4

Causal Comparative and Correlational Research

The school psychologist announced to the faculty that the school would participate in a research study to compare the effect of a new strategy for improving students' self-concepts. To control for the differential selection effects, all the names of the students in the school would be put in a hat and then randomly assigned to the high- and low-self-concept groups. Of course, this example is absurd. You can't assign people to different self-concept levels at random. Many characteristics of individuals are not manipulable—for example, disabilities, gender, ethnicity, age, cognitive abilities, and personality traits, such as aggression or anxiety. Causal comparative and correlational research strategies represent two approaches that are appropriate for studying such nonmanipulable variables.

Types of Variables

A variety of types of variables are appropriate for causal comparative and correlational research:

1. Inherent characteristics (organismic)—for example, gender, ethnicity, age, disability, socioeconomic class, ability, personality traits.
2. Characteristics that should not be manipulated for ethical reasons—for example, illegal drug use, cigarette smoking, alcohol consumption.
3. Characteristics that could be manipulated but that are not—for example, school placement, social promotion to the next grade, participation in psychotherapy.

When studying such characteristics, a researcher can use a causal comparative or a correlational approach. These types of research are quite common in education and psychology because of the frequency of comparisons of persons with different characteristics (such as gender, race, and disabilities). Although both approaches explore cause-and-effect relationships between variables, neither involves the experimental manipulation of treatment variables, and therefore, the results cannot be used as proof of a cause-and-effect relationship.

For example, Mexican American and European American college students were compared on the basis of their willingness to use, and their perceptions of the effectiveness of, four different alcoholism treatment programs (Atkinson, Abreu, Ortiz-Bush, & Brewer, 1994). The researchers found that Mexican Americans and European Americans reported being equally willing to use alcoholism treatment programs; however, the Mexican Americans perceived all four treatment programs to be more effective than did European American students. The authors did not conclude that Mexican American ethnicity causes a higher degree of perceived effectiveness. Rather, they speculated that cultural differences between the two groups might result in different ratings. To explore this hypothesis, they administered an acculturation scale on which low scores represented identification with Latinos and Hispanics and high scores represented European American identification for language use, media, and ethnic social relations. Their hypothesis was supported because the Mexican Americans with high acculturation scores rated the treatment effectiveness in a manner similar to the ratings of the European Americans. Atkinson et al. note that these results raise questions as to why current findings suggest that Mexican Americans tend to underuse, or prematurely terminate from, alcoholism treatment programs. They suggest alternative explanations, such as the use of college students (rather than recovering alcoholics) in their sample; counselor characteristics, such as ethnicity, cross-cultural competence, and credibility; and economic and physical inaccessibility of services. Thus, the authors did not assume causality but explored alternative hypothesis for explaining the study's results.

Although both causal comparative and correlational research are used to study phenomenon involving the inherent characteristics of participants, there is an important difference between the two approaches: Causal comparative research focuses on making group comparisons (e.g., comparing academic achievement in groups with high vs. low self-concepts). Although correlational research can also be used to make group comparisons, its main focus is on providing an estimate of the magnitude of the relationship between two variables (e.g., examining the relationship between the level of self-concept and academic achievement). The difference in focus of the two types of studies leads to a difference in the kinds of conclusions that can be drawn. In the causal comparative study, the researcher might conclude that a group of students with high self-concepts differed significantly on academic achievement compared with a group with low self-concepts. In a correlational study, a researcher might conclude that there is a strong, positive relationship between self-concept and academic achievement.

Inherent Characteristics: Challenging Issues

By their very nature of comparing individuals who differ based on inherent characteristics, such as ethnicity, gender, socioeconomic class, or disabling conditions, these approaches to research have serious implications for researchers in terms of how the research questions are framed and the basis that is used for group definition.

Focusing on Group Differences

As Campbell (1988, 1989), Fine and Gordon (1989), and others (Shakeshaft, Campbell, & Karp, 1992) have pointed out, the focus of much gender-related research has been on gender differences. Campbell (1988) notes, "Even the sound of 'sex similarities' sounds new and strange" (p. 5). These researchers point out that focusing on sex differences obscures the many areas in which males and females overlap.

Researchers have noted that gender differences in academic abilities, even in the areas of verbal and mathematics skills, have narrowed between males and females (Hyde, 1990). Lips (1993) used this basis of "no differences" in math abilities and aptitude to explore why women are underrepresented in mathematics and the physical and engineering sciences. Now, you might ask, how did the myth that males are superior in mathematics attain such stature as a scientific fact? Campbell (1988) proposes a number of factors related to research design and reporting that might help explain the focus and seeming credibility of differences between the sexes.

Campbell (1988) attributes the focus on gender differences to an artifact associated with the origins of research designs in the social sciences rooted in an agricultural model designed to investigate differences in effect of treatments on variables such as crop size and quality. Although this is an appropriate approach for research in agriculture, it fails to accommodate the needs of a researcher who is exploring the complexity of human beings in that it does not allow an examination of similarities *and* differences. Thus, the model is incomplete and inadequate for gaining knowledge about males and females or about people from minority groups and Whites. The second artifact of the research process that leads to an overemphasis on differences arises from the practice of journals in education and psychology to publish studies in which *statistically significant differences* are found (a term defined in Chapter 1 and discussed more fully in Chapter 12). Campbell (1988) points out that finding differences

> has been what counts in terms of publication, dissemination and ultimately research survival. Studies not finding significant differences are FOUR times less apt to be finished. Even when they are finished, they are less likely to be published than studies in which significant differences are found. (p. 3)

Fine and Gordon (1992) present a different basis for criticism of gender difference research. They write that "this almost exclusive construction of gender-as-difference functions inside psychology as a political and scientific diversion away from questions of power, social context, meaning, and braided subjectivities" (p. 8). They suggest that what is needed is a new language, because the issues are less about sex as biology or even gender as social construction and more about the politics of sex-gender relations that can transform oppressive social arrangements. Although sex and gender may feel biological or psychological, the more important focus is on the political implications.

Group Identification

A second area of challenge for the researcher in the causal comparative and correlational approaches to research is the definition of who belongs in which group. This is more a problem with race, class, and disability-related research at present than with gender-related research. Definitions of race and ethnicity somewhat parallel those of sex and gender; that is, one is considered to have a biological basis and the other a socially constructed basis. However, Stanfield (1993a) identifies the difficulties in answering questions such as, What is a White person? or What is a Black person? as stemming from the "extensiveness of ethnic mixing that has occurred in the United States in reciprocal acculturation and cultural assimilation processes, if not in miscegenation experiences" (p. 21).

Problems arise in terms of which basis for categorization to use. Stanfield (1993a) recognizes that most classifications of racial identity are based on skin color and other phenotypic characteristic: For example, a person who has dark skin, woolly hair, and a broad nose is readily identified as an African American. The problems for a researcher arise when the skin color of the African American person becomes lighter so that he or she might "pass" for European American or with people of mixed-race descent who do not readily identify with any of the standard racial categories.

Qualitative researchers tend to construct the meaning of a person's racial identity based on the respondents' self-perceptions of race and ethnicity and their influence on one's life experiences. Quantitative researchers, on the other hand, have tended to rely on statistical categories derived from government documents and survey coding. Stanfield (1993a), however, warns both groups of researchers that they have no way of knowing "whether an informant's expressed racial identification is a response to the objectified categorization derived from learning experiences in a race-saturated society or merely a subjective admission" (p. 18).

Bias can result when the method of determining racial or ethnic identity does not adequately represent the complexities of the situation. For example, if research is done on Latino populations and the respondents are selected based on a Spanish surname, children of Latino fathers who use the names of their fathers would be included but not the children of Latino mothers (Shakeshaft et al., 1992). The practice of lumping together biological and social definitions of race under a common racial label results in a biased sample. This is the situation that results when children of a Black and a White parent are identified as African American.

The use of a socially constructed, self-identification of race or ethnicity by authors such as Jensen (1969) and Herrnstein and Murray (1994) is problematic in terms of the types of genetically based interpretations that they offer. As Campbell (1989) notes, "When social definitions of race are used, *no* conclusions about genetic or biological differences can be made" (p. 11).

Stanfield (1993a) raises some important questions for researchers who choose to work in the causal comparative or correlational mode with respect to race and ethnicity:

- How do we conceptualize identity issues in race and ethnicity research that go beyond reified, simplistic stereotyping?
- How do we use official data sources with care in exploring racial identity questions, realizing the problematics of aggregate data and ill-defined circumstances of self-reporting versus actual self-identity?
- If we have to categorize people to understand who they are and how they define themselves, how do we do so in this area of research more in terms of self-definitions than in terms of what popular cultural folk wisdom dictates?
- How do we incorporate understanding in research designs regarding the interactional aspects of identity formation in dominant and subordinate populations that would make such considerations much more sociological? (p. 24).

The U.S. Office of Management and Budget (OMB) has recognized the complexity of multiracial groups and other individuals who do not neatly fit into existing categories for race on federal forms (Skrzycki, 1994). The OMB is currently struggling with how to collect and code information for people who might have one White and one African American parent or who might be from South America but are not Hispanic. The options they are considering include the following:

1. Add a new "multiracial" category
2. Add an "other" category
3. Provide an open-ended question to probe for information on race and ethnicity
4. Add categories for Native Americans, Native Hawaiians, and Middle Easterners

Other federal agencies are also struggling with this issue. For example, the U.S. Bureau of the Census and the Centers for Disease Control are considering changing the racial categories they use by allowing people to identify themselves as mixed race or by adding new categories for minorities, such as Cambodians or Arab Americans (Wheeler, 1995).

The whole debate over racial or ethnic identity is further complicated by the conflict between the reality of social oppression based on such phenotypic characteristics as skin color and the realization that no single gene can be used to define a race. The American Anthropological Association passed a resolution saying that "differentiating species into biologically defined 'races' has proven meaningless and unscientific" (cited in Wheeler, 1995, p. A9). Anthropologists have replaced the concept of race with a focus on how people identify themselves, by geographic origins or by other means.

Despite all the problems associated with categorizing people according to race or ethnicity, disabling conditions, and (sometimes) gender, researchers need to be aware of the benefits that have accrued from cross-group comparisons. Causal comparative and correlational research have been used to document oppression based on skin color and other phenotypic characteristics. Discontinuing such research based on the rationale that our understanding of race, gender, and disability is limited needs to be weighed against the benefit associated with revealing inequities in resources and outcomes in education, psychology, and the broader society.

Fallacy of Homogeneity

Campbell (1989) and Stanfield (1993a) both discuss the fallacy of homogeneity—that is, assuming similarities within racial and ethnic groups on other characteristics, such as socioeconomic class. Much of the research done in cross-race comparisons ignores the generally higher socioeconomic class associated with people of European American descent. Teasing out the effects of race and poverty is a complex and difficult task (if not impossible).

Problems similar to those associated with race can be found in Mertens and McLaughlin's (1995) discussion of the identification of persons with disabilities (p. 60).

Post Hoc Fallacy

Problems with the identification of group differences in causal comparative and correlational research occur with such regularity that researchers have developed a specific name for the inappropriate attribution of causation in such studies—that is, the post hoc fallacy (Campbell, 1989). The types of studies discussed in this chapter (e.g., comparisons of males and females, European Americans and African Americans) are particularly susceptible to the post hoc fallacy, and therefore, competing explanations for the results should be carefully examined. Lips's (1993) study is one example of causal comparative research that examined numerous competing explanations for differences in male and female participation in mathematics and engineering careers. Her study is used in the following section to explain strategies for conducting causal comparative research (see Sample Study 4.1 for a summary of Lips's study).

Causal Comparative Research

The steps for conducting causal comparative research are similar to those outlined in Chapter 1 for any research undertaking:

1. Identify a research problem.
2. Select a defined group and a comparison group.
3. Collect data on relevant independent and dependent variables and on relevant background characteristics.
4. Analyze and interpret the data.

SUMMARY STUDY 4.1 *Summary of a Causal Comparative Study*

Research problem: Compared with men, women are underrepresented in mathematics and science careers.

Research question: For people who have strong positive feelings about math and science, what are the gender differences in self-rated likelihood of pursuing a career in physical sciences or engineering?

Method: Male and female college students were asked to respond to yes-or-no questions on a computer screen related to their feelings about their own inclinations toward math and science and their interest in a career in the physical sciences or engineering.

Participants: Ninety-seven college students (55 females, 42 males) participated in the study.

Instruments and procedures: The researcher used computer-programmed and timed assessments. The first part involved yes-or-no responses to short self-descriptive phrases, including 13 related to mathematics and science affirmations, such as "good with numbers." The career portion of the assessment items related to a large number of careers, including some specific to mathematical, physical, and engineering sciences were presented. Again, a yes-or-no format was used to determine the likelihood that the respondent would pursue this career.

Results: For physical and engineering science careers, the pattern for women and men diverged sharply at high-interest levels. For men, a strong general endorsement of science and mathematics was significantly and positively predictive of an increasingly strong likelihood of interest in a career in the physical sciences and engineering. For women, there was little or no relationship between self-described interest in and liking for science and mathematics and self-reported likelihood of pursuing a career in the physical and engineering sciences.

Discussion: Lips suggests that future research is needed to explore factors that may be necessary to heighten women's chances of choosing scientific and engineering careers, such as psychological support at some critical juncture, a minimum number of female faculty, or a visible number of female student peers in departments such as science, mathematics, and engineering.

SOURCE: Based on Lips (1993).

First, a research problem is identified. Lips (1993) notes that women are underrepresented in mathematics and science careers. She speculates on possible causes for this discrepancy:

1. Differences in ability and performance might explain the discrepancy. But Lips rejected this because of the small to modest differences supported by previous research. Even small differences by gender are suspect because of the gender-sensitive nature of the measuring instruments, such as the SAT-M. Also, comparisons of group means resulted in biased estimates of true differences because the males' scores were "pulled up" by a relatively small group of high-scoring outlier men.
2. Social forces such as low peer and faculty support might lower the extent to which girls and women value achievement in math and science.
3. Females might feel personally alienated from the subject matter, methods of science, or both, thus leading to the development of a self-view that does not include interest in math or science. If this is the reason for underrepresentation in these fields, Lips reasons that redress could be made by establishing policies and procedures focused on awakening (or reawakening) women's self-perception of their abilities and interests in math and science.

Lips decided to pursue this problem in terms of the gender differences in the relationship between self-affirmed inclination toward mathematics and science and likelihood of choosing careers in the physical sciences and engineering.

The *second step* is to select a defined group and a comparison group. The researcher has a number of options for creating the groups that yield more or less control over the differential selection threat to internal validity. Because causal comparative research compares the performance of two (or more) intact groups, the threat of differential selection must be addressed. If the two groups differ significantly on characteristics other than the explanatory variable (e.g., gender), those other (extraneous) characteristics might explain the difference between the groups.

Some strategies that researchers can use to control the threat of differential selection include the following:

1. Matching on particular characteristics of relevance (discussed further in Chapter 10)
2. Using a statistical technique such as analysis of covariance to control for preexisting differences (discussed further in Chapter 12)
3. Eliminating subjects with specific characteristics (e.g., those with multiple disabilities)
4. Analysis of subgroups

The creation of homogeneity by elimination of people with specific characteristics comes at a cost in restricting the generalizability of the findings to that "homogeneous" population.

For her study, Lips (1993) chose 97 college-level students (55 females, 45 males). In causal comparative research, the researcher will often form subgroups based on other variables of interest (e.g., ethnicity) to better understand the phenomenon. Lips decided to explore inclinations for careers in the physical sciences and engineering for students with different self-affirmed inclinations toward math and science. She created five subgroups based on their self-affirmed math-science inclination: low, low-average, average, high-average, and high. Thus, she was able to determine if the self-reported likelihood of a career in physical science and engineering varied across levels of inclination toward these fields.

The *third step* involves collecting data on the independent and dependent variables as well as on relevant background characteristics. Lips (1993) asked her participants to respond to short, self-descriptive phrases, including 13 related to math and science affirmation, such as "good with numbers," and a yes-or-no format to reflect the likelihood of pursuing careers in various fields (including physical and engineering sciences).

The *fourth step* includes analyzing and interpreting the data. She reported that men and women with low levels of inclination shared low interests in careers in physical science and engineering. However, men with high inclinations reported high career likelihood, but women with high inclinations did not indicate high career likelihood. Thus, she concluded that for these women, liking mathematics and sciences did not positively predispose them to a career in the physical sciences and engineering.

As is necessary in causal comparative research, Lips (1993) explored alternative explanations for her results:

- A need to be aware of the effect of a scarcity of other women on women in math and science classes
- Perceived social norms and pressures associated with being a member of a highly visible minority
- Experiencing stereotyping and fewer opportunities for affiliation and social support
- Gender bias and inequities in the classroom

Thus, Lips's study demonstrates that a comparison of male and female characteristics would oversimplify interpretations if subgroup analysis and alternative explanations are not pursued.

Correlational Research

Correlational studies can be either *prediction* studies or *relationship* studies. In prediction studies, the researcher is interested in using one or more variables (the predictor variables) to predict performance on one or more other variables (the criterion variables). For example, kindergarten test scores can be used to predict first-grade test scores, if there is a strong relationship between the two sets of scores. In prediction studies, it is important to be aware of any other variables related to performance on the criterion variable. Relationship studies usually explore the relationships between measures of different variables obtained from the same individuals at approximately the same time to gain a better understanding of factors that contribute to a more complex characteristic.

It is important to realize that the correlation coefficient can range between plus and minus 1.00. The closer the correlation coefficient is to plus or minus 1.00, the stronger the relationship. A positive correlation means that the two variables increase or decrease together. For example, a positive correlation might exist between age and reading skills for deaf children, meaning that older children tend to exhibit higher reading skills. A negative correlation means that the two variables differ inversely; that is, as one goes up, the other goes down. For example, reading skills may be higher for children with less severe hearing losses—for example, as hearing loss goes up, reading skills go down. If the correlation coefficient is near zero, no relationship exists. For example, lipreading ability might be unrelated to reading skills in deaf children.

A word of caution should be entered here regarding the inadvisability of assuming cause and effect from correlational data. It is possible to calculate a correlation coefficient between any two sets of numbers:

- The number of PhDs in a state and the number of mules (it is strongly negative)
- The number of ice-cream cones sold and the number of deaths by drowning (it is strongly positive)
- The number of churches and bars in the same vicinity (it is strongly positive) (Beins, 1993)

There are obvious explanations other than causality for these correlations. Such high correlations that are due to some third variable (such as rural areas, hot weather, urban crowding) are called *spurious*. Nevertheless, it should be remembered that a high correlation does not in and of itself negate the possibility of a causal relationship (to wit: smoking and lung cancer).

An extension of this word of caution about assumptions of causality centers around the finding by researchers that the sum is somehow larger than the parts. In other words, even though a strong relationship may be found between a set of variables and an outcome measure, it is not always possible to then achieve the desired outcomes by manipulating the set of prediction variables. Wittrock (1986) uses the failure of the "input-output" model for effective teaching to make this point. He notes that researchers were able to find strong correlations between various teacher behaviors, such as use of positive reinforcement and student achievement. However, when teachers were trained to increase their use of such behaviors, corresponding increases in student achievement did not occur. He attributes the failure of the correlational approach to inappropriate theoretical assumptions that did not recognize cognitive variables inside the teacher and student and to contextual variables outside the teacher-student dyad.

Steps in Conducting Correlational Research: Relationship Studies

1. Identify an appropriate problem.
2. Identify variables to be included in the study.

3. Identify the appropriate research participants.
4. Collect quantifiable data.
5. Analyze the data and interpret the results.

The *first step* in correlational research, as in all other approaches, is to identify an appropriate problem. Remember, correlational research can be either for prediction purposes or to explain relationships between variables. Steps for conducting a relationship study are explained here, and the following section contains information specific to prediction studies.

One example of a relationship study is Solberg, Valdez, and Villarreal's (1994) investigation of variables related to Hispanic students' adjustment to college (see Sample Study 4.2 for a summary of this study). Their problem arose from the observation that Hispanic students had not fared well in postsecondary education; therefore, the researchers were interested in identifying variables related to college adjustment for this population.

The *second step* is to identify the variables to be included in the study. The variables in correlational research are sometimes called *explanatory* or *predictor* variables instead of independent variables because they are not experimentally manipulated. The dependent variable is then termed the *outcome* or *criterion* variable.

One advantage of correlational research is that several variables can be included in one study (more easily than in experimental or causal comparative designs). (Of course, the number of variables is moderated by sample size. The recommended number of participants per variable is 15, at a minimum.) However, the choice of variables should be done using a theoretical framework rather than a shotgun approach[1] (Borg & Gall, 1989). A researcher should give considerable thought to the variables chosen for inclusion for explanatory purposes. It is possible to "pour" many variables into the computer and then focus on those that come out statistically significant. Because statistics work on the theory of probability, with enough variables, it is probable that some will appear to be significant. It is more important that researchers include those variables that they have reason to believe are related to the outcome variable, based on previous research and theory.

In the Solberg et al. (1994) study, the authors wanted to test a "diathesis-stress" model that posits that mental health functions as an interaction between the amount of stress a person experiences and individual characteristics such as acculturation and social supports that could serve to minimize a person's negative experience of stress. Thus, theoretically, social support should moderate the relationship between stress and adjustment such that a person who perceives that social support is available will have a better adjustment level than someone without such support. Thus, the researchers used this theoretical model to select their independent (predictor or explanatory) variables: stress, social support, and cultural pride. Their dependent (criterion) variable was college adjustment.

The *third step* is to identify appropriate participants. Borg and Gall (1989) suggest that the groups be homogeneous or that subgroup analyses be done because variance in the criterion variable may be explained by different sets of variables for different subgroups. For example, in explaining high school dropout behavior for females, early pregnancy is an important variable; for males, economic need is a stronger predictor.

SUMMARY STUDY 4.2 *Summary of a Correlational Study of Relationship*

Research problem: Previous research suggests that Hispanic students had not fared well in postsecondary education; therefore, researchers were interested in identifying variables related to college adjustment for this population.

Research question: What is the relationship between Hispanic students' adjustment to college and individual characteristics, such as stress levels, acculturation, and social supports?

Method: The researchers used a correlational approach, collecting quantitative data on explanatory variables (stress, social support, and cultural pride), and examined their relationship to the participants' level of college adjustment.

Participants: The participants included 126 men and 268 women who attended a public university and who identified their family's country of origin as Mexico, a Latin American country, or an island in the Caribbean.

Instruments and procedures: Quantitative, paper-and-pencil scales were used to measure all variables. For example, stress level was measured using a 30-item scale that asked questions such as, "How often have you experienced difficulty trying to fulfill responsibilities at home and at school?" (p. 233). Students indicated their responses on a 5-point scale from 0 (*never*) to 4 (*very often*).

Results: Both stress and social support were significantly related to college adjustment; cultural pride was not found to be a significant variable. Academic and social stress and social support combined to account for 59% of the variance in college adjustment.

Conclusions: The diathesis-stress model was not supported in that social support was not found to buffer individuals from the negative effects of stress. "This research indicated that intervention and prevention programming aimed at addressing social factors such as living in the local community, handling relationships, and availability of social support are areas that are likely to have a positive impact on retention efforts targeted for Hispanic populations" (p. 237).

SOURCE: Based on Solberg, Valdez, and Villarreal (1994).

Solberg et al. (1994) chose to use 126 men and 268 women who attended a public university and who identified their family's country of origin as Mexico, a Latin American country, or an island in the Caribbean. They did not conduct subgroup analysis to determine if men's or women's college adjustment were explained by different patterns of variables.

The *fourth step* is to collect quantifiable data. For example, students' stress levels in the Solberg et al. (1994) study were measured using a 30-item scale that asked such questions

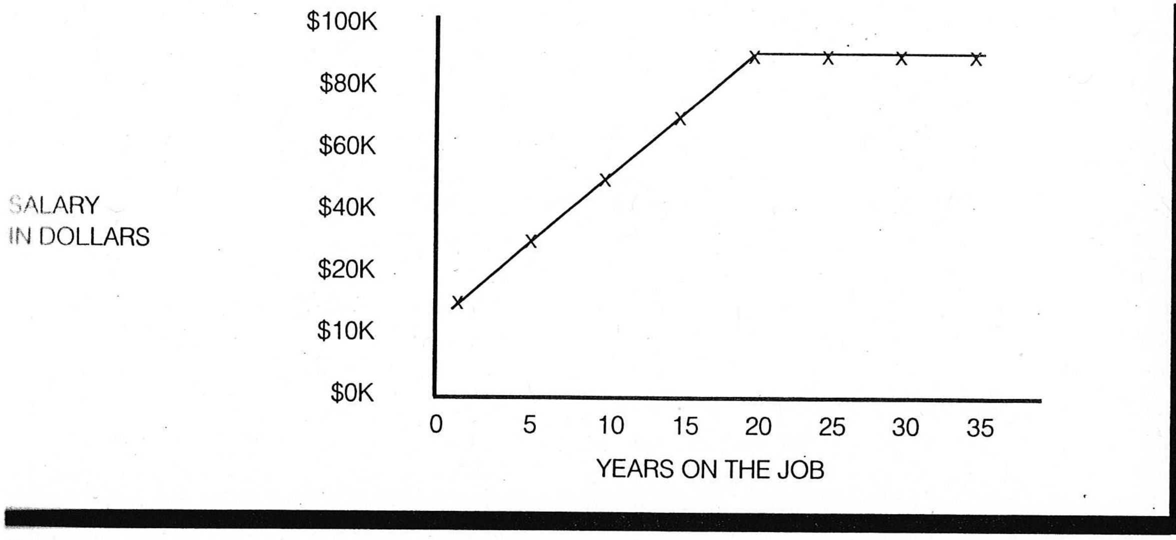

Figure 4.1.

as "How often have you experienced difficulty trying to fulfill responsibilities at home and at school?" (p. 233). The students indicated their responses on a 5-point scale from 0 (*never*) to 4 (*very often*).

The *fifth step* is to analyze the data and interpret the results. The researcher has a number of options for correlational analysis, including simple correlation, regression analysis, multiple regression analysis, discriminant function analysis, canonical correlation, path analysis, and factor analysis. These analytic techniques are described in Chapter 12. In this chapter, I explain some of the issues related to the use of statistical techniques for correlational research studies.

Graphs and Curvilinear Relationships. No matter what statistic is chosen, the researcher should always start with a graphic display of the relationships between the variables. One reason for this is that it gives you a commonsense base for interpreting the correlation coefficients that are subsequently calculated. Another *very* important reason is that simple correlation analysis is based on the assumption of a linear relationship between the variables. For example, as one's number of years in a job increases, one's salary increases. However, if a curvilinear relationship is depicted in the graph, simple correlation is an inappropriate statistical choice. For example, if a sample's ability to increase its earnings was restricted (because it had reached a "ceiling" within the organization or whatever), the relationship would be represented as shown in Figure 4.1.

The correlation coefficient would be low, suggesting a lack of relationship, when in actuality, a curvilinear relationship exists.

Choice of a Correlation Coefficient. The choice of a correlation coefficient depends on the scale of measurement. For variables with a continuous scale, the Pearson product-moment coefficient is typically used. For rank level data, Spearman's rho can be used. For nominal (dichotomous) data, a biserial correlation coefficient can be used. These are explained in more depth in Chapter 12.

Size and Interpretation. The interpretation of the size of the correlation depends on the purpose of the study. For relationship studies, a test of statistical significance can be applied to a correlation coefficient (see Chapter 12 for further discussion of this concept). For prediction studies, generally, a correlation above .60 is considered to be adequate for group predictions, and above .80 for individual predictions (e.g., school placement decisions).

Common or Explained Variance or r^2. Interpretation of correlation coefficients is often based on the amount of *common* or *explained variance* found by squaring the correlation coefficient. The explained or common variance refers to the variation in one variable that is attributable to its tendency to vary with the other (Gay, 1992). For example, Solberg et al. (1994) obtained a correlation coefficient of –.74 between stress and college adjustment scores for the Hispanic college students. Therefore, they reported that stress accounted for 40% of the variance in college adjustment. If stress was perfectly correlated with college adjustment, the two variables would have 100% common variance (and a correlation coefficient of –1.00). Because many variables other than stress influence college adjustment, the two variables have 40% shared or common variance.

Multiple Regression and Ordering Variables.[2] The order of entry for variables in multiple regression equations is important. When the predictor variables are correlated (a situation called *collinearity*), the amount of variance that each independent variable accounts for can change drastically with different orders of entry of the variables. Although there is no "correct" method for determining the order of variables (Kerlinger, 1973), the researcher must decide on a rationale for entry. If the researcher is interested in controlling for the effects of background characteristics before testing the effects of a treatment, it makes sense to enter the background characteristics first (e.g., see Andrews & Mason, 1986). Then the treatment variable will explain what is left of the variance.

Other possible rationales for entering variables include the following:

1. Enter the variables in order of their highest correlation with the criterion variable.
2. Enter them in chronological order.
3. Enter them in an order established by previous research.
4. Use a theoretical base for ordering.

To test their diathesis-stress model, Solberg et al. (1994) first entered the cultural pride variable into a hierarchial regression, followed by social support, stress, and the Social Support × Stress interaction term. They stated that if the entry of cultural pride and social support was statistically significant, the hypothesis of a direct relationship would be supported. However, for the diathesis-stress model to be considered, stress must also be directly related to adjustment, and the interaction of social support and stress would have to be significant.

In hierarchical regression, R^2 is generated as the amount of explained variance accounted for by the entry of variables in the equation. In the Solberg et al. (1994) study, statistical significance was determined by the level of R^2 change and whether the predictor or interaction term was significant (i.e., beta value) at the time it was entered into the equation. The R^2 changed from .00 to .21 with the entry of social support and from .21 to .58 with entry of the stress variable. These changes in R^2 indicated that stress and social support were significantly related to college adjustment. However, no significant changes in R^2 were found with the cultural pride variable or the interaction term of social support and stress. Thus, their diathesis-stress model was not supported in that social support was not found to buffer individuals from the negative effects of stress.

Discriminant Function Analysis. This statistical technique is used to predict group membership on the basis of a variety of predictor variables. For example, a number of different test scores could be used to see if they can discriminate between individuals who have mental retardation, learning disabilities, or no educational disability. Or it could also be used to see if measures of self-esteem, social skills, and participation in recreational activities can discriminate between people who are lonely or not lonely.

Canonical Correlation. Canonical correlation is also used to determine group membership; however, it can be used with multiple independent (explanatory or predictor) *and* multiple dependent (criterion) variables. For example, explanatory variables such as sex, socioeconomic status, and educational level can be combined with criterion variables such as income, employment, and prestige to determine if any discernible patterns emerge that could be used to separate people into groups.

Path Analysis. Path analysis is used when a researcher wants to test a causal theoretical model. For example, based on previous research, a casual model of academic achievement could be developed that included various student background characteristics (e.g., sex, ethnic status, presence or degree of disability), and instructional process variables (e.g., teacher expectation, degree of sensitivity to cultural differences). The researcher must specify the model in advance and then test to estimate the strength and direction of the relationships.

Factor Analysis. Factor analysis is an empirical way to reduce the number of variables by grouping variables that correlate highly with each other. For example, Solberg et al. (1994) identified 30 items related to stress. They factor-analyzed their participant's responses to these 30 items and determined that stress could be understood in terms of these factors: social stress, academic stress, and financial stress. Thus, they were able to reduce the number of variables from 30 to 3 (based on factor scores). For factor analysis, at least five participants per variable is recommended.

Cross-Validation. Perhaps all research seemingly could merit from replication to substantiate that the results are not a fluke. However, because of the lack of control and manipulation in correlational research, it is advisable to *cross-validate* the results of correlational analysis with a separate, independent sample. For example, would the stress variable divide into the same three factors if data were collected from another, similar sample?

Correlational Studies: Prediction

An example of a correlational study for predictive purposes is summarized in Sample Study 1.1. Doren, Bullis, and Benz (1996) were interested in making predictions as to which young adults with disabilities were more likely to be victimized within 1 year of leaving high school. In predictive correlational studies, the procedures are very similar to relationship studies. However, a few differences should be noted.

The *first step* is to identify the research problem. Doren et al. (1996) note that previous research suggests that people with mental retardation are vulnerable to economic, psychological, and physical abuse. They wanted to know more about the variables that might be used to predict who might become a victim.

The *second step* is to identify the variables to include in the study. Doren et al. (1996) selected variables based on previous research, including gender, minority status, serious emotional disturbance, specific learning disability, dropout status, family socioeconomic status, parent rating of academic skills, and a rating of personal and social skills.

In prediction studies, the researcher who focuses on one predictor variable (e.g., score on the Graduate Record Exam [GRE]) needs to be aware of the multiple criteria used to select people for admission to graduate school. A simple correlation between GRE scores and graduate school grade point average (GPA) would probably be low for a number of reasons:

1. Many criteria are used to select people for graduate school, including their undergraduate GPA, letters of reference, and personal position statement.
2. The range of scores for those accepted is restricted on a predictor variable such as the GPA.
3. The range of scores in the graduate school GPA (the criterion variable) is very restricted. (In many graduate schools, only A's and B's are acceptable for continuation in the program. If a student gets a C, he or she can be put on probation or dismissed.)

Thus, a high correlation between GRE and graduate school GPA could be obtained if a random sample of people took the GRE, all were accepted into graduate school, and all were allowed to remain in the program, no matter what grades they got. Not a likely scenario.

Researchers who conduct predictive correlational studies must be concerned not only with the number of predictor variables but also with the reliability and range of the criterion variable.

The *third step* is to identify appropriate participants for the study. Doren et al. (1996) included 422 adolescents from two Western states. The samples represented their respective populations in each state by primary disability category and gender. Population data on minority status was not available in the state's databases.

The *fourth step* is to collect quantitative data. One big difference that should be noted about prediction studies is that a time period must be allowed to elapse between the predictive variables and the criterion variables. In a prediction study, there is a need for an appropriate time delay between the measurement of your explanatory (predictor) variable(s) and your criterion variable. For example, suppose you wanted to use children's scores on a reading readiness measure to predict their ability to read at the end of first grade. You would need to administer the reading readiness measure at the beginning of the school year and then wait until the end of the school year to measure their reading abilities.

The researchers in the predictive study of victimization, quantified all the data concerning predictor and criterion variables. For example, victimization was coded as occurring if the students or their parents reported that the adolescent had experienced more than one of the following: being teased or bothered, having something stolen from him or her, or being hit hard or beaten up. The quantification of this variable allowed them to conduct the statistical analysis necessary for the prediction model. However, the researchers did acknowledge that this quantification was not able to capture the true nature of the adolescents' experiences of victimization. In terms of time delay, the researchers were able to interview students and parents once during the students' last year in school and once again when students were out of school for 1 year (Doren et al., 1996).

The *fifth step* is to analyze the data and interpret the results. The statistical choices for prediction studies are similar to those for relationship studies. One difference in predictive studies concerns the amount of variance that can be explained in the criterion variable. If predictor variables are to be useful, they should (in combination) explain about 64% of the variance in the criterion variable. This would translate into about .8 correlation between one predictor and one criterion variable. In the Doren et al. (1996) study, the researchers did not report the amount of variance accounted for by their model. However, they did report that a person who was both characterized as having a serious emotional disturbance and low on personal and social achievement was 20.48 times more likely to experience victimization.

▶ *Questions for Critically Analyzing Causal Comparative and Correlational Research*

1. Is a causal relationship assumed between the independent (predictor) variables and the dependent (response) variable? What unexpected or uncontrollable factors might have influenced the results? What competing explanations are explored?
2. How comparable are the groups in causal comparative studies?
3. Did the authors address group similarities and differences?
4. How did the authors operationally define who belonged in each group—for example, based on ethnicity or race or on disability? How did they address the issue of self-reporting versus actual self-identity? How were issues related to multiracial people addressed?
5. How did the authors address the fallacy of homogeneity?

6. How did the authors avoid the post hoc fallacy?

7. After the initial groups were defined, were subgroup analyses conducted, based on age, sex, socioeconomic status, or similar variables?

8. Could a third variable cause both the independent (predictor) and dependent (criterion) variables?

9. For correlational studies, what was the rationale for choosing and entering explanatory or predictor variables? What was the percentage of variance explained by the explanatory or predictor variables?

10. If a predictive relationship was studied, was the predictor variable the only criteria used to select participants in the study? Would combining the predictor variable with other screening criteria improve its predictive validity? (A predictive validity coefficient of about .8 is needed for an accurate prediction.)

11. What is the reliability of the criterion variable (compared with the test used to make the prediction)? Is there a restricted range for the criterion variable?

12. Were the results cross-validated with a separate, independent sample?

■ *Questions and Activities for Discussion and Application*

1. Through a computerized literature search or by going directly to the main journals that publish empirical, quantitative research studies, identify four research studies, two that use a causal comparative approach and two that use a correlational approach. For each study do the following:

   a. Identify the research problem.

   b. Identify the independent or predictor and dependent or criterion variables.

   c. Categorize the study as causal comparative or correlational.

   d. Explain the basis for your categorization.

   e. Critique the studies using the questions for critical analysis at the end of this chapter.

   f. For each study, note how the authors addressed the challenges of focusing on group differences, group identification, the fallacy of homogeneity, and the post hoc fallacy.

2. Brainstorm a number of different problems that would be appropriate for causal comparative or correlational research.

3. Select one research problem and explain how you would approach it using a causal comparative approach:

a. Identify the independent and dependent variables.

b. Explain how your approach to the study would satisfy the questions for critical analysis for causal comparative research.

4. Select one research problem and explain how you would approach it using a correlational approach:

a. Identify the independent or predictor and dependent or criterion variables.

b. Explain how your approach to the study would satisfy the questions for critical analysis for correlational research.

5. Under what circumstances would you *not* recommend using causal comparative or correlational approaches to research? What kind of alternative approach would you suggest?

Notes

1. A shotgun scatters the "shot" in a broad area; it does not hit a precise area. Thus, likelihood of hitting something is increased with a shot gun, but it may not be the precise thing that you intended to hit.

2. A sophisticated understanding of statistical practices is necessary to actually conduct multiple regression and make decisions about ordering variables. Nevertheless, you should be aware of the potential problems associated with inappropriate ordering of variables in research that uses this type of analysis.