevention of clinical dysfunction. These content areas encompass all age groups from infancy through late adulthood. Indeed, in some way, the field extends beyond these age limits by studying processes before birth (e.g., prenatal characteristics of mothers and families that may influence offspring) and after death (e.g., the impact of death on the surviving relatives, treatment of bereavement). Diverse populations are studied, including persons with special experiences (e.g., the homeless, prisoners of prior wars), with psychological or psychiatric impairment (e.g., children, adolescents, or adults with depression, anxiety, posttraumatic stress disorder, autism, schizophrenia), and with medical impairment and disease (e.g., cancer, Acquired Immune Deficiency Syndrome [AIDS], spinal cord injury, diabetes). Persons in contact with special populations are often studied (e.g., children of alcoholics, spouses of depressed patients, siblings of physically handicapped children) and hence also become special populations of interest. The different content areas, age groups, and populations lead to research in diverse settings (e.g., laboratories, day care centers, clinics, hospitals, prisons, schools, industry) and less well-specified "settings" (e.g., homeless families, runaway children).

The diversity and richness of clinical research have direct implications for research methods. Ideal methodological practices (e.g., random assignment, matching of cases to ensure group equivalence) are not always available. Also, special ethical considerations (e.g., withholding treatment) or characteristics of the subject matter (e.g., only small sample sizes available for a particular problem area) limit the researcher's options. The task of the scientist is to draw valid inferences from the situation and to use methodology, design, and statistics toward that end. In clinical psychology, the methodological, design, and data analytic strategies are more diverse than in basic research. The investigator often must engage in innovative strategies and methodological solutions to draw inferences about key phenomena. Clinical research is not in any way soft science; indeed, the processes involved in clinical research reflect science at its best precisely

because of the thinking and methodological ingenuity required that forces nature to reveal its secrets.

Multiple methods of study are required to meet the varied conditions in which clinical psychologists work and the special challenges in drawing valid scientific inferences from often complex situations. This chapter discusses key questions and concepts that serve as the impetus of research and the role that methodology plays in addressing these questions. Different types of designs and the conditions under which investigations can be conducted are highlighted. Special methodological challenges in research related to sampling, measurement selection, evaluation of constructs, and statistical evaluation are also presented.

GOALS OF RESEARCH

Key Questions and Concepts That Guide Research

The goals of research are to address substantive questions about the phenomena of interest. Research methods are inextricably related to these questions insofar as selection of the design, sample, measures, and methods of data analyses materially influence the answers. It is not possible to enumerate all of the questions that guide clinical research and their implications for research design. Yet, several key concepts often underlie research in clinical psychology. These concepts convey different levels of understanding phenomena and raise different design issues and strategies.

Table 1.1 presents key questions and concepts that pertain to the relations among variables of interest. At the most basic level of understanding, clinical research often focuses on identifying whether variables are *correlated*: Participants are tested on several measures at a given point in time to relate such variables as symptoms, cognitive processes, personality, stress, social support, family functioning, recall of past experiences (e.g., abuse, attachment), and physical health. The presence of a relation and the magnitude of the relations among such variables can be important for elaborating the nature of a problem and for testing or developing theories about its onset and course.

Establishing that two (or more) events are correlated is a preliminary level of understanding and, of course, does not establish their temporal relation to each other. The notion of *risk factor* represents a deeper level of understanding because the time line is established, namely, that one event or experience (e.g., abuse, prenatal exposure to drugs or alcohol, enriched educational or parenting experiences) is correlated with a later characteristic (e.g., dysfunction at school, marital bliss).¹ The early experience may not be a cause, but we know that the experience, for whatever reason, increases the likelihood of the outcome.

Demonstrating *cause* means, of course, that we have established that the relation is not merely a temporal ordering of events but rather that there is some direct influence of

¹ The term risk factor has emerged from epidemiology and public health, where the focus has emphasized adverse antecedents (e.g., exposure to toxins, poor nutrition) and untoward outcomes (e.g., morbidity, mortality). The term is used here more generically, as it has come to be used within clinical research, to denote antecedent factors that increase the likelihood of an outcome (see Chapter 18). The valence, desirability, or focus of either the antecedent or outcome is not critical.

Table 1.1. Sample questions and concepts that serve as the impetus for research

1. What is the relationship between (among) the variables of interest?
 - Correlate*—The two (or more) variables are associated at a given point in time in which there is no direct evidence that one variable precedes the other.
 - Risk factor*—A characteristic that is an antecedent to and increases the likelihood of an outcome of interest. A “correlate” in which the time sequence is established.
 - Cause*—One variable influences, either directly or through other variables, the appearance of the outcome. Changing one variable is shown to lead to a change in another variable (outcome).

2. What factors influence the relationship between variables, that is, the direction or magnitude of the relation?
 - Moderator*—A variable that influences the relationship of two variables of interest. The relationship between the variables (A and B) changes or is different as a function of some other variable (e.g., sex, age, ethnicity).
 - Protective factor*—A characteristic that reduces the likelihood of an outcome of interest among individuals identified as at risk. The relation between the risk factor and an outcome is altered by the presence of some other characteristic. Protective factors are moderators but worth distinguishing because of their interest in clinical research.

3. How does the phenomenon work; that is, through what relation or mechanism or process does A lead to B?
 - Mediator*—The process, mechanism, or means through which a variable produces a particular outcome. Beyond knowing that A may cause B, the mechanism elaborates precisely what happens (e.g., psychologically, biologically) that explains how B results.

4. Can we control or alter the outcome of interest?
 - Intervention*—Is there something we can do to decrease the likelihood that an undesired outcome will occur (prevention) or decrease or eliminate an undesired outcome that has already occurred (treatment)? Although these questions are usually framed as focusing on some undesirable outcome, they often promote positive, prosocial outcomes to achieve their end.

Note: These terms, the relations they reflect, and diverse research strategies their evaluation entails are detailed elsewhere (see Baron & Kenny, 1986; Holmbeck, 1997; Kazdin, Kraemer, Kessler, Kupfer, & Offord, 1997; Kraemer et al., 1997).

one event on the other. As a rule, several conditions usually need to be satisfied to infer cause, including evidence that the event or agent precedes the outcome, there is a strong association between the event and outcome, the relation is not due to confounding influences, the effects are consistent (replicable) across samples from the same population, and there is a plausible explanation of the mechanisms and processes through which the antecedent and outcome are related (see Haynes, 1992; Hill, 1965; Kenny, 1979; Schlesselman, 1982). No single study is likely to be sufficient to satisfy all of the criteria. From multiple studies, the causal role of the event becomes more plausible and parsimonious, as various competing interpretations are either confirmed or rejected. There are often many causes of a phenomenon of interest so that demonstrating that some

theory based research:
 predictive → causality
 intervention → efficacy

event is a cause does not mean that the event is *the only* cause. Demonstration of causality is a high level of understanding and a goal of research.

Knowing *the cause* or *a cause* of a phenomenon can reflect different levels of understanding because precision of the knowledge may vary. In the usual case in clinical research, cause refers to knowing how to change a phenomenon. For example, a great deal of research focuses on interventions (treatment, prevention, and educational programs) to reduce clinical dysfunction, to prevent the onset of dysfunction, and to promote well-being or adaptive functioning. These studies focus on causal relations, that is, making a change at the level of the individual, school, or community, for example, will lead to change in the outcome(s) of interest. We may know how to produce change even if we are not sure of the mechanisms involved. For example, successful prevention programs (e.g., for young children at risk for school failure and behavior problems) often rely on multimodal interventions, that is, programs that entail several techniques (e.g., counseling and medical care for the parents, special day care for the child, changes at home for the child such as reading, and others) (see Meisels & Shonkoff, 1990). If such an intervention produces reliable change, we can say that a causal relation was demonstrated, even though we may not know how change was produced, precisely what facet of the intervention produced change, or what intervening steps (e.g., affect, cognition, behavior) led to the change in the target domain.

Research often focuses on *moderators*, which are variables that influence the direction, magnitude, and nature of the relation (Baron & Kenny, 1986; Holmbeck, 1997). For example, early maturation among adolescents is a risk factor for later psychopathology (internalizing and externalizing disorders and suicide attempt) for girls but not for boys (e.g., Graber, Lewinsohn, Seeley, & Brooks-Gunn, 1997). That is, gender moderates the risk-factor outcome relation between early maturation and psychopathology. (For boys, the risk stems from late rather than early maturation.) Identification of moderators is an important advance in understanding because knowing what other factors influence the relation between variables often prompts theory and research to explain why the moderator has an influence. Also, the moderating influence may suggest different causal paths for one group rather than another group.

The study of *protective factors* is an excellent example of the study of moderators. The term "protective factor" has been used in different ways. The first way is to refer to a characteristic in the population that reduces the likelihood of the outcome. In this sense, a protective factor is similar (conceptually) to a risk factor and denotes an antecedent that is associated with a particular outcome. The second and more common use is in the context of a group identified as at risk or high risk for a particular outcome (Rutter, 1987). Among individuals identified as at risk, many may not show the undesired outcome. Characteristics that are associated with a decrease in risk for this group are referred to as protective factors. Stated another way, the relation between two variables is influenced by the presence of another variable, that is, a protective factor. For example, Werner and Smith (1992) found that early influences in childhood, such as parental conflict, alcohol abuse of a parent, and below normal intellectual functioning, placed youths at risk for delinquency in adolescence. The at-risk group that did not become delinquent showed a number of characteristics that seemed to "protect" them from deleterious influences. Those who did not evince delinquency by adolescence were more likely to be firstborn, to be perceived by their mothers as affectionate, to show higher self-esteem

and locus of control, and to have alternative caretakers in the family (than the parents) and a supportive same-sex model who played an important role in their development. Stated another way, the relations between several antecedents and delinquency were moderated by these other characteristics.

The focus on *mediators* or *mechanisms* represents a deeper level of understanding beyond the relations noted above, because this means we know *how* the problem unfolds, through what *processes*, and the *ways* in which one variable leads to another (see Holmbeck, 1997). As an example, an association with deviant peers during adolescence is related to subsequent antisocial and delinquent behavior (see Rutter & Giller, 1983; Stoff, Breiling, & Maser, 1997). What mediates this relation or through what mechanisms or processes does one characteristic (association with deviant peers) lead to the other (delinquent behavior)? No doubt there are multiple mechanisms involved; however, evidence suggests that in day-to-day interactions, deviant peers directly reinforce deviant behavior (Dishion & Patterson, 1997). For example, in a delinquent peer group, when the conversation of an individual focuses on rule breaking, there is a positive reaction (e.g., laughter) which serves to reinforce and extend discussion of deviance. In addition, deviant peers tend not to react positively to more normative talk. This evidence suggests that there is a deviance training process in the interactions of delinquent peers and this process leads to escalation of rule-breaking acts. Multiple influences are likely to contribute to and explain the connection between association with deviant peers and delinquent behavior.

Consider another example where the mediator refers to a circumscribed process that explains more precisely the relation of an antecedent and an outcome. Research on the relation of cigarette smoking and lung cancer has spanned the range of concepts included in Table 1.1. That is, both human and animal studies have shown that cigarette smoking is a correlate of, a risk factor for, and a cause of lung cancer. This still leaves open the question of *how* cigarette smoking leads to the disease, that is, through what mechanisms or intervening processes. Recent research has elaborated the mechanism involved in the causal relation. A chemical (benzo[*a*]pyrene) found in cigarette smoke induces genetic mutation (at specific regions of the gene's DNA) that is identical to the damage evident in lung cancer cells (Denissenko, Pao, Tang, & Pfeifer, 1996). This finding is considered to convey precisely how cigarette smoking leads to cancer at the molecular level. Thus, beyond the demonstration of a causal relation, a fine-grained analysis of mechanisms is important as well. Knowing the mediator of a relation between variables does not require knowing the biological substrates. The mechanism or process through which two variables are related may involve all sorts of psychological constructs.

Our understanding, of course, is optimal when we know all of the above in great detail about a particular phenomenon of interest, that is, correlates, risk factors, moderators, causes, and mechanisms or processes through which outcomes are produced. There are few areas that are so well established. An example that encompasses many of the key concepts is the relation of parenting practices and the development of aggressive and antisocial behavior in children. Research has shown that inept parental discipline practices foster aggressive and antisocial behavior at home (see Dishion, Patterson, & Kavanagh, 1992; Patterson, Reid, & Dishion, 1992). These discipline practices include parental attention to deviant behavior, interactions in which increasingly aggressive child behavior is reinforced, inattention to prosocial behavior,

Handwritten notes in the left margin:

knowing how or how
 (e.g., laughter) which serves to reinforce and extend discussion of deviance
 important
 "mechanism" in this specific
 is

causal-
 under-
 ical re-
 eat deal
 grams)
 promote
 that is,
 le, will
 ge even
 on pro-
 s) often
 es (e.g.,
 nges at
 If such
 demon-
 at facet
 gnition,
 e direc-
 1997).
 10path-
 but not
 gender
 10path-
 ication
 t other
 arch to
 ay sug-
 rs. The
 er to a
 sense,
 cedent
 in the
 Rutter,
 d out-
 re re-
 bles is
 mple,
 l con-
 ouths
 quent
 us in-
 to be
 steem

coercive punishment, poor child supervision, and failure to set limits. Initial studies showed that inept discipline practices were correlated with child antisocial behavior, as measured by parent, teacher, and peer ratings, both in community and clinic samples of boys and girls (Dishion et al., 1992; Forgatch, 1991). Further studies helped to establish the time line, namely that inept child-rearing practices preceded child antisocial behavior.

The causal role of parenting practices and antisocial behavior was suggested in *randomized controlled clinical trials* in which parents were randomly assigned to various treatment and control conditions.² Those conditions involving direct alteration of parent discipline practices, compared to treatment and control conditions without this focus, led to decreases in child antisocial behavior (Dishion & Andrews, 1995; Dishion et al., 1992). Several controlled studies have shown similar results, namely, changes in parenting skills lead to changes in child behavior (see Kazdin, 1997). From multiple studies, we know that adverse (especially harsh) parenting practices lead to aggressive and antisocial child behavior, even though this does not imply such practices are *the* cause of aggressive behavior, the *only* cause of these behaviors, or even a *necessary or sufficient* cause of the behaviors. Although less well studied, research has identified some of the factors that influence the effectiveness of parent training, including age of the child, severity of child antisocial behavior, socioeconomic disadvantage, stress, and dysfunction in the family (e.g., Kazdin, 1995; Webster-Stratton, 1996). There remains a great deal to learn. The precise mechanisms are not entirely clear. That is, changes in child behavior might result through a number of processes including the direct operation of reinforcement and punishment contingencies and/or improved parental responsiveness to their children (see Wahler & Meginnis, 1997). Yet, from the progression of research we know a great deal about parenting practices, their impact, and what we can do about them.

Importance of Theory

The progression of research from correlation to cause and hence from description to explanation, as described to this point, may inadvertently imply a crass empiricism, that is, one merely tests different types of relations among variables to see what role they play, if any. Underlying the concepts that guide research highlighted in Table 1.1 is the investigator's theory that focuses the research idea. Theory, broadly defined, refers to a conceptualization of the phenomenon of interest. The conceptualization may encompass views about the nature, antecedents, causes, correlates, and consequences of a particular characteristic or aspect of functioning. Also, the theory may specify the relations of various constructs to each other. There are different levels of theory in terms of their breadth, generality, and scope of affect, cognition, and behavior they attempt to encompass. In contemporary work, theories tend to be relatively circumscribed in an effort to explain the relation between selected characteristics and a disorder (e.g., hopelessness and helplessness in relation to depression) and how these characteristics lead to other features of dysfunction. Theories that attempt to explain large segments of functioning

² A randomized, controlled clinical trial refers to an outcome study in which clients with a particular problem are randomly assigned to various treatment and control conditions.

(e.g., all of clinical depression) are likely to be very general and are very unlikely to adequately account for varied patterns. The focus on segments of a problem, subtypes, and different developmental periods is more common. Once there is support for a theory within a narrow domain, it is then likely to be extended.

A goal is to understand human functioning and to accomplish that we do not merely accumulate facts or empirical findings. Rather, or in addition, we strive to relate these findings to each other and to other phenomena in a cohesive way. For example, an investigator may demonstrate that there are gender differences regarding a particular disorder, personality characteristic, or cognitive style. However, such differences are not necessarily inherently interesting. A theoretical understanding would pose how this difference develops and what implications the difference may have for understanding biological or psychosocial development. Inevitably, there will be many exceptions to the theory and these will require posing moderators and new lines of work. From the standpoint of research, theoretical explanations guide further studies and the data generated by the studies require emendations of the theory. This activity is important because theory moves us to implications beyond the confines of the specific relations that have been demonstrated and the restricted conditions in which these relations may have been demonstrated.

THE ROLE OF RESEARCH METHODS

Drawing Valid Inferences

The purposes of empirical research are to demonstrate relations among variables, as highlighted in Table 1.1, and to test theoretical propositions about those relations. Research methods play a central role because most relations of interest cannot be readily dissected when viewed in their full complexity as they appear in nature. Research design and statistical evaluation help simplify the situation in which the influence of many variables, often operating simultaneously, can be separated from the variable(s) of interest to the investigator. Without such simplification and isolation of variables, many if not an unlimited number of interpretations could explain a particular phenomenon. Among these interpretations is the possibility that in fact there is no relation between the variables of interest, that there is a relation but it is obscured by other influences, and that there is no relation but one appears to exist in light of other influences that are operating. Research methods make a special contribution by helping to rule out or make implausible several different factors that might explain the relation that the investigator wishes to evaluate. An experiment does not necessarily rule out all possible explanations; the extent to which it is successful in ruling out alternative explanations is a matter of degree. From a methodological standpoint, a well-designed experiment is one in which competing hypotheses that might explain the results are made relatively implausible or ruled out.

The purpose of research is to reach well-founded (i.e., valid) conclusions about the effects of a given intervention, experimental manipulation, or condition. Four types of experimental validity have been delineated as a useful way to identify major design issues that can emerge and interfere with drawing valid inferences. These consist of internal, external, construct, and statistical conclusion validity (Cook & Campbell,

1979). Table 1.2 lists each type of validity, the methodological question each is designed to address, and sources of influence or threats to validity that can interfere with conclusions the investigator wishes to reach. Research design, methodological practices, and statistical evaluation assist in rendering implausible or less plausible sources of threat that might otherwise explain the results.

Brief Illustrations

The threats to validity and how they interface with substantive questions can be illustrated by considering a few basic experimental arrangements. Consider that an investigator is interested in establishing the effectiveness of a new treatment—in fact, a

Table 1.2. Types of experimental validity and the issues they raise

Internal validity—To what extent can the intervention, rather than extraneous influences, be considered to account for the results, changes, or group differences?

Validity is threatened to the extent that the results can be explained by history, maturation, repeated testing, changes in the measure in some way, regression toward the mean, selection differences/biases among groups, attrition, special influences affecting one group but not another, diffusion of treatment, and special treatment or reactions of control participants.

External validity—To what extent can the results be generalized or extended to people, settings, times, measures, and characteristics other than those in this particular experimental arrangement?

Validity is threatened to the extent that the results are restricted because of special characteristics of the sample, stimulus conditions of the experiment or setting, reactivity (awareness of participating in a study or of the measurement procedures), the way in which the intervention was presented (e.g., in the context of multiple treatments), novelty effects, and test sensitization (the results due in part to the measures that influenced receptivity to the intervention).

Construct validity—Given that the intervention was responsible for change, what specific aspect(s) was the causal agent; that is, what is the conceptual basis (construct) underlying the effect?

Validity is threatened to the extent that group differences could be explained by differential attention and contact with the subjects, experimenter expectancies, cues of the experimental situation, single operations (e.g., one therapist), and narrow stimulus sampling that cannot be separated from the intervention.

Statistical conclusion validity—To what extent is a relation shown, demonstrated, or evident, and how well can the investigation detect effects if they exist?

Validity is threatened by low statistical power, variability in the procedures, subject heterogeneity, unreliability of the measures, restrictive or lenient error rates due to multiple tests.

Note: Additional discussion and illustrations of these threats and the issues they raise are provided elsewhere (Kazdin, 1998).

