Rea, L. M., & Parker, R. A. (2005). *Designing and conducting survey research: A comprehensive guide* (3rd ed.). San Francisco, CA: Jossey-Bass.

Selltiz, C., Wrightsman, L. S., & Cook, S. W. (1981). *Research methods in social relations* (4th ed.). New York: Holt, Rinehart & Winston.

Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1989). *An introduction to survey research and data analysis* (2nd ed.). Glenview, IL: Scott, Foresman.

Chapter 5

Correlation Studies

Correlation studies attempt to establish relationships between two or more quantitative variables. Participants with the characteristics of interest are selected and measured, and relationships are assessed via correlation coefficients. It is difficult—if not impossible—to establish causal statements from simple correlations alone. Observed correlations may be **spurious correlations**—i.e., due to unmeasured variables. A classic example is the positive correlation between ice cream sales and number of drownings. In this case, both ice cream sales and number of children swimming (and thus drowning) increases with daily temperature. The observed correlation disappears when temperature is held constant. Additionally, the direction of any presumed causal relationship may be difficult to establish because the variables may influence each other (e.g., loneliness and depression).

If only two variables are involved, we typically assess the relationship with a Pearson's product moment correlation coefficient r or **Pearson's *r***. It is considered a **zero-order correlation** because it does not control for any other variable. The magnitude of the coefficient describes the strength of the relationship, and the sign (positive or negative) describes the direction of the relationship. **Positive relationships** indicate that both characteristics covary in the same direction (e.g., amount of exercise and fatigue). **Negative relationships** indicate that both characteristics covary in opposite directions; an increase in one is accompanied by a decrease in the other (e.g., hours of sleep and fatigue). The magnitude ranges from zero (no relationship) to one (perfect relationship). The closer the value is to 1.00, the stronger the linear relationship or degree of covariation.

One of the most important assumptions made when r is computed is that the relationship between the two variables

is linear (i.e., a straight line). It is further assumed that the variables are accurately measured. The final assumption, known as **homoscedasticity**, is that the amount of scatter of the plotted points around that line is the same for the entire range of X scores. This indicates that the relationship between the two variables (X and Y) is equally strong for all values of X. Should the amount of scatter differ, it would indicate that the units of measurement in X, Y, or both have to be transformed or that if a relationship between X and Y exists, the relationship is not linear.

The assumptions become important when statistical tests are applied to the correlation coefficients. Once the sample r is calculated, it is tested statistically to determine whether it is large enough to infer that the population value is not zero (i.e., a linear relationship exists). Fortunately, the statistical test is somewhat robust with respect to violations. The degrees of freedom (df) associated with the test are $N - 2$, where N = number of pairs of scores. If a sample r is sufficiently large that it would be expected to occur rarely by chance (say 5 or fewer times in 100), we declare it significantly different from zero and conclude that a linear relationship exists.

Note that failure to obtain a significant coefficient may be due to a range of X values that is too restricted. The correlation between IQ scores and job proficiency among a group of computer programmers likely will be low, because computer programmers have relatively high IQ scores, leaving no room for variation in the X scores, even though proficiency scores may vary. Another factor that can lower the value of an existing correlation is the use of measuring instruments that are unreliable. The effect is called *attenuation*. Finally, if the sample is more heterogeneous or homogeneous than the population from which it was drawn, the sample correlation coefficient may not accurately estimate the population coefficient.

Whenever a correlation coefficient is declared significantly greater than zero, we run a risk that we've made an error in the decision we've reached: The coefficient may, in fact, be a chance occurrence. Such a value, on rare occasions, may occur even when two events are not correlated. When working at the .05 level of significance, we risk making this (Type I) error 5 times out of 100. If two coefficients are computed, each tested at the .05 level of significance, the risk of a Type I error with either decision is greater than .05. And the more coefficients we compute, the greater the risk of a Type I error when reaching a decision (i.e., declaring any coefficient significant when it is not) about a coefficient in the entire set. Behavioral and social scientists prefer to keep that overall risk of a Type I error at about .05. This can be accomplished in several ways. One way is to test each coefficient at a significance level that equals .05/C, with C = number of coefficients computed. If five coefficients are computed, each would be tested for significance at alpha = .05/5 = .01. Therefore, any coefficient that can occur by chance (when variables really are not related) more than one time in 100 (e.g., .02 or even .05) would be declared "not significantly greater than zero." Because such a procedure may be too conservative, an alternate remedy was suggested by Larzelere and Mulaik (1977). Coefficients are arranged in order of magnitude. The highest coefficient is tested at .05/C (e.g., .05/5 = .01). The next one is tested at .05/C − 1 (e.g., .05/4 = .0125). The next is tested at .05/C − 2 (e.g., .05/3 = .0167), and so forth, with the lowest coefficient tested at .05/1 = .05. This is extremely important because many authors have not made the adjustment. Therefore, some of the coefficients they declare significant, especially at the .05 level, may not be significantly greater than zero. For those wanting to be more precise, Box 5.1 presents a way of calculating an exact probability value for a reported correlation coefficient.

Box 5.1 Obtaining Exact Probabilities for Correlation Coefficients

In many cases, in the absence of the exact p value actually reported for each coefficient, that value can be approximated. If N is sufficiently large (at least 50), then we can determine the Z score for each r and read off the associated p value from the normal curve table. Here, $Z = r/[1/\sqrt{N}]$. If the actual p value (depending on Z) is equal to or lower than the critical p, then the r is significant.

The most common type of covariation is linear. And the most common linear correlation coefficient is the Pearson r. It measures the degree to which the two variables are linearly related. A more user-friendly description of degree of linear relation is r^2, the **coefficient of determination**. This describes the percentage of shared variance or percentage of variability in one characteristic that is accounted for by the other variable. It is an extremely useful statistic, which allows you to judge the usefulness of a coefficient declared to be significant. Thus, if r between IQ and grade point average (GPA) equals .45, then $r^2 = (.45)^2 = 20.25\%$ of variability in GPA is associated with IQ. Those with higher GPAs are likely to have the higher IQs. The remaining variability is due to time spent studying, other activities, and so forth. And if a coefficient of .25 is significant, you know that it only accounts for $(.25)^2 = .0625 = 6.25\%$ of variability.

Although cause and effect relationships are not likely to result from correlation studies, the effects of some obvious "third variables" can be removed statistically, by means of a **partial correlation**. The simplest removes the effect of one variable from the ones of interest and answers the following question: If participants are equal on variable X, what is the true correlation between A and B? Such a correlation that controls for the effect of one variable also is called a *first-order correlation*. One that controls for two variables would be a second-order correlation and so on.

(Continued)

Box 5.1 (Continued)

Pearson's *r* describes the linear relationship between a dependent and independent variable, but other sorts of data can be gathered and related to each other. Commonly, ordinal data, in the form of actual or imposed ranks, are assessed for their degree of relationship. The **Spearman rank correlation coefficient** is based on actual ranks or on data that have been ranked; it shows the amount of agreement between pairs of ranks. Because the degree of association between the ranks can be calculated by Pearson's method (provided that no ranks are tied), it is a correlation. However, the correlation cannot be used as a predictor (regression analysis) and generally is computed for descriptions of extent of agreement between ranks. The other popular method of determining a relationship between ranked or ordinal data is **Kendall's tau**, which is not a correlation. It measures the number of times pairs of ranks agree and disagree, relative to the number of pairs of ranks, a coefficient of agreement.

Box 5.2 Caution Factors in Correlation Studies

- Participants are randomly selected.
- Interviewers or data gatherers do not affect participant performance.
- Testing conditions are uniform.
- Measuring instruments are reliable and valid.
- Multiple instruments are presented in counterbalanced order.
- Measures vary sufficiently to yield a correlation.
- Variables show a straight line relationship (a scatter plot helps to see this).
- Important "third variables" are controlled.
- Adjustments in alpha are made when multiple correlation coefficients are presented.

Before we look at a correlation study, it is worth noting that we have to consider more than just the correlation coefficients in evaluating a correlation study. If individuals are selected to participate, they should be randomly selected. If tests are used to obtain the measurements, they should be valid and reliable. If more than one test is used, the order of presenting the tests should be counterbalanced (e.g., AB to half and BA to the other half). If other methods are used (e.g., ratings of some behavior), they should be valid and reliable. The data gatherers also have to be considered (sex, experience, training, etc.), as well as the location in which the data are gathered. All of these factors can differentially affect the performance of participants, leaving the correlation coefficients suspect. The major caution factors were summarized in Box 5.2.

The studies to be reviewed are concerned with establishing relationships between two variables at a time. We'll look at the first study together, and you'll evaluate the second one.

STUDY EXAMPLE 5.1: "CONTEMPORANEOUS AND LONGITUDINAL RELATIONS OF DISPOSITIONAL SYMPATHY TO EMOTIONALITY, REGULATION, AND SOCIAL FUNCTIONING"

In this study, the focus was on relating characteristics of young adolescents that measure a disposition to feeling sympathy with prosocial behavior. It was part of a longitudinal study that began when the participants were 4 through 6 years old, who were measured again when they were 6 through 8 years old, and again when they were 8 through 10 years old. Although longitudinal studies are favored for studying developmental characteristics, they risk losing participants and possibly end up with a selective loss. The other popular technique, a cross-sectional study, has its own risks: personality differences, because different individuals are measured at each level, and differential exposures to social factors, because the older individuals have lived longer.

The Study

Murphy, B. C., Shepard, S. A., Eisenberg, N., Fabes, R. A., & Guthrie, I. K. (1999). Contemporaneous and longitudinal relations of dispositional sympathy to emotionality, regulation, and social functioning. *Journal of Early Adolescence, 19*(1), 66–97. Copyright © 1999 by Sage.

In recent years, there has been increasing interest in children's vicariously induced emotional responses in relation to their social functioning. . . . *Empathy* often is defined as an emotional reaction resulting from the comprehension of another's emotional state. . . . A possible consequence of empathizing with another individual is sympathy. . . . *Sympathy* is an other-oriented reaction that involves feelings of sorrow or concern for another based on the other's perceived emotional state. However . . . another possible consequence of empathizing with another is a self-focused, aversive reaction, which has been referred to as personal distress. . . .

Empathy and sympathy theoretically have been linked to the quality of social functioning, particularly to positive social behaviors. . . . Theorists have posited that empathy and sympathy contribute to positive behavior because individuals who experience others' negative emotional states are motivated to reduce others' distress and to inhibit aggression toward others. . . . Unlike sympathy, personal distress reactions have been linked conceptually to low levels of prosocial behavior because the focus is on the self rather than on others. . . .

. . . Because empathy can result in sympathy as well as personal distress, there is a clearer conceptual link between sympathy and positive behaviors than between

empathy and positive behavior. However, researchers generally have focused on empathy rather than distinguishing between sympathy and personal distress. Therefore, it might be beneficial for researchers to focus specifically on sympathy rather than on empathy in general.

Although research exists on the relations of empathy and sympathy to prosocial behavior . . . there has been little research regarding the relation between sympathy and the quality of children's social functioning. . . . Sympathetic people would be expected to be relatively socially appropriate and sensitive. Therefore, one purpose for the present study was to examine young adolescents' dispositional sympathetic tendencies . . . in relation to a variety of aspects of their social functioning and adjustment.

Despite the conceptual importance of sympathy in children's social functioning . . . little is known about the characteristics of children who dispositionally are sympathetic. Thus, another purpose for the present study was to examine personal characteristics that are associated with individual differences in dispositional sympathy during early adolescence. . . . Sympathetic tendencies are likely to be associated with the abilities to manage vicariously induced emotions. Specifically, individuals who are able to maintain their emotional reactions within a tolerable range (i.e., not so arousing as to be highly aversive) would be expected to experience sympathy in empathy-inducing situations. . . . In contrast, individuals who experience high emotional arousal in empathy-inducing contexts would be likely to experience vicariously induced emotions as aversive. Thus, they would be expected to be self-focused and to experience personal distress rather than sympathy. . . .

The primary purpose for the present study was to examine further the relations of dispositional sympathy to emotional intensity, regulation, and social functioning during early adolescence. Much of the research to date has focused on younger children or adults, not on adolescents. . . . A variety of measures pertaining to 10- through 12-year-olds' dispositional sympathy, emotionality, regulation, and social functioning were obtained from teachers (school setting) and parents (home setting). Moreover, because participants were part of a longitudinal study . . . reports of emotionality, regulation, and social functioning were obtained 2, 4, and 6 years prior to the current assessment when participants ranged in age from 8 though 10 years, 6 through 8 years, and 4 through 6 years, respectively.

. . . Relatively high levels of regulation were expected to be associated with sympathetic tendencies, whereas negative emotional tendencies were expected to be associated with relatively low levels of dispositional sympathy. Emotionality and regulation are considered aspects of temperament and are relatively stable across time. . . . Thus, contemporaneous reports of young adolescents' emotionality and regulation as well as reports obtained 2, 4, and 6 years earlier were expected to be associated with dispositional sympathy at 10 through 12 years of age. . . . Young adolescents' sympathetic tendencies were expected to be related to the quality of their social functioning. Specifically, young adolescents viewed by parents or by teachers as possessing relatively high sympathetic tendencies were expected to be relatively high on constructive social behaviors as well as popularity and relatively low on problem behaviors. . . .

Gender also was considered in the present study to obtain a better understanding of young adolescents' sympathetic tendencies. . . . Boys and girls are exposed to differing socialization pressures . . . and these differences affect the ways in which boys and girls negotiate their social worlds. . . . There might be a stronger association between sympathy and regulation, emotionality, and social functioning for girls than for boys.

1. What was the rationale for the study?

Recent research has focused on relating children's emotional responses to their level of social functioning. Empathy can lead to sympathy or to personal distress. Empathy and sympathy are related to positive behaviors, whereas personal distress is related to low levels of prosocial behavior. Because empathy can lead to sympathy and personal distress, there is a clearer link between sympathy and positive behaviors than between empathy and positive behaviors. Whereas stress has been placed on empathy, it seems wiser to focus on sympathy.

Although empathy and sympathy have been related to prosocial behavior, little is known about sympathy and the quality of social behavior in children. Moreover, little is known about the characteristics of children who are sympathetic. One such characteristic is the ability to manage vicariously aroused emotions in an empathy-producing situation. Those who manage their emotions would be sympathetic; those whose emotional arousal is aversive would experience personal distress. Research has related sympathy to emotional intensity, emotional regulation, and social functioning of young children and adults but not of adolescents.

2. What were the purposes of the study?

One purpose was to examine young adolescents' tendencies toward sympathy. A second purpose was to examine the characteristics of young adolescents who are associated with different tendencies toward sympathy. The main purpose was to relate sympathy in young adolescents to intensity and regulation of emotion and to social functioning. Participants were 10 through 12 years old but had been studied at ages 4 through 6, 6 through 8, and 8 through 10 years.

High regulation of emotions was predicted to be associated with sympathy and low regulation (negative emotion) with low levels of sympathy, and the relationship

was predicted to be stable over time. Sympathy was expected to correlate with positive social behavior at home and at school. Girls were expected to show stronger relationships than boys.

Method

Participants

Participants were 33 girls and 31 boys who were part of an original group of 94 children studied approximately 6 years previously. For this group of 64 children, data were collected from 62 parents and 59 teachers. Participants were in the fourth, fifth, or sixth grade (= 10.96 years, *SD* = .60, range = 10.08 through 12.17), and 93% were non-Hispanic Caucasian (2% were Asian, 3% were Black, and 2% were Hispanic). Of the children, 89% lived in two-parent households, and annual household income ranged from $12,000 through $200,000 (= $81,170, *SD* = $30,710). Of the parents who participated, mothers' and fathers' years of education averaged 17.06 (*SD* = 2.40, range 12 [high school] through 20 [graduate school]) and 17.53 years (*SD* = 2.11, range 12 through 20), respectively. Of the 13 children who did not participate in the current follow-up (henceforth called T5) but had participated at the previous assessment, 1 refused participation, 5 could not be contacted, and 7 agreed to participate but did not return the questionnaires.

▶ *(Note that the bulk of children came from two-parent homes and were in high middle-income brackets. All parents had finished high school, and about one third of both mothers and fathers had completed graduate school.)*

The original group of participants was studied during two consecutive academic semesters approximately 6 years prior to the current assessment (henceforth called T1 and T2) when they were attending university kindergarten or preschools (45 girls and 49 boys; age = 5.37 years, *SD* = .56, range = 4.42 through 6.58). . . . Two years later (henceforth called T3), data were collected for 82 of the original participants (38 girls and 44 boys; age = 7.25 years, *SD* = .59, range = 6.42 through 8.50). Data at T3 were collected in the laboratory for 74 children and through the mail for 8 children. Seventy-seven children (36 girls, 41 boys; age = 8.84 years, *SD* = .58, range = 8.08 through 10.17) participated in the 4-year follow-up (henceforth called T4). Of the participants, 65 came in to the laboratory, and data were obtained through the mail for an additional 12 children.

▶ *(Note that there were considerable losses of male and female participants from the original group. A large percentage was lost 2 years after the first testing: 7/45 = 15% of the girls and 5/49 = 11.36% of the boys. Relatively fewer were lost at the second testing: 2/38 = 5.3% of the girls and 3/44 = 6.8% of the boys. But by the final testing, 6 years after the longitudinal study began, 3/36 = 8.3% of the girls and 10/41 = 24.4% of the boys were lost. In total, 12/45 = 26.6% of the girls and 18/49 = 36.7% of the boys from the original sample were lost. Because almost one third of the parents (at least in the final sample) were professionals [and 6% of the children were from single-parent homes], part of the loss could be attributed to a change in residence. But, if any of the children were among extreme scorers [although not outliers], this might have the effect of reducing the magnitude of a correlation. Note, too, that original testing was done at the school, and later testings were done in a laboratory and by mail. The final testing was done by mail.)*

3. Who were the participants?

The original sample (T1 and T2) consisted of 45 girls and 49 boys from university kindergarten or preschool. The second testing (T3) included 38 girls and 44 boys from the original sample, who were tested in the laboratory (74/82 = 90.25%) or via mail (8/82 = 9.75%). The third testing (T4) included 36 girls and 41 boys of the original sample, who were tested in the laboratory (65/77 = 84.42%) or via mail (12/77 = 15.58%). The final sample included 33 girls and 31 boys of the original sample. They were in Grade 4, 5, or 6. Most were Caucasian, from high-middle-income and two-parent homes. All were tested by mailed questionnaires.

4. Are there any potential sore spots to keep in mind?

A large percentage of participants were lost, which may or may not be a selective loss of high or low scorers. And testing conditions varied, which may or may not affect responses on the questionnaires.

Procedure

Unlike in previous studies involving this sample, all measures at T5 were collected through the mail. Primary caregiving parents were sent the same packet of questionnaires that they had completed in the laboratory 2 years previously. Many of the same or similar measures had been completed in the laboratory 4 and 6 years prior to this data collection. . . . Typically packets were completed by mothers (*N* = 57 at T5); however, five fathers at T5, one father at T4, and two fathers at T3 also completed the primary parent packet. All fathers were sent the problem behavior checklist to complete and return (*N* = 55 at T5). Toward the end of the

data collection at all assessments, participants' current teachers completed some of the same questionnaires as were completed by parents ($N = 59$ at T5). As in previous time periods involving this sample, parents and teachers responded to questionnaires in random order.

▶ *(This is a good feature of the design. Carryover effects from one questionnaire to the next are not a problem.)*

5. What was the general procedure?

Questionnaires were mailed to the primary care parent (mainly mothers), who were the same as or similar to those of prior testings. Current teachers also completed some of the same questionnaires. Orders of questionnaires were randomized.

Measures

. . . A summary of measures of emotionality, regulation, and social functioning is presented in Table 5.1.

Table 5.1 Composite Scores for Emotionality, Regulation, and Social Competence

Group	*Parent*	*Teacher*
	Negative emotionality	
T1/T2	Emotional intensity (averaged across T1/T2) and autonomic arousal (T2 only)	Emotional intensity (averaged across T1/T2) and autonomic arousal (T2 only)
T3, T4	Negative affectivity, negative emotional intensity, anger/frustration, attention shifting (reversed), falling reactivity/soothability (reversed)	Negative affectivity, negative emotional intensity
T5	Negative affectivity, negative emotional intensity, anger/frustration, falling reactivity/soothability (reversed)	Same as T3 and T4
	Regulation	
T1/T2	Attention control (attention shifting and attention focusing averaged at T2)	Attention control (attention shifting and attention focusing averaged at T2)
T3	Attention focusing, inhibition control, self-control, impulsivity (reversed)	Attention focusing, attention shifting, self-control
T4	Same as T3	Attention focusing, attention shifting, inhibition control, impulsivity (reversed)
T5	Attention focusing, inhibition control, self-control, impulsivity (reversed), attention shifting	Same as T4
	Social competence	
T1/T2		Same-gender and other-gender sociometric ratings
T3, T4	Mother and father ratings of problem behavior	Social skills, prosocial behavior, disruptive behavior (reversed), popularity, aggression (reversed), puppet responses (friendliness minus aggression)
T5	Same as T3 and T4	Social skills, prosocial behavior, disruptive behavior (reversed), popularity, aggression (reversed)

Measures of Sympathy

At T5, teachers and parents reported on young adolescents' sympathetic tendencies with six items. Teachers and parents responded to five sympathy items . . . using the . . . 4-point response scale. Teachers and parents were instructed to select one of the two statements (e.g., "This [my] child often feels sorry for others who are less fortunate" or "This [my] child does not feel sorry for those who are less fortunate") and indicate whether the chosen statement was *really true* or *sort of true* of each child. An additional sympathy item ("in general, to what degree does this [your] child feel sympathetic?") was rated using a 5-point scale (1 = *very slightly or not at all* [and] 5 = *extremely*). Items were standardized and averaged for teachers and parents (alpha = .92 and .84, respectively).

▶ *(Tests, based on items that were rated on a Likert scale, were administered to measure emotionality [negative emotional intensity, dispositional negative affectivity, and autonomic reactivity] and regulation [attention control, impulsivity, inhibition control, and global self-control]. There were standardized scales, so it is assumed that reliability and validity are established, and in all cases, alpha coefficients for the present samples are presented.)*

Measures of Social Functioning

Socially appropriate behavior. At T1 and T2, teachers and teacher aides completed seven items adapted from the Perceived Social Competence Scale for Children . . . to assess children's socially appropriate behavior. The Harter 4-point response scale was used to evaluate statements such as "This child is usually well behaved" or "This child is not well behaved." Teachers and aides selected one of the two statements and indicated whether it was *really true* or *sort of true* of each child. . . . At T3, T4, and T5, teachers completed four items from the original scale . . . .

Popularity. At T3, T4, and T5, as part of the measure of socially appropriate behavior, teachers rated children's popularity with three items (e.g., "This child has a lot of friends" or "This child doesn't have a lot of friends") using the same 4-point scale used for the socially appropriate behavior scale. Items were adapted from the Perceived Social Competence Scale for Children. . . .

Social behavior. At T3, T4, and T5, teachers rated disruptive behavior (e.g., "This child bothers kids when they are trying to work"; 8 items . . .), aggression (e.g., "This child says mean things to peers, such as teasing or name calling"; 8 items . . .), and prosocial behavior (e.g., "This child is good to have in a group, shares things, and is helpful"; 4 items . . .) with items from the Coie, Terry, Dodge, and Underwood . . . Teacher Checklist using a 7-point scale (1 = *never* through 7 = *almost always*). . . .

Problem behavior. Mothers and fathers completed the Lochman and Conduct Problems Prevention Research Group . . . Child Problem Behavior Checklist. Twenty-three items (e.g., "breaks things on purpose," "defiant toward adults") were rated on a 1 = *never* through 4 = *often* point scale. . . .

Children's ratings of sociometric status. At T1 and T2, procedures adapted from Asher, Singleton, Tinsley, and Hymel . . . were used to assess children's sociometric status. . . . Children sorted pictures of their classmates into three piles marked with smiling faces. The smiles ranged from a lot to virtually not at all to represent distinctions between peers that the target child "really likes to play with," "likes to play with some," and "likes to play with only a little bit." Children were assigned a score of 3 every time their peers placed them in the "really like" pile, a 2 for the middle pile, and a 1 for the least positive pile. Scores were summed and standardized within class and gender and then were averaged across T1 and T2 to create same-gender and other-gender sociometric status composites.

Enacted puppet procedure. At T3 and T4, children used puppets to act out how they would behave in five peer interaction situations. . . . Experimenters presented the following five stories in random order to the child (with props): (a) a peer criticizes and marks on the child's picture, (b) the child is excluded from a game, (c) the child is pushed into a water fountain, (d) the child is called "a baby" for playing with "baby toys," and (e) a peer grabs a toy away from the child. After each story, children used their puppet to show their response.

Children's responses to the puppet vignettes were coded using criteria adapted from Mize and Ladd. . . . Responses to each story were rated on a scale of 1 (*hostility*; i.e., negative outcome likely for the peer) through 5 (*friendliness*; i.e., prosocial outcome likely for the peer). A 5 [rating] included polite suggestions and consideration of positive outcomes for both parties, whereas a rating of 1 included responses such as name calling and destruction of the peer's property. Presence or absence of physical aggression also was coded. . . .

▶ *(Note that the enacted puppet test at T3 and T4 was a live procedure that had to be performed in person; hence, it was performed in the laboratory. Recall, however, that 8 children at T3 and 12 children at T4 were contacted through the mail and presumably did not take part in this procedure. Note, too, that there were many items to complete at each time period, and had the questionnaires not been randomized, there would have been a serious question about fatigue effects.)*

6. What was the general nature of the measures used in this study?

Standardized tests or some adaptations of them, all valid and reliable, were used to measure sympathy, emotionality, regulation of emotions, and social functioning. These scales were used to rate the children by the parents, primary care parent, and/or the children's teachers and teacher aides. Composite scores that reflected distinct characteristics measured by all similar tests were used in the final analyses.

Results

Initial analyses were conducted to examine participant attrition and relations of T5 variables with age and gender. In addition, across-reporter (parents and teachers) relations of sympathy, emotionality, regulation, and social functioning were examined. Subsequent analyses tested the predictions that in early adolescence, dispositional sympathy would be associated positively with regulation and constructive social behavior and associated negatively with negative emotionality and problem behavior, both contemporaneously and over time. . . . Results are presented for teachers' as well as for parents' rating of sympathy, negative emotionality, regulation, and social functioning. . . .

Participant Attrition

Of the original 94 children at T1/T2, 64 remained at T5. To examine whether T5 participants differed from those who discontinued participation throughout the longitudinal study, a series of *t* tests was computed to compare the ratings (i.e., sympathy, emotionality, regulation, and social functioning) that T5 participants received at T1/T2, T3, and T4 to the ratings of all of the children who subsequently dropped out of the study after T1/T2, T3, and T4.

Young adolescents (T5) did not differ on sympathetic tendencies (as rated by parents and teachers) from the children who discontinued participation throughout the longitudinal study. However, at T4, the T5 participants were rated by teachers as lower in negative emotionality than were the 13 children who did not participate subsequently at T5, $t(73) = -1.97$, $p < .05$. In addition, at T3 and T4, T5 participants were rated by teachers as higher on school social competence than were the 17 and 13 children, respectively, who did not participate at T5, *t*s(80 and 73) = 2.12 and 2.96, $p < .04$ and .004, respectively. Finally, at T1/T2, parents rated the 30 children who did not participate at T5 as lower on regulation than they did the children who participated subsequently at T5, $t(77) = 1.97$, $p < .05$.

▶ *(This check for selective loss is a very good feature of the study. It showed that no extreme scorers were lost with respect to sympathy, but some may have been lost regarding other measures.)*

7. What did the check for selective loss reveal? What does it suggest?

Compared to the current participants, participants lost at T4 were more emotional and showed less regulation of their emotions (the same was true of those lost from T1/T2). This suggests that correlations between sympathy and emotionality and sympathy and regulation may be lower than anticipated. Furthermore, participants lost at T3 were less socially competent than the present participants. This may affect the correlation between sympathy and social functioning.

Relations With Age and Gender

In general, gender differences obtained in the present study were similar to gender differences found at previous time periods. . . . Specifically at T5, teachers and parents rated girls higher than they rated boys on regulation,

*t*s(57 and 59) = 3.41 and 2.92, $p < .001$ and .01, respectively. However, teachers and parents did not rate boys and girls differently on negative emotionality. In addition, mothers rated boys higher than they rated girls on problem behaviors, $t(55) = -2.14$, $p < .05$. Finally, teachers and parents viewed girls as more sympathetic than they viewed boys at T5, *t*s(55 and 60) = 2.91 and 3.48, $p < .01$ and .001, respectively.

Because of the gender differences on numerous measures, results are presented by gender as well as for the total sample, and gender-related differences in findings also are noted.

8. What were the general findings regarding gender differences in ratings by teachers and parents?

At T5, girls were found to be more capable of regulating their emotions, in ratings from teachers and parents. Boys and girls did not differ in teacher or parent ratings on negative emotions. Mothers rated boys higher than girls on problem behaviors. Girls were rated as more sympathetic than boys by teachers and parents.

Correlations of Analogous Teacher and Parent Measures of Sympathy, Emotionality, Regulation, and Social Functioning at T5

Sympathy. Teacher and parent reports of sympathy were not related significantly at T5, $r(53) = .14$, ns. Findings across reporter did not differ markedly for boys and for girls.

▶ *(Note that expression of sympathy may have differed at home and in school. Moreover, lack of correlation means that correlations have to be calculated separately for parents and teachers.)*

Regulation and emotionality. Although regulation was related significantly across reporter (teacher and parent) at T5, negative emotionality was not related significantly across reporter at T5, *r*s(55) = .44 and .12, $p < .001$ and ns. Correlations were similar for boys and for girls.

Social functioning. Mothers' reports of young adolescents' problem behaviors were associated positively with analogous fathers' reports, $r(49) = .65$, $p < .001$. However, mothers' and fathers' reports of young adolescents' problem behaviors were not related significantly to teachers' reports of young adolescents' school social competence, *r*s(48, 47) = −.5 and .01, respectively, ns. None of the correlations differed significantly by gender.

9. What were initial results regarding parent and teacher ratings of sympathy, regulation and emotionality, and social functioning?

There was no relationship between parent and teacher ratings of sympathy, negative emotionality, nor social functioning. Ratings of regulation of emotionality by parents and teachers were correlated. And ratings of problem behaviors by mothers and fathers were correlated.

Relations of Dispositional Sympathy to Measures of Regulation and Emotionality

Reports of dispositional sympathy during early adolescence (at T5) were examined in relation to contemporaneous reports of their regulation and negative emotionality, as well as in relation to reports of regulation and emotionality obtained at previous time periods.... Zero-order correlations were computed separately for teachers' and parents' ratings to examine predictions within the different contexts of home and school.

Teachers' reports of sympathy, regulation, and emotionality. Contemporaneous and longitudinal zero-order correlations of teachers' ratings of early adolescents' dispositional sympathy with negative emotionality and regulation are presented in Table 5.2. Predictions pertaining to negative emotionality and regulation generally were supported....

As expected, contemporaneous ratings of young adolescents' sympathy and regulation at T5 were associated positively. In addition, children viewed as relatively high in regulatory abilities at ages 8 through 10 (T4) were viewed as relatively sympathetic 2 years later (T5). Furthermore, at T5, girls' but not boys' dispositional sympathy was associated with low negative emotionality.... Regarding longitudinal relations, teacher-reported sympathy during early adolescence generally was predicted by negative emotionality at younger ages. Specifically, for the total sample, young adolescents who were rated by their teachers as relatively high in sympathy at T5 were viewed by their teachers as relatively low in negative emotionality 2 ($p < .10$), 4, and 6 years earlier. Relations were somewhat stronger for girls than for boys (see Table 5.2).

Parents' reports of sympathy, regulation, and emotionality.... Hypotheses regarding contemporaneous ratings of parent-rated sympathy in relation to negative emotionality and regulation were supported, especially for boys. Moreover, regulation ... during younger years was associated with sympathetic tendencies during early adolescence.

Specifically, young adolescents (especially boys) who were viewed by their parents as relatively high in sympathy tended to be relatively high in regulation and

Table 5.2 Correlations of Time 5 Teacher and Parent Ratings of Sympathy With Analogous Reporter Ratings of Regulation and Emotionality Contemporaneously and Across Time

	Young Adolescents' Dispositional Sympathy as Rated by Analogous Reporter		
Measure	*Overall*	*Girls*	*Boys*
Teacher ratings			
Negative emotionality			
Time 5	–.26*	–.38**	–.06
Time 4	–.23*	–.42**	.08[a]
Time 3	–.31**	–28	–.25
Time 1/Time 2	–.27**	–.46***	–.01[b]
Regulation			
Time 5	.46****	.48***	.27
Time 4	.38***	.55***	.12[c]
Time 3	.26*	.21	.12
Time 1/Time 2	.19	–.03	.23
Parent ratings			
Negative emotionality			
Time 5	–.28**	.02	–.35*
Time 4	–.11	.18	–.18
Time 3	–.13	.08	.21
Time 1/Time 2	–.17	–.01	–.18
Regulation			
Time 5	.42***	.08	.48***
Time 4	.40***	.08	.47***
Time 3	.23*	.00	.26
Time 1/Time 2	–.12	–.24	.01

a. The correlations for girls and for boys were marginally significantly different, $z = 1.85$, $p < .10$.

b. The correlations for girls and for boys were marginally significantly different, $z = 1.73$, $p < .10$.

c. The correlations for girls and for boys were marginally significantly different, $z = 1.74$, $p < .10$. **$p < .05$; ***$p < .01$; ****$p < .001$ (*$p < .10$, trend).

relatively low in negative emotionality. Furthermore, parents who viewed their children as relatively high in regulatory abilities 2 years earlier (T4) rated their young adolescents as sympathetic at T5....

(Look at Table 5.2. Because raters were the same at any given time period [e.g., T5] and because N differed for the overall group, girls, and boys, we can look at three groups of four correlation coefficients, one for overall ratings at T5, one for ratings of girls at T5, and one for ratings of boys at T5. The first group [overall] consists of the following ordered coefficients: .46, .42, –.28, and –.26. The .46 should be tested at alpha = .05/4 = .0125, .42, at alpha = .05/3 = .0167; –.28, at alpha = .05/2 = .025; and –.26, at alpha = .05. On this basis, the two highest coefficients are significant but not –.28 nor –.26.)

10. What were the authors' results for the T5 correlations between sympathy and regulation and between sympathy and negative emotions for the group as a whole? What changes when alpha levels are adjusted for multiple correlations?

The coefficients were as predicted: There were positive correlations between sympathy and regulation of emotions as rated by parents and teachers and negative correlations between sympathy and display of negative emotions. When alphas are adjusted, the correlations between sympathy and negative emotions no longer are significant.

11. Correct the alpha levels for those coefficients reported to be marginally significant or significant at the .05 level. Do any conclusions change?

For T4 ratings by teachers, –.23 is not significant: Sympathy is not related to negative emotions shown 2 years earlier. Likewise, ratings of regulation for the group by teachers and parents at T3 (.26 and .23) are not significant. And –.27 at T1/T2 (ratings of negative emotions by teachers) is not significant. Regarding ratings of girls, the correlation between teacher ratings of negative emotions at T4 and sympathy at T5 (–.42) is not significant. And regarding the boys, the correlation between parent ratings of negative emotionality and sympathy (–.35) is not significant.

12. How do these reevaluations change the reported results?

With one exception, none of the reported results regarding negative emotionality was significant. Sympathy shown in adolescent girls, as rated by teachers, was negatively correlated with negative emotionality shown at the earliest age (T1/T5).

13. Consider the r = .46 (teacher-rated regularity and sympathy for overall group) and r = –.46 (teacher-rated negative emotionality at T1/T2 and sympathy). Interpret them in terms of r^2.

$(.46)^2 = .2116$ indicates that 21.16% of variability in teacher-rated sympathy in adolescents is associated with ratings of their current ability to control their emotions. Likewise, 21.16% of variability in teacher-rated sympathy of adolescent girls is associated with ratings of negative emotions displayed 6 years earlier; the higher the rating of emotional behavior, the lower the rating of sympathy by these girls.

Relations of Dispositional Sympathy to Measures of Social Functioning

Young adolescents' dispositional sympathy was examined in relation to contemporaneous ratings of their social functioning as well as social functioning measured 2, 4, and 6 years earlier. Zero-order correlations were computed to examine relations within the school setting (as rated by teacher or by peers) as well as the home setting (as rated by mothers and by fathers). Overall, the findings supported predictions.

Teachers' reports of sympathy and school functioning. In general, teachers' ratings of young adolescents' sympathy at T5 were associated positively with their school functioning contemporaneously as well as across time (see Table 5.3).

Table 5.3 Correlations of Time 5 Teacher Ratings of Sympathy With Social Competence and Sociometric Status Contemporaneously and Across Time

	Teacher Reports of Sympathy		
Measure	*Overall*	*Girls*	*Boys*
School social competence			
Time 5	.59****	.64****	.45**
Time 4	.33**	.63****	–.08[a]
Time 3	.26*	.25	.10
Time 1/Time 2	.36***	.55***	.06[b]
Same-gender sociometric status			
Time 1/Time 2	.31*	.22	.42**
Other-gender sociometric status			
Time 1/Time 2	–.02	.11	–.03

a. The correlations for girls and boys were significantly different, $z = 2.88$, $p < .01$.

b. The correlations for girls and boys were significantly different, $z = 1.98$, $p < .05$; **$p < .05$; ***$p < .01$; ****$p < .001$ (*$p < .10$, trend).

Specifically, young adolescents who were rated as sympathetic tended to be viewed as socially competent by their present teachers as well as by teachers 2, 4, and 6 years earlier. Furthermore, young children who were well liked by their same-gender peers at T1/T2 were viewed by their teachers as sympathetic young adolescents. The pattern of findings generally was stronger for girls than for boys.

14. What were the reported results regarding the correlation between social functioning, as measured or rated by teachers, and sympathy?

There was a positive correlation between the two variables at T5 for the overall group and for girls and boys. There was a positive correlation, overall, for ratings at the different times and current sympathy. The relationship between the two variables was stronger for girls than for boys.

15. Because teachers differed at the different times, consider the three coefficients (Table 5.3) at a single time as one group. The adjusted alpha levels at T5 should be the following: for .64, alpha = .05/3 = .016; for .59, alpha = .05/2 = .025; and for .45, alpha = .05. By these criteria, all three coefficients are significant. Perform similar adjustments for the rest of the table. What do you conclude?

At T4, the correlation between social functioning for the whole group and sympathy (.33) is not significant. At T3, the correlation between social functioning of the group and sympathy (.26) is not significant. And measures using sociometric scales at T1/T2 yielded no significant correlations with adolescents' sympathy.

Table 5.4 Correlations of Time 5 Parent Ratings of Sympathy With Mother and Father Ratings of Problem Behaviors Contemporaneously and Across Time

	Parent Reports of Sympathy		
Measure	*Overall*	*Girls*	*Boys*
Mothers' reports of problem behavior			
Time 5	−.46****	−.24	−.52***
Time 4	−.42***	−.17	−.43**
Time 3	−.42***	−.18	−.38**
Fathers' reports of problem behavior			
Time 5	−.46***	−.40*	−.47**
Time 4	−.17	−.11	−.32
Time 3	−.31**	−.14	−.39*

**$p < .05$; ***$p < .01$; ****$p < .001$ (*$p < .10$, trend).

Parents' reports of sympathy and problem behavior. As can be seen in Table 5.4, parents' reports of young adolescents' sympathy at T5 were associated negatively with mothers' contemporaneous reports of problem behavior as well as reports obtained 2 and 4 years earlier. . . . Parents' (mothers') reports of young adolescents' sympathy at T5 were associated negatively with fathers' reports of young adolescents' problem behavior contemporaneously and 4 years earlier but not 2 years earlier.

16. What did the authors report regarding parental ratings of sympathy and their reports of problem behavior?

Mothers' problem behavior reports correlated with adolescents' sympathy at T5, T4, and T3. Fathers' problem behavior reports correlated with adolescents' sympathy only at T5 and T3.

17. Adjust the alpha levels across each time interval. What do you conclude?

The correlations were significant for mothers' reports for the whole group, but for boys, this was true only at T5. The correlation was significant for fathers' reports only for the group at T5. All significant correlations indicate that the fewer the problem behaviors, the greater the rating of sympathy in the adolescents.

Discussion

Findings from the present study provide further support for the view that dispositional sympathy is associated with high regulation, low negative emotionality, and constructive social functioning across development. . . . Young adolescents' dispositional sympathy was examined in relation to their regulation, emotionality, and social functioning contemporaneously and as much as 6 years earlier. Individual differences in emotionality and regulation continued to affect sympathetic tendencies into early adolescence despite the numerous biological, psychological, and social changes individuals experience during the transition into adolescence. . . .

In interpreting the findings, it is important to note that relations might have been difficult to detect during early adolescence because there was less variability in emotionality, regulation, and social functioning at T5 than at previous assessments due to the loss of participants throughout the years. . . . It should be noted that the majority of the findings at school were for girls, whereas the pattern of findings at home was somewhat stronger for boys than for girls.

. . . Young adolescents who were viewed by adult raters (teachers or parents) as relatively high in sympathetic tendencies were viewed by the same raters as relatively high in regulatory abilities during early adolescence. . . . Across development, regulatory abilities likely prevent individuals from becoming overwhelmed when they experience another's distress, allowing them to focus on the other's distress rather than their own distress. . . .

In the school context, sympathy was associated particularly with high regulation for girls. . . . Teachers are likely to expect higher levels of regulatory abilities with age and increasingly might evaluate adolescents on the basis of the regulation of their vicarious emotional responding. This might be especially true for girls who might be expected to be more regulated than are boys in interpersonal contexts. Perhaps highly regulated girls increasingly are viewed as sympathetic because regulation and sympathy both are consistent with the adult stereotypic feminine role. In contrast, boys who are emotionally responsive to others might not be viewed as regulated because displays of emotion and caring are inconsistent with the adult masculine role. . . .

It is of interest that teachers' current reports of sympathy and previous reports of negative emotionality tended to be related, whereas this was not the case for parents' reports. . . . Teachers' ratings of young children's negative emotionality appeared to reflect displays of overt negative emotion such as anger and frustration, whereas mothers' ratings reflected those types of emotions to a lesser degree. Perhaps teachers continue to attend primarily to disruptive externalizing negative emotions; if so, it is not surprising that sympathy at school would be negatively related to displays of anger. In contrast, parents might attend to a greater range of children's negative emotions, making the relation of negative emotionality to sympathy somewhat complicated. . . .

. . . In the present study, young adolescent girls' dispositional sympathy was associated negatively with contemporaneous and longitudinal reports of their negative emotionality. . . .

Given that the pattern of findings at school between sympathy and emotionality/regulation was more consistent for girls than for boys, it is surprising that the pattern in the home context was somewhat more consistent for boys than for girls. However, parents rated girls higher than they rated boys in sympathy. To examine whether the different pattern of findings for boys and for girls could be due, in part, to differences in the distributions of parents' ratings of boys' and girls' sympathetic tendencies, a test of variances was conducted. Indeed, there was less variability in parents' ratings of sympathy for girls than for boys. . . . Thus, the restricted range of parent-reported sympathy scores could explain why findings in the home context were primarily for the total sample and boys rather than for girls.

As expected, both in the contemporaneous and longitudinal data, young adolescents' sympathetic tendencies were associated positively with appropriate social behavior, social competence, and peer sociometric status (T1/T2 only) in the school context (primarily for girls), and negatively associated with problem behaviors in the home context (primarily for boys). In the home context, the relation of sympathy to low levels of problem behavior held across reporter (i.e., when mothers reported on sympathy and fathers on problem behavior). . . .

In summary, the present findings are consistent with the view that dispositional sympathy is associated with individual differences in regulation, emotionality, and social functioning across development. . . .

18. What did the authors conclude? Are the conclusions justified?

Dispositional sympathy in adolescents is associated with high regulation of emotions, low negative emotionality, and constructive social functioning. But the relationships between sympathy and the other variables may have been difficult to detect because of the loss of participants over the years, which would reduce variability in ratings. Relationships at school were stronger for girls (especially between sympathy and regulation), whereas relationships at home tended to be stronger for boys.

Teachers' but not parents' current rating of sympathy tended to be related to present and past ratings of negative emotions. This was particularly true of adolescent girls. More consistent relationships for girls at school and for boys at home is attributed to a more restricted range of ratings, at home, for girls than for boys.

Adolescent sympathy was positively associated with present and past social competence and sociometric ratings and negatively associated with problem behavior at home. Positive relationships were stronger for girls, and negative relationships were stronger for boys.

Conclusions basically are justified. Associations involving negative emotions mainly were not significant, although they were in the predicted direction, when alpha levels were adjusted. And correlations involving sociometric ratings and adolescent sympathy were not significant, although they were in the expected direction, when alpha levels were adjusted.

19. What probably accounts for lack of stronger, predicted relationships? What might be done to overcome the problem?

As noted by the authors, the loss of high (particularly in negative emotions) and low (particularly in social competence) scores could have the effect of reducing variability in the remaining scores, leaving a more restricted range and lower correlation coefficients. Because this was part of a longitudinal study, not much could

be done to restore participant scores that have been lost. But a new, larger group of adolescents might be selected and rated on the same variables. Larger groups make it more likely that a range of scores will be obtained, and if the correlations do exist, as this study suggests, they would be evident with the new sample.

STUDY EXAMPLE 5.2: "DESIGNING SUCCESSFUL SERVICE LEARNING PROJECTS FOR URBAN SCHOOLS"

This study, which you will evaluate, concerns the effects of course community service projects on various attitudes toward learning and academic performance. The statistical analyses included correlations and more advanced techniques. You'll only have to focus on the correlations.

The Study

Moore, K. P., & Sandholtz, J. H. (1999). Designing successful service learning projects for urban schools. *Urban Education*, *34*(4), 480–498. Copyright © 1999 by Sage.

Few people will argue against the value of community service but not all see the connection to the public school system.... In the past decade, hundreds of schools across the country implemented service learning programs or required community service for graduation.... Service learning is part of a widespread movement aimed at rebuilding communities and reforming public education.

As the name implies, service learning is different from community service in that it has two inseparable components: service and learning. Service learning is typically defined as a "pedagogical technique for combining authentic community service with integrated academic outcomes."... The premise is that students' involvement in community service can also contribute to their learning and growth. ...Some supporters suggest that the balance between service learning and academic learning in schools needs to shift, with service learning becoming an integral part of the school curriculum rather than an extracurricular activity....

As part of the curriculum, service opportunities... [are] employed as a curricular tool to help students bridge the gap between theory and reality....

A primary advantage of using service learning as a teaching methodology is that... learning stems from direct experience.... However, data from the National Service-Learning Clearinghouse indicate that the primary focus of most programs is personal and interpersonal knowledge with the least focus on academic or subject matter learning....

Those directly involved in service learning programs report a variety of benefits for participating students, including enhanced self-esteem,... clarification of values,... social and personal development,... and enhanced academic performance.... Service learning can be an effective way to increase the connection of youth with their society and to develop good habits of citizenship.... However, to achieve these purposes, service learning programs must... encourage and require student reflection on the dynamics of volunteerism, the responsibilities of citizenship, and an understanding of multicultural societies and the larger meaning of community. Through this process, students' attitudes about themselves (academically and socially) and their neighborhoods may change over the course of their participation in such a program.

Despite their potential, not all service learning programs achieve positive changes in students' attitudes or performance.... If service learning programs are to achieve their promising potential, researchers need to continue examining the experiences of students involved in these projects and to expand the knowledge of the key features needed for successful programs.

In this article, the authors draw on data from a 3-year evaluation to compare the strengths and weaknesses of 10 different service learning programs. The authors provide a brief background about the program, describe the research design, examine critical elements in designing effective service programs, and discuss the connections among them.

Background

The service learning programs included in this research took place in a large urban school district in Southern California. The district is located in an economically depressed community with high rates of unemployment and crime. Approximately one third of the county's population is younger than the age of 18 years. School and community leaders are concerned with trends in the youth population such as gang violence, drive-by shootings, teenage parenting, and failure to graduate from high school. One primary objective of the service learning programs was to help change these trends by involving the youth in citizenship training through service.

The district began its venture into service learning by piloting the concept at a single high school during the first year. In the second year, the program expanded to three high schools. During the third and final year of the evaluation, the district implemented service learning at all six high schools in the district. The 172 participants included approximately an equal number of male and female students, but the majority were either freshmen or sophomores. The ethnic composition of the program participants was 60% Hispanic, 20% Caucasian, 9% African American, 9% Asian American, and 2% other.

At the beginning of each year, interested teachers at each high school attended a training session on service learning where they were encouraged to adopt service

learning as a part of their regular curriculum. Service learning was broadly defined to give teachers the maximum amount of flexibility and to encourage creative projects that would fit both the service and learning aspects of the definition. Because service was strongly recommended but not mandated, participation by teachers was voluntary.

The 10 projects included in this study . . . are contained in Table 5.5.

Table 5.5 Description of Projects

Project	*Description*
World History Awareness (world history class)	This project involved high school students in making presentations to elementary students on the importance of world history.
Head Start (interdisciplinary English and social studies class)	This project placed students with Head Start preschoolers for a limited time. Students worked with the preschoolers as program volunteers.
Miles of Smiles (English-as-second-language class)	This project trained limited-English-proficient students as guides for non-English-speaking guests at the high school.
Peace Coaching (English class)	This project involved high school students acting as coaches and mentors for kindergarten student-designed children's books to help elementary students learn various subjects.
Quilt Blocks (mathematics class)	This project required students to design a 12-inch-square quilt block using at least one geometric transformation. The students were each given a packet of materials and a brief lesson on sewing. Each student was responsible for cutting out pattern pieces for his or her design. Most of the students tried to sew their own square with the help of family and friends. The blocks were sewn together into unified quilts that were donated to local organizations.
School Guides (English-as second-language class)	Duplicate of Miles of Smiles at a different school.
Snagology (integrated science class)	A short-term project that encouraged environmentalism through recycling.
Spanish Song Books (Spanish class)	This project involved a Spanish honors class. The class [was] divided into groups of four students and created song boards and audiocassettes that were sent to elementary schools in the area. This project lasted 1 week and was completed in class.
Stereotypes Presentation (English class)	This project focused on the negative power and energy of stereotyping. Students were trained to present a workshop on tolerance to their peers.
Teachers for the Dream (English class and after-school club)	This project gave future teachers an opportunity to practice their skills with younger students (middle school and elementary). Student teachers addressed the whole child during the project, helping with academic and social problems. The project lasted the entire year.

Method

For this study, the authors used both quantitative and qualitative data from the third year of the program to examine the greatest number of service learning projects. The 172 students participating in service learning completed a 35-question attitudinal survey at the end of their projects. The survey addressed five main categories of student attitudes: academic success, school socialization, future planning, self-perception, and community pride. Academic success encompassed issues of individual academic achievement, grades, participation in class discussions, and students' perception of the efficacy of instruction. This category referred to all classes, not just the specific course that included service learning. School socialization included students' relationships with other students, attitudes toward multiculturalism in the school setting, participation in school activities, and general perceptions about the high school experience. In the area of future plan[ning], students responded to inquiries regarding their postsecondary educational plans and the part altruism might play in their life goals. The self-perception category included students' perceptions regarding their appearance, achievements, happiness, and respect for themselves and others. Finally, questions regarding community pride [were] asked [of] students about their perceptions of their neighborhoods with regard to safety, desirability, and pride.

▶ *(Please keep in mind that influence cannot be determined in a correlational study.)*

To determine the extent to which program features influenced student attitudes in each of the five areas, the authors rated each of the 10 projects according to four design elements: duration, location, personal contact, and focus of the project. The design elements were identified based on 3 years of studying and evaluating both successful (as judged by teacher and student input) and less-than-successful service learning programs. However, there may be a number of other possible factors that might have influenced the attitudes of the students involved in the service learning projects studied: teacher quality, student interest in the subject matter, socioeconomic status [SES], teacher interest, and so on. Although this study does not determine definitely all of the design elements for successful service learning programs, it does begin to identify key characteristics of programs and provide a position for future research.

The authors' experience suggests that duration, location, personal contact, and focus of the project are elements that may contribute to the success (or failure) of service learning endeavors. The authors developed a rating scale of 1 to 5 (*low* to *high*) to delineate the range in program components found in the projects studied. Table 5.6 details the rating scale. . . .

Table 5.6 Scoring of Four Critical Elements

Scale	Duration	Location	Personal Contact	Focus of Project
1	Project lasts less than 4 weeks	Project done at home and then brought to school	No direct contact with beneficiaries of project	Includes little or no service
2	Project lasts from 4 weeks to 9 weeks	Project done in the classroom	A representative makes contact (i.e., contact by another student or the teacher) with the beneficiaries of the project	Some focus on service
3	Project lasts from 9 weeks to 18 weeks	Project done at the high school campus	Students have a single contact with the project's beneficiaries	Moderate amount of service
4	Project lasts from 1 semester to 1 year	Project done at another school's campus (often an elementary or middle school)	Students have sporadic contact with the project's beneficiaries	Significant amount of service
5	Project lasts longer than 1 year	Project conducted in the community	Students have repeated and sustained contact with the beneficiaries of the service learning project	Extensive focus on service

The ratings for each project on duration, location, personal contact, and focus are given in Table 5.7.

Table 5.7 Rating of Each Project

Project	Duration Rating	Location Rating	Personal Contact Rating	Focus of the Project Rating	Total Points	Rank of Project
World History Awareness	2	4	3	3	12	4
Head Start	2	4	4	4	14	3
Miles of Smiles	1	1	1	1	4	9
Peace Coaching	3	4	5	4	16	2
Quilt Blocks	3	2	2	1	8	5
School Guides	1	1	1	1	4	9
Snagology	1	3	1	3	8	5
Spanish Song Books	1	2	1	3	7	7
Stereotypes Presentation	1	2	1	1	5	8
Teachers for the Dream	5	4	5	5	19	1

. . . The authors . . . looked at the correlations between the five attitudinal categories and the four criteria. . . .

Quantitative data provided corroborating information regarding student gains and program effectiveness. Students wrote an essay prior to participation in service learning addressing the following query: "Give a brief description of what you hope to gain from this service learning project." Following participation in a service learning project, the students responded to a similar essay: "Give a brief description of what you gained from this service learning project. How does what you gained compare to what you thought you might gain before starting this project?" Qualitative data were analyzed in conjunction with the quantitative data to assist in determining the extent of each project's effectiveness, based on the four design elements. . . .

Results

There was broad variation in the 10 service learning projects evaluated, resulting in a wide range of attitudinal outcomes in the five areas measured (i.e., attitudes toward academic achievement, school socialization, self-esteem, future plans, and one's community). The results of the . . . correlational analysis (Table 5.8) . . . indicate that the higher the four criteria (duration, location, personal contact, and focus), the higher the students' attitudes toward academic achievement, plans for the future, self-esteem, school socialization, and their communities.

Table 5.8 Correlations of Attitudes With Criteria

	Academic Achievement	Plans for the Future	Self-Esteem	School Socialization	Attitudes Toward the Community
Duration of project	$r = .09$	$r = .34$	$r = .14$	$r = .25$	$r = .06$
	$p < .24$	$p < .0001$	$p < .06$	$p < .001$	$p < .45$
Location of service learning	$r = .14$	$r = .32$	$r = .32$	$r = .38$	$r = .14$
	$p < .07$	$p < .0001$	$p < .0001$	$p < .0001$	$p < .07$
Extent of personal contact	$r = .16$	$r = .39$	$r = .21$	$r = .29$	$r = .08$
	$p < .03$	$p < .0001$	$p < .005$	$p < .0001$	$p < .27$
Focus of the project	$r = .08$	$r = .16$	$r = .27$	$r = .24$	$r = .11$
	$p < .33$	$p < .04$	$p < .0004$	$p < .002$	$p < .15$

Duration

Results indicate that the longer the service learning project, the greater the positive attitudinal scores regarding future plans and school socialization.

Students participating in projects that lasted less than 1 month showed the smallest attitudinal scores in these areas. Projects that lasted continuously over the student's entire high school experience corresponded with the higher scores. . . .

Students who spent the longest time working with people in their communities who needed their help reported getting along better with the students at their own schools and were better able to solidify their goals for the future. The data in this study support what proponents of service learning have long maintained: Longer is better.

Location

Service learning projects took place in three locations: at the student's home high school, at a school district other than the home school of the student, and in the community at large. Those projects that involved the student working with others away from his or her home campus corresponded with higher student attitudinal scores in the areas of future plans, self-esteem, and school socialization. Projects that took place on the home campus yielded low scores in these areas. Those projects that involved students in the neighborhood, such as those run in conjunction with the local volunteer center, produced the greatest overall scores.

Personal Contact With Recipients

Although some projects involved an interaction with residents of the community, they did not involve personal interaction. Projects of this type usually consisted of producing a product in class that was then distributed to community recipients. . . . The data indicate students had higher attitudinal scores when they had direct personal contact with the beneficiaries of their efforts. . . .

Of the four criteria, personal contact was the most significant. Although face-to-face contact was not the norm among the service learning projects studied, those projects that included this aspect showed higher student attitudinal scores in four of the five areas: academic achievement, plans for the future, self-esteem, and school socialization. . . .

Focus of the Project

Greater positive attitudinal scores occurred in students who participated in projects that had an emphasis on service with learning as a necessary [by-product] rather than in projects that had an emphasis on learning with service as an added component. Projects that were designed to provide a service to others generally contained a learning component. . . . Projects that concentrated most extensively on service showed significant results in three of the five areas (plans for the future, self-esteem, and school socialization).

. . . Comments from student essays point out the differences in students' perspectives. Projects with an academic focus alone elicited responses such as "I am doing this project because it is a part of my grade" and "I am participating because my teacher made me." In describing what was gained from participating, students emphasized academic learning. . . . In contrast, students who participated in more service-oriented projects focused on personal values. For example, one student wrote, "When I first started the project I thought it was stupid. But later after I did the assignment I realized that I was doing a good deed by taking a little time out to make other people's days brighter." . . .

Discussion

The two key components of service learning, service and learning, directly correspond with the two most common goals of service learning projects in public schools: an increase in academic achievement and the inculcation of citizenship. This analysis supports the proposition that service learning projects of longer duration that place students in direct contact with service beneficiaries away from the school campus result in higher student attitudinal scores. However, the critical question is . . . Do higher attitudinal scores indicate or lead to increases in academic achievement and citizenship? The authors argue that they can.

With regard to citizenship, students who participated in service learning projects that scored high on the four criteria (duration, location, personal contact, and focus) reported that they were more involved in school activities, got along better with their peers, enjoyed being in school, and cared more about what their teachers thought of them than students who participated in service learning projects with low ratings. At the same time, these students reported that they knew what career paths they would follow after high school and that they believed helping other people was important to their futures. . . .

The challenge for service learning has not been in demonstrating its effect on citizenship but in establishing its connection to increased academic achievement. The data in this study indicate that . . . an increase in personal contact with the beneficiaries of service correlates with an increase in attitudes toward academic achievement, and students who travel away from the school site to perform service show an increase in self-esteem. . . .

. . . Researchers examining motivational issues argue that students who feel that they control their learning perform at higher academic levels than those who do not. One suggested method for maximizing control is to increase project-based learning. . . . Students who were engaged in a project over a longer period of time and who controlled the project by engaging directly with the beneficiaries reported higher levels of self-esteem and more positive attitudes toward academic achievement.

Although these data do not provide direct evidence of improved academic achievement, they do demonstrate increases in the motivational elements that lead to achievement.

Conclusions

The results of this research lead to three main conclusions regarding the implementation of service learning programs into the public school curriculum. First, to have a meaningful influence on students, service learning projects must achieve a delicate balance between the service and learning components. . . . The process of combining both service and learning component into a viable program is complex. . . . The results of this study suggest that projects with a clear emphasis on service fare better than those with service as an adjunct feature. . . . The learning component must be meaningfully incorporated into the service component. . . .

Second, many of the critical features of successful service learning projects are challenging to implement within the current school structure. This research suggests that personal contact with the beneficiaries, locating the project away from the school site, and conducting the project over a long period of time are crucial if both service and learning goals are to be achieved. . . . Once back in the classroom, the integration of service and learning must be completed and solidified. . . .

Third, the benefits of service learning may not be readily apparent over the short term. Although service learning may not result in immediate, short-term increases in academic achievement (as measured by standardized tests), it can more immediately affect student attitudes. . . . These results indicate that the longer the service learning project, the greater the positive attitudinal scores for students. Service learning programs that begin early and are continuous throughout schooling appear to achieve the greatest gains with respect to both citizenship and academic performance. . . .

CRITIQUE OF STUDY EXAMPLE 5.2

1. What was the rationale for the study?
2. What was the purpose of the study?
3. Where did the programs take place over the course of 3 years? Why were they introduced?
4. What are the general characteristics of the participating students? What information is missing? Of what consequence might it be?

5. Only teachers who volunteered to participate introduced programs in their curriculum. What problem may this present when evaluating the conclusions?
6. What was the general nature of the programs?
7. What five attitude-related variables did students rate after the project was finished?
8. What can be said about reliability and validity of the questionnaire?
9. What were critical aspects of the projects? How were they determined?
10. Are there any weaknesses of this rating procedure?
11. Are there any questionable ratings in the rating scale, found in Table 5.6?
12. Are there any questionable ratings of the projects, found in Table 5.7?
13. What was the time relationship between administration of the questionnaire and the essay written at the end of a project?
14. Did all students fill in the questionnaire at the same time? At the same place?
15. Look at the correlation coefficients in Table 5.8. There are 20 reported, presumably without adjusted alpha levels. Arrange them from highest to lowest and adjust each alpha, starting with $\alpha = .05/20 = .0025$ for $r = .39$. What do you conclude?
16. Interpret $r = .39$ and $r = .24$ in terms of r^2.
17. Is the statement regarding the relationship between duration of the project and attitudinal scores consistent with the results of the correlation analysis?
18. Is the statement regarding the relationship between location of the project and attitudinal scores consistent with the results of the correlation analysis?
19. Is the statement regarding the relationship between degree of personal contact and attitudinal scores consistent with the results of the correlation analysis?
20. Is the statement regarding the relationship between focus of the project and attitudinal scores consistent with the results of the correlational analysis?
21. What did the authors maintain the results showed?
22. Are these statements justified?
23. What did the authors conclude?

24. Are the conclusions justified?

25. What might account for the low correlation coefficients that were obtained?

26. What factors, other than the projects, might have resulted in positive attitudes?

For answers to these questions, see page 358.

BIBLIOGRAPHY

Larzelere, R. E., & Mulaik, S. A. (1977). Single-sample tests for many correlations. *Psychological Bulletin, 84*(3), 557–569.

SUGGESTED READINGS

Ary, D., Jacobs, L. C., Razavieh, A., & Sorensen, C. K. (2009). *Introduction to research in education* (8th ed.). Orlando, FL: Harcourt Brace.

Chen, P. Y., & Popovich, P. M. (2002). *Correlation: Parametric and nonparametric measures*. Thousand Oaks, CA: Sage.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Hays, W. L. (1994). *Statistics* (5th ed.). Belmont, CA: Wadsworth.

Howell, D. D. (2009). *Statistical methods for psychology* (7th ed.). Belmont, CA: Wadsworth.

Keppel, G., & Zedeck, S. (1989). *Data analysis for research designs: Analysis of variance and multiple regression/correlation approaches*. New York: Freeman.

McMillan, J. H., & Schumacher, S. (1997). *Research in education: A conceptual introduction* (4th ed.). White Plains, NY: Longman.

Chapter 6

Regression Analysis Studies

When two variables are found to be correlated, one may be used to predict the other. If extroversion is shown to be correlated with salesmanship (both measured at the same time), extroversion now may be used to predict salesmanship (which will be measured at a later time) of prospective salespersons. **Regression analysis** deals with predicting certain continuous, behavioral variables on the basis of knowledge about other, independent variables, some of which may be **dummy coded** categorical variables. Or it describes the role played by certain independent variables in a particular dependent variable. Thus, it may describe how well graduate record exam (GRE) scores predict performance in graduate school, or it may describe the average amount of change in blood glucose (BG) for each unit change in stress. The variable that is being predicted (Ŷ: salesmanship, graduate school success, BG) is called the **criterion variable**. The variable from which the prediction is made (X: extroversion, GRE, stress) is called the **predictor variable**. When we are concerned with a single dependent variable but more than one independent (or predictor) variable, the problem is called **multiple regression analysis (MRA)**. In this instance, the regression equation is

$$\hat{Y}=b_0+b_1X_1+b_2X_2+\ldots+b_pX_p$$

Here, the *predicted* criterion variable is designated Ŷ. Assuming a straight line relationship between Y (the *actual* criterion variable) and Xs, b_0 is the Y intercept, the average value of Y when each X equals 0, and the remaining *b*s are regression coefficients or, more accurately, partial regression coefficients. Each reflects the average amount of change predicted in Y for each unit change in the corresponding X when all other independent variables are held constant. That is, each *b* is a partial