David M. Wulff (1996). *Psychology of religion. Classic & contemporary.* New York: John Wiley & Sins, (Chapter 6, pp. 205-257).

6

THE CORRELATIONAL STUDY OF RELIGION

Psychologists of religion have long preferred the more naturalistic approach offered by correlational methods. The earliest American contributors of the modern period, including Stanley Hall, James Leuba, Edwin Starbuck, and George Coe, made extensive use of frequency counts and tabulations of questionnaire replies to discern trends in religious behavior and experience. Starbuck (1899), for example, reports the frequency of conversion by age, sex, and theological background, and then correlates these data with changes in height and weight during the childhood and adolescent years. He tabulates the reported "mental and bodily affections" that characterized the preconversion states of his subjects; the motives and forces that led to conversion; the elements that composed the conversion experience; the feelings that followed the experience; and the character of the postconversion struggle. Starbuck also describes statistically the course of religious belief through the life-span and the relative prominence of religious feelings and ideals in adults of various ages.

Galton and the Efficacy of Prayer

The earliest systematic investigation of religion by correlational techniques, such as they were in the latter half of the nineteenth century, was undertaken by Francis Galton. Charmed by statistics, Galton became convinced of their value in exploring and understanding human variation.

> Some people hate the very name of statistics, but I find them full of beauty and interest. Whenever they are not brutalized, but delicately handled by the higher methods, and are warily interpreted, their power of dealing with complicated phenomena is extraordinary. They are the only tools by which an opening can be cut through the formidable thicket of difficulties that bars the path of those who pursue the Science of man (Galton, 1889, 62–63).

According to Edwin Boring (1950), Galton was the first to work out the method of statistical correlation. He was also a pioneer in the use of questionnaires and rating scales, devices that lie at the heart of the correlational psychology of religion.

Does Prayer Have Objective Effects?: Galton's most famous contribution to the psychology of religion is his study of the objective efficacy of petitionary prayer. Rather than involving himself in the difficulties and limitations of an experimental investigation, Galton (1872) chose simply to tabulate the evidence he could find around him. Did public prayers on behalf of the reigning sovereign, for health and longevity, have an evident effect? Apparently not, reports Galton, who cites statistics showing that male members of royal houses, from 1758 to 1843, were the shortest-lived among various affluent classes. The alternative hypothesis that prayers are efficacious in saving royalty from otherwise even shorter life-spans seemed to Galton rather improbable.

Statistics also demonstrated that a prayerful life gave no advantage to members of the clergy, at least the eminent ones. Galton (1869) documents this conclusion in detail in a comparative study of "Divines," based on a four-volume encyclopedia of evangelical biographies (Middleton, 1779–1786) that includes such figures as John Calvin, John Donne, Jonathan Edwards, Martin Luther, Cotton Mather, and John Wycliffe, as well as many others less well known today. From his study of these biographies, Galton draws several conclusions: (1) Divines are not founders of notably influential families, whether on the basis of wealth, social position, or abilities; (2) they tend, if anything, to have fewer children than average; (3) they are slightly less long-lived than the eminent men in Galton's other groups; and (4) they usually have wretched constitutions.

Galton gives special emphasis to the last conclusion. Many of these religious figures were sickly in their youth, a fate for which they compensated, he suggests, by turning to bookish pursuits. As adults they remained infirm, even taking pleasure in dwelling on the morbid details of their suffering. Indeed, says Galton, "There is an air of invalidism about most religious biographies" (p. 256). The majority of the few Divines who possessed vigorous constitutions had, according to Galton, been wild in their youth. Anticipating the findings of James as well as later research results, he concludes that "Robustness of constitution is antagonistic, in a very marked degree, to an extremely pious disposition" (p. 260).

In no observed respect, then, were the clergy that Galton studied exceptionally favored; on the contrary, they were among the less fortunate of the human race. That the exceptionally pious—combining, according to Galton, the unrelated trends of high moral character and instability of disposition—may pass on to their children only one of these traits accounts for the absence of extreme piety, or the presence of notoriously bad behavior, in the next generation. The slight advantage in life-span that *ordinary* clergy had, on the average, over most other professional groups can be explained, says Galton (1872, p. 129), by "the easy country life and family repose."

Galton reports a variety of other findings that likewise suggest that petitionary prayer does not have consistent objective effects. Missionaries, in spite of their worthy objectives, proved not to live longer than other people; Galton reports that many of them died, in fact, shortly after arriving in their host countries. Similarly, Galton found that the newborn children of the "pious classes" were no more likely to survive than children born to parents less inclined to pray for the well-being of their offspring. Compared with other buildings, churches were no less frequently damaged by lightning, fire, earthquakes, or avalanches. Medical doctors and insurance companies, presumably eager to discover the slightest advantage that some factor might give for health or safety, seemed nowhere to take into account the influence of prayer or pious disposition.

A Biased Selection: Both the rather unsympathetic chapter on Divines and the essay on prayer, which Galton published three years later, after it had been rejected three times because of its controversial nature, were sharply criticized (Forrest, 1974, pp. 111, 172). Perhaps his most judicious critic was Karl Pearson (1924), one of Galton's associates and his biographer, who notes the bias in Middleton's selection of men of piety. By limiting himself to Middleton's encyclopedia, Pearson observes, Galton omitted nearly all the founders and many of the central figures of the great religious movements, including Saint Augustine, Saint Thomas Aquinas, Saint Francis, Meister Eckhardt, George Fox, Emanuel Swedenborg, and John Wesley. Entirely excluded, too, are a number of influential women. Middleton's selection includes, Pearson observes, individuals eminent in piety but undistinguished in intellectual ability. During most of the years covered by this collection of biographies, many churchmen became eminent largely because they were politicians or statesmen, not because they were profoundly religious. Galton claims that practical considerations forced him to limit himself largely to persons of English nationality, or at least to persons well known in England. He excluded Roman Catholic Divines because, as celibates, they provided no data for his statistics on heredity. Thus his conclusions about Divines may be generalized strictly to male Protestant Christians who lived at a particular time and were eminent in a particular part of the world.

The Subjective Effects of Prayer: Although many of his contemporaries considered Galton a "flippant freethinker," he was in his own way a religious man: reverent, deeply committed to high ideals, and even subject to mystical intimations (Blacker, 1946). He perceived a unity within all of life and he held that "Men and all other living animals are active workers and sharers in a vastly more extended system of cosmic action than any of ourselves, much less any of them, can possibly comprehend" (Galton, 1869, p. 361). "Man," he wrote elsewhere (1908, p. 323), "is gifted with pity and other kindly feelings; he has also the power of preventing many kinds of suffering. I conceive it to fall well within his province to replace Natural Selection by other processes that are more merciful and not less effective. This is precisely the aim of Eugenics." Galton felt that the principles of EUGENICS should be embraced as though they were religious tenets.

Although prayer may not have strictly objective results within this lawful cosmos, Galton had no doubt that it can yield subjective effects, whether they include "a confident sense of communion with God" or the no less powerful feeling of solidarity with what surrounds us, a world governed by physical laws that include the hereditary ones that so interested Galton. Either outcome can "ennoble the resolves" and "give serenity during the trials of life and in the shadow of approaching death" (Galton, 1872, p. 135). Galton regularly led his own household in prayer, albeit with some reservations, and he "always made it a habit to *pray* before writing anything for publication, that there be no self-seeking in it, and perfect candour together with respect for the feelings of others" (quoted in Pearson, 1930, p. 272). With prayers of aspiration rather than solicitation he clearly had no quarrel.

The Emergence of Modern Correlational Psychology

Since Galton's day correlational psychology has become remarkably more sophisticated. The impetus for developing tests that measure individual differences at the level of complex functions came in particular from two sources: the interest of the French government at the beginning of this century in identifying and educating the mentally retarded; and, a decade or so later, the sudden need of the war-

pressed United States Army to classify in short order more than a million recruits. From an initial interest limited largely to intelligence and special abilities, psychological testing grew to encompass an enormous diversity of instruments—some commercially published, many others not—serving education, industry, clinical psychology, and scientific research (Anastasi, 1988).

Over the same period of time, Galton's relatively simple statistical approach evolved into highly complex methods, including such widely used procedures as FACTOR ANALYSIS, ANALYSIS OF VARIANCE, and a variety of other tests for evaluating statistical significance. Advances in the design of electronic computers in the 1950s rapidly made practical the use of the more tediously complex statistical methods, and they spurred on the development of ever more elaborate quantitative theories and procedures. Drawing on these new statistical methods as well as modern principles of test construction and validation, contemporary correlational psychologists have continued Galton's exploration of individual differences, including differences in religiousness.

THE PROBLEM OF MEASURING RELIGION

The first major task of the correlational psychologists of religion is to develop precise and reliable means for assessing individual piety. Research on religious behavior and experience, they maintain, requires an "operational definition" of these events. That is, the researcher must specify precisely the operations or procedures that will be used to observe them. Because a person can be religious to varying degrees, these procedures are typically quantitative, yielding scores that can be correlated statistically with a range of other dimensions. The search among personality characteristics, attitudes, and other variables for correlates of religiousness is the second major task of these correlational investigators (see Hood, Spilka, Hunsberger, and Gorsuch, 1996).

In estimating the piety of a new family in the neighborhood, for example, we would ordinarily resort to a predictable series of observations. Are they members of a religious organization? If so, how often do they attend religious services and other functions? Have they furnished their house with religious pictures, statues, or other sacred objects? Are there religious books or magazines lying about? Do they bow their heads before meals, cross themselves, or engage in other obvious ritual behavior? In asking questions directed at observable and even quantifiable behavior, we are conducting ourselves much like the correlational psychologists, although our observations are likely to be less systematic and precise.

As most social scientists will acknowledge, such indicators of piety are crude at best, no matter how carefully the observations are made. Persons may be active in religious organizations, not primarily as an expression of religious faith, but as a means of gaining friendship, making business contacts, or expressing their socioeconomic status and values. Sacred objects around the house may carry little religious meaning for its occupants, serving rather as sources of aesthetic pleasure or nostalgic links to the past. The books and magazines, some of them perhaps gifts from persons outside the family, may be largely unread. Pious gestures may occur out of sheer habit or to fulfill the expectations of others. The absence of these indicators is likewise ambiguous: some profoundly religious people may exhibit none of them.

Although psychologists and sociologists of religion employ similarly rough-and-

ready measures of religiousness (see Argyle and Beit-Hallahmi, 1975), many of these researchers have sought more subtle means of assessment. On the one hand, they recognize the ambiguity of any outward expression and therefore seek to develop questionnaires that also evaluate the less obvious underlying beliefs and attitudes. On the other hand, many are convinced that these indicators are not parallel measures of a single dimension, religiousness, but that they reflect, rather, several dimensions of piety varying more or less independently of each other.

Early Measures of the Beliefs of Scientists

The earliest questionnaires were relatively long and complex, inviting detailed answers that could not easily be quantified (see Lehmann, 1915; Pratt, 1907; Starbuck, 1899). Leuba's (1916, 1934) studies of the beliefs of American scientists are a famous exception. In 1914 and again in 1933, Leuba sent out a brief questionnaire to a large number of scientists chosen at random from the current edition of *American Men of Science* and the directories of the American Sociological Society and the American Psychological Association. Leuba included in his samples a number of "greater" contributors who had been identified as eminent in their fields by a small group of distinguished peers. Recipients of the questionnaire were asked to indicate whether they believed "in a God to whom one may pray in the expectation of receiving an answer. By 'answer' I mean more than the natural, subjective, psychological effect of prayer." As alternatives they were offered "I do not believe in a God as defined above" or "I have no definite belief regarding this question." Three corresponding statements regarding immortality were also presented (Leuba, 1916, p. 223).

Leuba's findings, summarized in Table 6.1, indicate that the probability of belief declines (1) the more knowledge the scientist has about "living matter, society, and the mind," (2) the greater the scientist's peer-rated eminence, and (3) the more recent the response to Leuba's questionnaire. Here is evidence, Leuba (1950) concludes, of a revolution in thought and a readiness for a transformation of the churches into liberal religious societies that recognize natural spiritual forces and promote "spiritual hygiene and culture" (p. 124). A follow-up study carried out more than two decades later by Ronald Mayer (1959), using Leuba's scales

Table 6.1 **PERCENTAGES OF AMERICAN SCIENTISTS AFFIRMING BELIEF IN GOD AND IMMORTALITY, 1914 AND 1933**

	Lesser Scientists		Greater Scientists	
	1914	1933	1914	1933
A. Belief in a God Who Answers Prayer				
Physical scientists	50	43	34	17
Biological scientists	39	31	17	12
Sociologists	29	30	19	13
Psychologists	32	13	13	12
B. Belief in Personal Survival after Death				
Physical scientists	57	46	40	20
Biological scientists	45	32	25	15
Sociologists	52	31	27	10
Psychologists	27	12	9	2

Source: From Leuba, 1934, page 297.

along with two others, confirmed most of the same trends, although Mayer did not find significant differences among disciplines for the eminent scientists.

As we might anticipate, many of Leuba's subjects objected to the limitations imposed by his questions, often adding in marginalia or separate letters the qualifications or alternatives they felt were needed to represent their own views. One wrote, for example, "I do believe in a God and in prayer, but *not as you have outlined it*" (Leuba, 1916, p. 235). Of course, any questionnaire may inspire such dissent. Even the more open-ended set of questions sent by C. L. Drawbridge (1932) and his associates to the fellows of the Royal Society drew its share of protests. A surer way of representing the individual views of scientists was employed by Edward Long (1952), who turned to published works—many of them obscure—for evidence of religious outlook. What he demonstrates, however, is the possibility of finding scientists with conservative religious views, not the validity of his conclusion that among scientists "the lack of pattern is obvious" (p. 145). Having limited himself to scientists who had published their religious views, and then having chosen which of them to include, Long has no basis for generalizing about scientists as a whole.

For such generalizations, a simple device such as Leuba's questionnaire, combined with an appropriate sampling technique, is far more serviceable. Yet it is possible to simplify matters still further. Rather than sending out a questionnaire, Harvey Lehman and Paul Witty (1931) used *American Men of Science* and *Who's Who in America* in combination to establish the frequency and nature of religious affiliation among eminent scientists. They found that these distinguished researchers acknowledged a denominational affiliation half as often as all others listed in *Who's Who* (25 percent as compared with about 50 percent, a figure close to the level of affiliation in the general population of the United States at that time [Argyle, 1958]). More revealing was the discovery that, among these scientists, the liberal denominations, such as the Unitarians, Congregationalists, and Episcopalians, were markedly overrepresented, and the conservative groups, including the Baptists, Lutherans, and above all the Roman Catholics, were strikingly underrepresented. Lehman and Witty speculate that these findings are symptomatic of the conflict between conservative religious doctrine and scientific thought.

The Thurstone Scales of Religious Attitudes

In the late 1920s and early 1930s, as the psychologist's measurement techniques became more sophisticated, so also did the questionnaire of the psychologist of religion. When the eminent University of Chicago psychologist L. L. Thurstone (1887–1955) made his famous adaptation of psychophysical methods—developed for studying the relation of physical stimuli and bodily sensation—to the quantification of social values, he chose as his illustrative case the Scale for Measuring Attitude Toward the Church.

With an initial list of 130 carefully edited statements derived from solicited opinions and published literature, Thurstone and Ernest Chave (1929) asked 341 subjects to sort slips of paper bearing these brief statements according to how appreciative each appeared to be. The subjects were *not* to respond to them in terms of their own opinions about the church. The arrangements of the statements into 11 piles, ranging from "highest appreciation" to "strongest depreciation," gave these researchers an objective means of determining each statement's *ambiguity* (as reflected in the variability of ratings) and its *scale value,* equivalent to the median ranking of the 300 subjects whose sortings appeared to be properly carried out. A second group of 300 subjects was asked to endorse the statements that

expressed their own sentiments. Their responses provided a basis for determining the consistency with which each item was answered in relation to the others, a measure that was used to estimate an item's *relevance.*

The final set of 45 items that compose the published questionnaire (reprinted in Chave, 1939) was chosen to minimize ambiguity and irrelevance, as well as to distribute the median rankings of the items more or less uniformly along the entire scale. This elaborate procedure created an attitude scale consisting of "equal-appearing intervals." The questionnaire is scored by calculating the average scale value (or, alternatively, the average rank order) of all the statements an individual has endorsed. Two estimates of the RELIABILITY, calculated by the split-half method (see Anastasi, 1988), yielded values of .92 and .94—satisfactory evidence by any standards that the test's scores are highly consistent. A striking correlation with reported church attendance provides some evidence of the scale's VALIDITY.

Using similar procedures, Chave and Thurstone constructed and published several other scales for use by psychologists of religion. These include five forms of the Scale of Attitude Toward God, which assesses individual conceptions of God, belief in the reality of God, and the degree to which belief in or ideas of God influence conduct, and two forms of the Scale of Attitude Toward the Bible, on which scores may be interpreted in a range from "Strong prejudice against the Bible" to "Strong belief [in] and devotion to the Bible" (Chave, 1939). Two forms of yet another scale, on Sunday observance, were prepared by Charles Wang and Thurstone (reprinted, with the church scale, in Shaw and Wright, 1967). The particular content of these scales was contributed by Thurstone's collaborators, not Thurstone himself, who, though a Lutheran minister's son, was apparently less than enthusiastic about orthodox observance even as a child. His short-lived interest in attitude measurement was itself largely derivative, serving his wish "to introduce some life and interest in psychophysics" (Thurstone, 1952, p. 310). He left it to others to use and evaluate these scales, of which the Attitude Toward the Church scale has been most widely employed.

What advantages do these painstakingly derived tests have over the far simpler indicators that Leuba or Lehman and Witty used? Could the researcher not just as well ask the subjects to rate themselves on a continuum labeled at one end "Strongly favorable to the church" and at the other end, "Strongly against the church," with the midpoint designated "Neutral"? Such is precisely what Thurstone and Chave did. Converting this self-rating line into a simple ten-point scale, they correlated it with their more elaborate measurement device. The outcome, a correlation of .67, indicates that the two measures have approximately 45 percent of their VARIANCE in common (see Nunnally, 1978)—a sizable amount, by the standards of the social sciences, but still less than half. Moreover, from examining individual questionnaires, the authors inferred that the subjects tended to evaluate themselves on the self-rating line as more favorably disposed to the church than is indicated by the scale statements they endorse. Thurstone and Chave are properly cautious in interpreting these observations, yet it would seem safe to conclude that the brief measure cannot serve as a substitute for the longer measure without a change in meaning.

Brevity is not necessarily a virtue, of course, as we saw in Leuba's research. Too brief a questionnaire may leave subjects doubtful that their views are properly represented. Too long a questionnaire, on the other hand, runs the risk of losing the subject's sustained cooperation. From a PSYCHOMETRIC perspective, the extended questionnaire is the more valuable one. Other things being equal, the more

items a psychological test has, the more reliable it will be (although a point of diminishing returns will soon be reached). A variety of questions also allows a far more adequate sampling of the domain to be measured, and the scores they yield, because they range more widely than those of brief questionnaires, permit finer discriminations among the persons answering them. The broader range of scores correspondingly increases the possibility of finding significant correlations with other measures.

FACTOR ANALYSIS AND THE DIMENSIONALITY OF RELIGION

A longer questionnaire also invites evaluation through FACTOR ANALYSIS, a term that refers to a group of elaborate statistical methods for which we are again indebted to Thurstone, among others. Although itself complex, factor analysis is employed as a means of simplification. It is used to find the few "factors" that account for the pattern of correlation among a large number of variables—in the present case, usually the individual items of a questionnaire.

Factor analysis begins with a matrix of correlation coefficients, which indicate the degree of relationship of every questionnaire item with each of the others. Correlation coefficients range from +1.00, which indicates a perfect *positive* relation between two variables, to −1.00, which designates a perfect *negative* (or inverse) relation. The closer the correlation coefficient is to zero, the more nearly independent of each other the two variables are. Thus if two statements on a questionnaire showed a correlation of .68, we would conclude that respondents who agree with one of these statements are likely to agree with the other. If the correlation were −.68, on the other hand, agreement with one statement would signal the likelihood of disagreement with the other. If the correlation coefficient were close to zero—say, .06—we would conclude that the statements are unrelated to each other.

During the process of factor analysis, the computer scans the table of correlation coefficients and extracts in succession the small number of factors that account for the interrelationships among the items. In a new table, then, the computer indicates the "loading," or correlation, of every questionnaire item with each of the extracted factors. The researcher will note which items load most heavily on each factor and, by studying their content, infer what the factors are. Identifying factors is rarely easy, and coming up with succinct yet accurate labels is likewise a challenge. Also of interest in the table of factor loadings are the numbers indicating how much variance, or variability, in the test scores is accounted for by each of the factors. In every instance of factor analysis, some variance will remain unexplained by any of the factors.

It is routine today to factor-analyze any questionnaire that consists of more than a few questions. Researchers developing complex inventories with multiple scales frequently use factor analysis as a way of checking to be sure that the scales possess the statistical coherence their content suggests they should have. Scale items that do not show an appropriate pattern of factor loadings—relatively high on the factor they are intended to measure and correspondingly low on all the others—can be either rewritten or replaced. It is also common to factor-analyze scales written to measure only one variable, especially if the investigator suspects that the variable has two or more facets and thus is factorially complex. A number of religiosity scales have been factor-analyzed to check out their suspected complexity—and, by extension, the complexity of human piety.

Religion as a Single Factor

Like other researchers, psychologists and sociologists of religion do not agree about the factorial complexity of the object of their study. Some continue to assume that religiousness is a unidimensional quality that can be adequately assessed by any number of indicators. There is some evidence supporting this position, most of it obtained through factor analysis.

The earliest factor-analytic study providing such support was carried out by Thelma Thurstone, who administered 11 of her husband's attitude scales to several hundred university students. She found that the church, God, and Sunday observance scales clustered together on the far end of one factor, a dimension identifiable as either conservatism or as religion (Thurstone, 1934). The scales measuring positive attitudes toward evolution and birth control loaded negative on the same factor, just as they did in investigations by Leonard Ferguson (1939, 1944a).

A variety of studies, drawing on disparate materials, have confirmed these early findings. Each of three factor-analytic evaluations of test items written to represent Eduard Spranger's (1914) six personality types found a single religion-related factor (Brogden, 1952; Gordon, 1972; Lurie, 1937). The longer and more homogeneous scale of attitudes toward the church and religion administered by Adolf Holl and Gerhard Fischer (1968) to a group of Austrian soldiers similarly yielded a single religiosity factor. Two other factors, Distance from the Church as Organization and Social Contact, although correlated with church attendance, were interpreted as nonreligious determinants. Unidimensionality is also confirmed in a series of studies by L. B. Brown (1962, 1966) and A. J. Wearing and Brown (1972), who consistently found a single factor regardless of the combination of measures they employed.

The Case for Multidimensionality

Most psychologists and sociologists of religion favor a multidimensional perspective. So also do other scholars. Well before the advent of factor analysis, the Roman Catholic lay theologian Baron Friedrich von Hügel (1908) identified three "elements of religion," three successively developing modes of apprehension or outlook: (1) the *traditional,* or *historical,* dependent largely on the senses, imagination, and memory, and originating in childhood; (2) the *rational,* or *systematic,* emerging with the capacities for reflection, argument, and abstraction; and (3) the *intuitive* and *volitional,* signaling the maturation of inner experience and outer action. James Pratt (1920) takes a similar approach, although he separates the two aspects composing von Hügel's third element in his own four "typical aspects of religion": the traditional, the rational, the mystical, and the practical or moral. Both von Hügel and Pratt argue that each element or aspect, however much it may be emphasized in a particular case and to whatever degree it may conflict with the others, remains fundamentally bound up with the rest.

Yet another variation on the facets of religion is offered by sociologist Charles Glock (1962), whose *experiential* and *consequential* dimensions closely parallel Pratt's mystical and practical aspects, whereas his *ritualistic, ideological,* and *intellectual* (that is, knowledge) dimensions seem largely to expand von Hügel's and Pratt's traditional category and touch little if at all on the rational (see Table 6.2). Although Glock thinks "it scarcely plausible" that the five dimensions he discerns are wholly independent of one another, he leaves the question of their interrelation to quantitative research.

Religion in Five Dimensions: The task of operationalizing Glock's dimensions

Table 6.2 ASPECTS OF RELIGION ACCORDING TO VON HÜGEL, PRATT, AND GLOCK

von Hügel	Pratt	Glock
1. Traditional (historical)	1. Traditional	1. Ideological (belief) 2. Ritualistic (practice) 3. Intellectual (knowledge)
2. Rational	2. Rational	
3. Intuitive and volitional	3. Mystical 4. Practical (moral)	4. Experiential (feeling) 5. Consequential (effects)

was taken up by Joseph Faulkner and Gordon DeJong (1966), who painstakingly developed five scales of four or five items each, employing the cumulative scaling technique originated by Louis Guttman (see Nunnally, 1978; Shaw and Wright, 1967). For their subject group as a whole, the intercorrelations among the five scales and thus presumably among the five dimensions ranged from a high of .58 (between the Ideological and Intellectual scales) to a low of .36 (between the Experiential and Consequential scales). When the subject group was broken down by sex, religious affiliation, or parents' church membership, some correlations ranged slightly higher and others considerably lower. These researchers found, for example, that the Ideological and Experiential scales correlated more highly for females than for males (.60 versus .39); that the Jewish students showed the lowest correlation by far between the Ritual and Consequential scales; and that half of the ten correlations among the five scales were statistically insignificant for students whose parents were not church members. Although most of the scale intercorrelations *are* statistically significant for the total group of 375 students, they appear at the same time to be low enough to support the multidimensional position.

The response both to Glock's proposed dimensions and to Faulkner and DeJong's efforts to measure them might be taken as a case study in the complexities of this area of research. Rodney Stark and Glock (1968) themselves had operationalized a slightly modified set of Glock's dimensions, using a rather different collection of items. Nonetheless, the pattern of intercorrelations they found was "extremely similar" to Faulkner and DeJong's (p. 179). They also discovered that the ideological commitment or Orthodoxy scale, because of its relatively high correlations with the others, is the best single measure of the five. Andrew Weigert and Darwin Thomas (1969), unconvinced by these apparent convergences, argue that the Faulkner–DeJong scales are inadequate measures of Glock's dimensions. Three of the four items constituting the Intellectual or knowledge scale, they point out, actually measure belief, and only one of the five items composing the Experiential scale assesses the individual's own religious emotion. Most of the 23 items might be described as ideological in format, so that the preeminence of the Ideological scale is no surprise. Similar criticisms are made by James Gibbs and Kelly Crader (1970), who fault the Experiential scale on the same grounds as Weigert and Thomas and argue that the items on the Consequential scale are contaminated by the religious context in which they are framed. Yet the data they collected using revised Experiential and Consequential scales "tend to support the pattern of relationships reported by Faulkner and DeJong" (p. 111).

In their reply to Weigert and Thomas, Faulkner and DeJong (1969) emphasize the difficulty of finding items of religious knowledge that form a proper scale, and they remind their critics of the breadth of Glock's original characterization of the

experiential side of religion. They do not justify, however, their counting of literal belief in the story of creation or biblical miracles as positive evidence of "an intellectual orientation toward one's faith." Nor do they acknowledge that for only one of their five experiential items does an affirmative answer unequivocally indicate that the respondent has had the experience in question. Whatever their items do assess, they are correct in pointing out that the relatively low intercorrelations are evidence that the five subscales do not measure the same thing.

Challenges to the 5-D Scales: Most criticisms of Glock's dimensions and the Faulkner–DeJong "5-D" scales are based on patterns of statistical relationship rather than on content analysis of questionnaire items. Although some researchers have also found low scale intercorrelations (e.g., Kuhre, 1971; Lehman and Shriver, 1968; Ruppel, 1969), others have discovered intercorrelations sufficiently high to call into question the assumption of multidimensionality (see Cardwell, 1969; Clayton, 1971a; Finner and Gamache, 1969; Gibbs and Crader, 1970; Rohrbaugh and Jessor, 1975). Several factor-analytic studies make the challenge especially clear. Richard Clayton and James Gladden (1974) factor-analyzed the responses of two groups of largely Protestant university students to the 5-D scales and found five and four factors, respectively. In both instances, however, only the first factor—Ideological Commitment—is clearly defined. This factor accounts for an astonishing 78 percent of the common variance in one group of subjects and 83 percent in the other, a finding that demonstrates that it is by far the most important factor. A second-order factor analysis of the factors themselves produced one general factor. "Religiosity," they conclude, "is essentially a single-dimensional phenomenon composed primarily of Ideological Commitment," the strength of which is reflected in experience and practice (p. 141).

Ursula Boos-Nünning's (1972) investigation in West Germany yielded findings that likewise raise doubts about Glock's frequently cited formulation. When Boos-Nünning factor-analyzed the responses of randomly selected Catholics in a large city of the Ruhr Valley to a 78-item questionnaire designed to measure six dimensions of religiosity, including Glock's five, only two of the six appeared as factors in her results: Religious Knowledge and Tie to the Parish. Most of the items representing Glock's ritualistic, ideological, experiential, and consequential dimensions loaded on Boos-Nünning's first factor, called General Religiosity, which accounts for 51 percent of the common variance. The second of her six factors, Tie to the Parish (or Parish Communication and Information, as she came to call it), accounted for 18 percent, and the third, Marital and Sexual Morality, accounted for only 9 percent. The last three—Belief in God, Public Religious Practice, and Religious Knowledge—were responsible for only 7 or 8 percent each.

Glock's dimensions were also put to the test by two other German researchers, Albert Fuchs and Reinhard Oppermann (1975). They factor-analyzed similarity ratings for 16 of the most frequently used religious words in the German language, such as *Glaube* (faith), *Gott* (God), *Engel* (angel), *Predigt* (sermon), and *Gnade* (grace), which were presented to their heterogeneous sample in paired combinations. Their two factors, accounting, respectively, for 53 and 17 percent of the total variance, confirm at best only the ideological and ritualistic dimensions of Glock's schema. Once again, therefore, the evidence argues against the complexity hypothesis. It tends, rather, to support the notion of one major factor, either general in character or ideological in emphasis.

The Multiplicity of Evidence: What, then, are we to make of the numerous factor-analytic studies that report two or more major factors largely independent

of each other? William Broen (1967), for example, found that the pattern of responses of 24 religious subjects to 133 statements yielded two factors: Nearness to God, reflecting a sense of "the Deity's loving presence and guidance," and Fundamentalism–Humanitarianism, a bipolar factor emphasizing at the "positive" end humankind's sinfulness and need for a punishing God and at the "negative" end the potential goodness and self-sufficiency of human beings. Broen found a high enough correlation (.32) between these two factors to allow him to combine them as a measure of "general religiosity," yet one sufficiently low to support his argument for multidimensionality.

At the other extreme of factorial complexity is the work of Morton King and Richard Hunt (1969, 1972a, 1972b, 1975a, 1990; King, 1967). Their basic set of items, derived from other researchers' work as well as from interviews and questionnaire replies, was assembled as an instrument for measuring 11 hypothetical dimensions. The responses of members of 27 congregations representing four mainline Protestant denominations in Dallas County, Texas, yielded—in spite of modifications in item composition made for each denomination—a remarkably stable set of ten religious factors: (1) Creedal Assent, (2) Devotionalism, (3) Church Attendance, (4) Organizational Activity, (5) Financial Support, (6) Religious Knowledge, (7) Orientation to Growth and Striving, (8) Extrinsic Orientation, (9) Salience: Behavior (earlier called Talking and Reading about Religion), and (10) Salience: Cognition (indicating the prominence of religion in everyday thoughts and feelings).

Items selected to measure eight of these factors were included, with others, in a questionnaire distributed to a nationwide sample of members of the United Presbyterian Church in the U.S.A. The entire questionnaire was then factor-analyzed (King and Hunt, 1975a). Among the 20 interpretable factors that were found, all eight of the King and Hunt factors reappeared; moreover, their item compositions were almost identical to those found in earlier investigations. Although nearly all the correlations among the King–Hunt scales are statistically significant (King and Hunt, 1972b, p. 119), the relatively low average of these correlations (.37) supports the case for multidimensionality. Group differences on these scales, as well as correlations with other measures, demonstrate the potential usefulness of so complex and thorough a measuring instrument.

Evidence for multidimensionality continues to accumulate. Starting with a new religiosity questionnaire rather more sophisticated than earlier instruments and administering it to both American and German students, DeJong, Faulkner, and Rex Warland (1976) found six familiar dimensions. Yet they also report evidence of "generic religiosity," encompassing belief, experience, and practice. Using the same questionnaire, Dale Hilty and Sue Stockman (1986) replicated five of the six factors as well as the higher-order generic religiosity. Various other scales intended as general measures of religion have upon analysis yielded two or more factors (Cline and Richards, 1965; Himmelfarb, 1975; Maranell, 1968; Shand, 1961; Tapp, 1971).

More specialized scales have proved to be factorially complex as well. Robert Coursey (1974) found six factors in a scale designed to measure the liberal–conservative dimensions among Roman Catholics. Leon Gorlow and Harold Schroeder (1968), Robert Monaghan (1967), and Sam Webb (1965) investigated the motives or needs underlying church attendance or religious activities more generally and report, respectively, seven, three, and eleven factors. In a succession of studies using many of the same adjectives, Bernard Spilka, Philip Armatas, and June Nussbaum

(1964), Richard Gorsuch (1968), and Walter Broughton (1975) found three or more major dimensions in the conceptualization of God, with notable convergence of factors, whereas Godelieve Vercruysse (1972), employing descriptive phrases, found six such factors with his adolescent subjects and four rather different factors with his adults. Finally, Ralph Hood's 32-item Mysticism Scale, which he derived from Stace's (1960) analysis, yielded two factors—General Mysticism and Religious Interpretation—when factor-analyzed by Hood (1975) and three when analyzed by Dale Caird (1988) using college students and also by Duane Reinert and Kenneth Stifler (1993) using three groups of older adults varying widely in religious background and psychiatric status. In both of the latter studies, Hood's second factor split in two: a noetic (Caird) or noetic/ineffability (Reinert and Stifler) factor, accenting the experience of insight if not also the indescribableness of the experience, and a religious factor, emphasizing the sense of the holy.

Support for the multidimensional position can also be garnered from studies reporting relatively low intercorrelations among different measures of religiousness (e.g., Finner, 1970; Fukuyama, 1961; Lenski, 1961; Vernon, 1962). The difficulty here is deciding how small such correlations must be, given the inevitable imperfections of these scales, to argue that they are measuring different dimensions.

With evidence and advocates on both sides of the dimensionality issue, and what may seem to be a blur of factors identified by those in the multidimensional camp, what firm conclusions can we hope to draw? For those who assume that factor analysis of religiosity scales is a way of uncovering an unchanging truth about all human beings, regardless of time or place, the diversity of outcomes may be interpreted as a disheartening yet temporary state of affairs, ultimately resolvable through further refinements in the scales and statistical procedures. For those of a postmodern perspective, on the other hand, the scales, procedures, and factors are all fallible human constructions—useful for certain purposes, no doubt, but not mirrors that reflect a psychological reality unchanged by context or history. What they do reveal is how social scientists think about religion today, and what their subjects make of it as well.

Factor analysis, all will agree, is a sophisticated procedure for discerning a pattern within a particular set of data. What the researcher includes among these data will determine the configuration of factors reported by the computer. If we assemble a questionnaire composed only in part of religious items, all of which express conventional attitudes or practices, and then give this questionnaire to a heterogeneous sample of persons, varying especially in religious background and commitment, we may be fairly confident that the religious items will cluster together to form a single "religious factor." If, by contrast, the questionnaire consists entirely of religious items, varied and subtle in content, and it is given to a relatively homogeneous and religiously sophisticated group of persons—members of the clergy, for example, such as Jack Shand (1961) employed—then we may naturally expect a range of religious factors. That religious "insiders" may provide a more differentiated factor structure than "outsiders" is demonstrated in Gary Maranell and Nevell Razak's (1970) comparison of priests and ministers with college professors.

Religion's Singular Complexity

No one has yet found *the* fundamental dimensions of religiousness—nor will they ever be found. Scales of any kind should be understood as conventional devices that serve the particular needs of the researchers (King and Hunt, 1975b, 1990;

Nunnally, 1978). Although Dittes (1969) argues cogently that the multidimensional approach is likely to be the more fruitful one for the social scientist, he also takes seriously the evidence for a single common factor, suggesting that it "be construed simply as 'religion as seen by the general population' " (p. 618). "The resolution," writes Gorsuch (1984), "could be *both/and* rather than either/or" (p. 232). As he demonstrates by reference to Maranell's (1974) results, a single, general factor of religiosity may relate well to other broadly conceived variables (e.g., age), whereas more specific religious factors may serve best to identify exceptions to these general trends.

The paradox we face here—that religion may be many things and yet one, a *unitas multiplex*—was addressed by von Hügel (1908) long before factor analysis and the dimensionality issue emerged. If religion were a mere multiplicity, a collection of parts without a whole, he says, it would be neither persuasive nor effective; yet if it were a simple unity, a whole without parts, it could not be a source of truth (p. 50). It may be, however, that both von Hügel and contemporary researchers are misled by the word *religion.* As we noted in Chapter 1, this word has become a reified abstraction that misleadingly suggests an unchanging essence, either within each tradition or underlying them all. If we substitute Wilfred Smith's (1963) preferred terms, cumulative tradition and faith, we immediately see that multiplicity is the inescapable fact, overwhelmingly so in the case of the historic traditions but unmistakable, too, in the realm of faith. It is personal faith that researchers aspire to measure, and it will always be expressed in a diversity of individual forms.

No set of scales, however complex, can hope to represent the nuances of personal faith. Every scale is a compromise: its statements must be general enough to be usable by a range of people, yet they must be specific enough to distinguish among the respondents. For example, the statement "I believe in God" would in the United States be too general, for around 95 percent of respondents are likely to agree. On the other hand, the declaration "I am regularly visited by an invisible Presence to whom I can pour out my heart," although crucial for representing the experience of Flournoy's (1915) mystic, would be either too specific for a general questionnaire or too easily reinterpreted to include any vague sense of a spiritual presence.

Some researchers have aspired to develop questionnaires sufficiently broad to be usable across religious traditions (Bhushan, 1970, 1971; Yinger, 1969, 1977); others, daunted by the obvious difficulties of encompassing so wide a range of concepts and practices, propose the development of parallel scales for the various traditions, including in each the specific expressions shared by persons of similar faith (King and Hunt, 1972b). Glock (1962) clearly has a program of this breadth in mind, but only after we have found adequate means for assessing religious commitment in our own culture.

Most existing religiosity scales have been developed for use by Christian subjects. Yet even more broadly conceived questionnaires may not provide sufficient options for thoughtful, imaginative, and religiously complex subjects. Noting that it has long been typical to provide respondents with only two options in relation to creedal statements—literal affirmation and literal disaffirmation—Hunt (1972a) has proposed adding a mythological alternative. For example, to the statement "Jesus was born of the Virgin in a manner different from human beings," a subject may give one of three responses: (1) a literal one: "Agree, since God conceived Jesus in Mary's womb before she had sexual relations with Joseph, her husband"; (2) an antiliteral one: "Disagree, although most religions claim a virgin birth for

their founder, we know that such an event is physically impossible"; and (3) a mythological one: "Agree, but only in the sense that this is an ancient mythological way of talking about the Ultimate Reality as manifested in Jesus" (Hunt, 1972a, p. 49).

Although Andrew Greeley (1972) asserts that Hunt's contribution "ought to mark a decisive turning point in religious research," he urges the inclusion of a fourth option, a "HERMENEUTIC dimension," which would more clearly and consistently retain the transcendental referent. Thus on the question of the virgin birth, respondents might agree that "The Ultimate Reality was present in Jesus in a way decisively different from the way It is present in the rest of us." The LAM (Literal, Antiliteral, and Mythological) scales, replies Hunt (1972b), in themselves represent three hermeneutical perspectives; a fourth would increase the options but not exhaust them. It is a question, therefore, of how many and which perspectives to incorporate.

In his research with Hunt's LAM scales, Norman Poythress (1975) did find a sizable group of people—28.2 percent of 234 college undergraduates—who could be classified as mythological types and thus as having a proreligious but nonliteral outlook. But contrary to Hunt's expectations, the Mythological scale did not show a distinctive pattern of relations with personality variables (in this study, intelligence, authoritarianism, and racial prejudice). Using a revised, Dutch version of the scales, in which metaphorical—and thus more clearly symbolical—alternatives were substituted for the mythological ones, Jan van der Lans (1991) found that nursing students who preferred the metaphorical alternatives gave a larger number of associations to a series of photographs and thus have, van der Lans infers, a greater tendency toward imaginative thinking. Associates of van der Lans have prepared an English version of his LAM scales for use with Hindu subjects and a Dutch version for research with Muslims.

The challenge of representing the more sophisticated, symbolic understandings of religion that have emerged in the postmodern world has also been taken up by Dirk Hutsebaut (1996a). Guided primarily by philosopher Paul Ricoeur's notion of the second naïveté—a new, more sophisticated openness to the objects of religious faith that is made possible through prior critical reinterpretation—Hutsebaut and a group of seminar students set about to formulate questionnaire items that measure this religious attitude, along with others represented in the summary schema of this book (see page 635). Factor analysis yielded three dimensions or attitudes, which correspond to the first, third, and fourth quadrants of the schema (the atheistic stance of literal disaffirmation, represented by the second quadrant, did not appear in their data). The attitude of *orthodoxy* is characterized by a predominance of literal thinking, a high level of religious certainty, positive god images, and a tendency toward anxiety and guilt. The *external critique* attitude is marked by religious ambivalence, a preference for the objective certainties of science, and expressions of revolt against God. The attitude of *historical relativism,* finally, recognizes the historically conditioned character of religious statements, interprets the Bible symbolically, and is open to complexity and uncertainty.

THE QUEST FOR THE CORRELATES OF RELIGION

Even as correlational psychologists labor to create more adequate ways of operationalizing religion, they are moving forward with their main task: identifying the

psychological, social, and other correlates of religious behaviors, experiences, and attitudes. Initially, the finding of correlates may help to validate the measures of religion being used. Yoshio Fukuyama (1961), for example, found that his measures of four of Glock's dimensions showed meaningful patterns of relation to age, sex, education, socioeconomic status, and selected attitudes. Such correlations provide evidence that his scales are assessing what they are supposed to.

With the validity of their measures reasonably well established, researchers may then proceed to test a variety of hypotheses, some derived from theories about the causal origins of individual piety, others pertaining to its consequences. As we noted in the last chapter, experimental procedures are far superior for establishing causal relations, yet they are difficult to apply in the realm of religion. Thus correlational methods may be cautiously used in their place. Much correlational research, it may be said, has been undertaken without the support of a well-developed theoretical or conceptual framework. Some is carried out with no particular hypothesis or expectation in mind. Although psychological methodologists are frequently disdainful of such DUSTBOWL EMPIRICISM, it can provoke helpful new thinking and research.

The literature making up the correlational approach to religion is today enormous and still growing rapidly. Books taking stock of this literature as it existed around 1970 (Argyle and Beit-Hallahmi, 1975; Scobie, 1975; Strommen, 1971) were already filled with references to research on a large diversity of topics. More recent book-length overviews document what even insiders find to be a bewildering number and variety of research findings, a fair number of which contradict each other (Batson, Schoenrade, and Ventis, 1993; Grom, 1992; Hood, Spilka, Hunsberger, and Gorsuch, 1996; Paloutzian, 1996). Essays on the problems of measuring religion have become more common (Gorsuch, 1984, 1990; Haub, 1992; Williams, 1994) at the same time that literature reviews on specific topics, especially mental health or well-being, have multiplied (e.g., Bergin, 1983; L. B. Brown, 1994b; Corveleyn and Lietaer, 1994; Gartner, Larson, and Allen, 1991; Koenig, 1990, 1993; Payne, Bergin, Bielema, and Jenkins, 1991; Paloutzian and Kirkpatrick, 1995; Schmitz, 1992; Schumaker, 1992; Shafranske, 1996). Such reviews are sufficiently numerous, in fact, to have become themselves the subject of critical analysis (Larson, Sherrill, and Lyons, 1994).

RELIGION AND SOCIAL ATTITUDES

The complications facing correlational psychologists of religion are well illustrated in an area of research that early attracted the attention of social scientists: the relation of piety to social attitudes and behavior. For many, including William James, religion must finally be evaluated by such fruits. Yet according to this criterion, say some of its sharpest critics, religion has been mainly a disaster. Morris Cohen (1946), for instance, in scanning the history of the Christian, Jewish, and Islamic traditions, is struck primarily by the "fierce intolerance," the self-serving disregard for truth, the "pretended certainties" that prevent "needed change and cause tension and violent reaction." In making a "virtue of cruelty" and a "duty of hatred," religion has proved itself "effective for evil," not good. It is a fact, he says, "that there is not a single loathsome human practice that has not at some time or other been regarded as a religious duty" (pp. 351–352).

Philosopher Alfred North Whitehead (1926) is more explicit about some of religion's harmful effects.

in 1997, the mass suicide by all 37 members of a group known as Heaven's Gate, in anticipation of being taken to a higher evolutionary level by a heavenly space craft thought to be following the Hale-Bopp comet.

> History, down to the present day, is a melancholy record of the horrors which can attend religion: human sacrifice, and in particular the slaughter of children, cannibalism, sensual orgies, abject superstition, hatred as between races, the maintenance of degrading customs, hysteria, bigotry, can all be laid at its charge. Religion is the last refuge of human savagery. The uncritical association of religion with goodness is directly negatived by plain facts. Religion can be, and has been, the main instrument for progress. But if we survey the whole race, we must pronounce that generally it has not been so (pp. 37–38).

Religion's Dark Side Today

While it is tempting to assume that the "dark side of religion," as Cohen calls it, is mostly a thing of the past, today's newspapers regularly remind us of its continuing existence. In some cases the effects are dramatic and lethal: in 1978, the suicide of 913 followers of Jim Jones in the jungles of Guyana; in 1990, the death of more than 2000 people when Hindu fundamentalists in Ayodhya, a small town in India, tore down a sixteenth-century mosque that they believed was built on the birthsite of Lord Rama; in 1993, the death of six people and the injury of more than a thousand in the bombing of the World Trade Center in New York City by individuals associated with Islamic fundamentalism; later the same year, the death by fire and bullets of 89 adults and children in the well-fortified Waco, Texas, compound of the Branch Davidians, whose leader, David Koresh, claimed to be the son of God; in 1994, the mass suicide-murder of 53 members of the doomsday cult Order of the Solar Temple, located in Switzerland and Canada, and then 16 more, including three children, the following year; in 1995, the gassing of Tokyo subway stations by members of a Buddhist sect, resulting in the death of 11 people and the sickening of ~~four thousand~~ 5500 others; and again and again in the 1990s, the murderous suicide bombings in Israel by dozens of young Palestinians who believed that such acts of "martyrdom" would win for them—and their friends and relatives—unimaginable physical and spiritual bliss in heaven.

Less singularly dramatic expressions of the dark side of religion sometimes also make it into the newspapers, though many times not. There are the recurrent episodes of clergy malfeasance, for example, including a conservatively estimated 400 Roman Catholic priests and brothers charged with sexual abuse since 1982. Long reluctant to acknowlege the problem, the leadership of the Church now faces serious financial problems from the many millions of dollars awarded by the courts to the numerous victims. Yet clergy abuse of congregants is not limited to Roman Catholic circles: various studies have found that it is common in Protestant churches as well, including both conservative and liberal denominations (Shupe, 1995). Adding insult to injury in many cases is the refusal of the church hierarchies to deal appropriately with the abuse and the tendency of many congregants to blame and harass the victims when they make their charges known (Fortune, 1989). Victims of child sexual abuse within their own families often suffer the same treatment by church authorities (Imbens and Jonker, 1985). callous

Deeply troubling, too, are the stories of the resurgence of repressive fundamentalism in various parts of the world. In Afghanistan, for example, more than half the country is now dominated by the Taliban, an Islamic fundamentalist political and military force that, emerging from the chaos of civil war, is abruptly wrenching this country's millions of inhabitants back into a dark past. Women long accustomed to various freedoms are now being forced back into purdah, the Muslim tradition of secluding women in their homes. Allowed only to work in hospitals

and clinics, and then only with members of their own sex, women must shroud themselves from head to toe whenever they go out. Girls, told that education is only for boys, have been expelled from schools and colleges. Television sets and stereo systems have been publicly "hanged"and books thought to be tainted by Western influence confiscated. The Koran is harshly interpreted and enforced through modern means: murderers and "enemies" of the Taliban are publicly hanged from cranes, and the hands and arms of convicted thieves are surgically amputated. Remarked one elderly Afghan scholar, "We are ruled now by men who offer us nothing but the Koran, even though many of them cannot read; who call themselves Muslims, and know nothing of the true greatness of our faith. There are no words for such people. We are in despair" (Burns, 1996, p. A8).

A dark side indeed—but might it not be the case, as the Muslim scholar seems to suggest, that what we see here are not genuine forms of religiousness, but aberrations and perversions of it? Or perhaps they reflect a radically different human impulse that insidiously masquerades as piety. Are not most religious people basically good and kind souls, even those who think of themselves as sinful? Might it not be that, when they misstep, it is in spite of their religiousness, not because of it?

An Elusive Search for Humanitarianism

More than a half century ago, social scientists began exploring the relationship of religiousness to a variety of moral and humanitarian concerns. What they uncovered was disturbing, especially to those who were religious themselves. Abraham Franzblau (1934), for example, found a negative relation between acceptance of religious beliefs and all three of his measures of honesty. Furthermore, religious belief bore no relation to his test of character. Among the nearly 2000 associates of the Young Men's Christian Association who responded to Murray Ross's (1950) questionnaire, the agnostics and atheists were more likely than the deeply religious to express willingness to help the needy and to support radical reform. Hirschi and Stark (1969) discovered that children who attended church regularly were no less likely to commit illegal acts, according to their own estimations, and Ronald Smith, Gregory Wheeler, and Edward Diener (1975), in a QUASI-EXPERIMENTAL situation, found that religious college students—including a group of Jesus people—were no less likely to cheat on a multiple-choice test and no more likely to volunteer to help mentally retarded children than atheists and other "nonreligious" persons. The religious subjects in Russell Middleton and Snell Putney's (1962) investigation even reported a *higher* frequency of cheating on examinations than did the skeptics.

In only two areas of moral concern do religious subjects consistently distinguish themselves: drugs and sex. However religiosity is measured, it has proved to be negatively related to the use of illicit drugs (Benson, 1992; Gorsuch, 1995). Similarly, high scorers on religiosity measures are significantly less likely than low ones to approve of or engage in any form of sexual behavior that has traditionally met with social disapproval. Responding to a variety of religiosity scales, unmarried college students have consistently shown themselves to be less permissive concerning premarital sex the more conservative they are religiously (Cardwell, 1969; Clayton, 1971b; Heltsley and Broderick, 1969; Sutker, Sutker, and Kilpatrick, 1970; Woodroof, 1985). Religiousness in married subjects is likewise associated with avoidance of traditionally disapproved practices, such as extramarital sex. Apparently, however, it is not correlated with any other measured dimensions, such as frequency

of intercourse or consistency of attaining orgasm (Bell, 1974; Fisher, 1973; Martin and Westbrook, 1973).

Abstention from illicit drugs and disapproved sexual behavior is not paralleled by a corresponding attentiveness to traditional humanitarian ideals. Studies investigating the relation of measures of religiousness to scales labeled "humanitarianism" have consistently found either no relation (Ferguson, 1944b) or a slightly negative one (Defronzo, 1972; Kirkpatrick, 1949). Similarly, Victor Cline and James Richards (1965) report no correlation between their Compassionate Samaritan factor and acceptance of conventional religious teachings, and Bruce Hunsberger and Ellen Platonow (1986) found that higher scorers on a Christian orthodoxy scale were no more likely than lower scorers to volunteer to help charitable groups. Rokeach (1969) found his religious subjects to be preoccupied with personal salvation and relatively indifferent to social inequality and injustice.

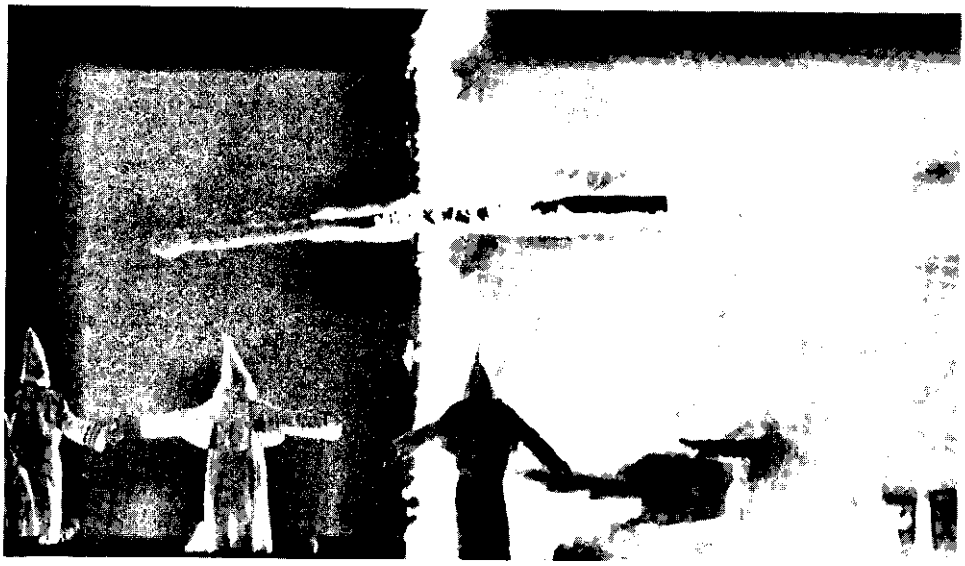
The presumed connection between piety and "prosocial behavior" has been equally elusive in the laboratory. Lawrence Annis (1976) found that none of his four measures of religiousness predicted who among his subjects would investigate nearby sounds of a "lady in distress." Similarly, Ralph McKenna (1976) reports that when a "stranded female motorist" claimed to have misdialed in trying to contact her garage, clergymen (or whoever answered their telephones) were no more willing than control subjects to place the needed call.

On the other hand, a few studies do report findings more nearly consistent with religion's traditional image. Warner Wilson and Wallace Kawamura's (1967) student subjects showed a small, positive correlation between measures of religiousness and social responsibility. In L. D. Nelson and Russell Dynes's (1976) study, both devotionalism and church attendance showed low, positive correlations with self-reported helping behavior, both ordinary and emergency, in a city that had earlier been struck by a tornado. Robert Friedrichs (1960) similarly reports a low but positive relation of religiousness to a rating of altruism, but only when piety is measured by belief in God, not by church attendance. In all three studies, however, correlations are so small that religiousness accounts for less than 5 percent of the variability in humanitarian concern.

Upwards to 14 percent of humanitarianism's variance is accounted for by religiosity among the British and American students who participated in Wesley Perkins's (1992) study in 1978–79. But in comparable student samples ten years later, the proportion dropped to 9 and 1 percent, respectively. In the same study, egalitarianism proved to be virtually unrelated to religiosity, which was operationalized as affirming religion as a source of guidance and of answers to a variety of problems.

Religion and Prejudice

The apparent failure of religious involvement to foster a humanitarian outlook has received its closest assessment in research on conservative social attitudes. The overall pattern of findings here is much like the one we have just reviewed. Using a variety of measures of piety—religious affiliation, church attendance, doctrinal orthodoxy, rated importance of religion, and so on—researchers have consistently found positive correlations with ETHNOCENTRISM, AUTHORITARIANISM, dogmatism, social distance, rigidity, intolerance of ambiguity, and specific forms of prejudice, especially against Jews and blacks (Batson and Burris, 1994; Batson, Schoenrade, and Ventis, 1993; Dittes, 1969; Gorsuch, 1988; Gorsuch and Aleshire, 1974; Hunsberger, 1995).



This traditional cross burning by members of the Ku Klux Klan at a 1987 rally led by the group's national leader in Rumford Point, Maine, roused the anger of many local religious and civic leaders. The Klan is well known for its campaigns of hate against Jews, blacks, and Catholics.

There are very occasional exceptions. In an ethnically diverse group of introductory psychology students, Gerald Meredith (1968) found no significant correlation between two measures of religious attitude and a pair of ethnocentrism and dogmatism scales. Dean Hoge and Jackson Carroll (1975) found that, after they statistically removed the effects of other variables, all their religious measures together accounted for only 5 percent of the variance in their index of ANTI-SEMITISM. Similarly controlling for social–psychological factors, Middleton (1973) reduced the common variance between these dimensions to a mere 2 percent. The factor-analytic study of Christopher Bagley, Roger Boshier, and David Nias (1974) yielded two factors, a religious one and a "racialist-punitive" one, which were wholly independent of each other; identical results were found for diverse samples in England, the Netherlands, and New Zealand. In a sample of 652 students in six Southern colleges, Terry Prothro and John Jensen (1950) found a slight but significant correlation between positive attitudes toward the church and toward blacks and Jews.

A Curvilinear Relation: Some of these exceptional studies suggest that, among the subject groups studied thus far, the positive relation of religion and prejudice holds mainly for white Americans. Other studies show a pattern that continues to recur: the positive correlation becomes negative at the upper extreme of piety. Gregory Shinert and Charles Ford (1958), for example, compared a group of daily communicants (a majority of whom were seminarians and nuns) with a group of laypersons at the same Roman Catholic university who did not receive communion daily. The daily communicants proved, on the average, to have significantly *lower* ethnocentrism scores. Similarly, Glenn Wilson and Francis Lillie (1972; Wilson, 1973) found that two groups at the extremes of conventional religiosity—officer cadets of the Salvation Army and members of the Young Humanist Association—both showed exceptionally low levels of racial prejudice.

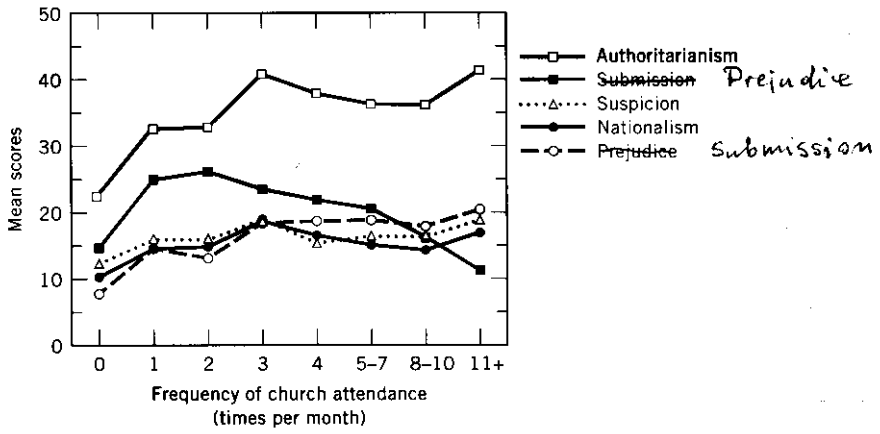


Figure 6.1. Church attendance in relation to four factors of Authoritarianism and to Prejudice. Source: From H. H. Remmers (Ed.), *Anti-Democratic Attitudes in American Schools.* Evanston, Ill.: Northwestern University Press, 1963, page 253.

The CURVILINEAR relation between piety and prejudice is well illustrated in a study by Elmer Struening (1963). Using as his subjects the instructional, administrative, and professional service staffs of a large Midwestern university, Struening sought to explore the relation of several variables to scales of authoritarianism and prejudice. The authoritarianism scale, taken largely from the F SCALE of Theodor Adorno and his associates (1950), was scored for a general factor (in Figure 6.1, Authoritarianism) as well as three specific ones: "authoritarian religious submission" (Submission), "cynical and suspicious view of human environment" (Suspicion), and "aggressive authoritarian nationalism" (NATIONALISM). The prejudice scale, which yielded a single score (Prejudice), contained items assessing both generalized prejudice and discriminatory attitudes toward specific groups. The relation of these five variables to frequency of church attendance is shown in Figure 6.1.

First, let us note the obvious curvilinear relation between church attendance and prejudice. As in other studies, those who report attending church once or twice a month have strikingly higher prejudice scores than those who do not attend at all. As the frequency of attendance rises beyond the twice-monthly point, the level of prejudice falls off, until—at an attendance rate of more than twice a week—it reaches a point slightly lower than the mean prejudice score of those who never attend.

The same general pattern is evident in other studies. Friedrichs (1959) found that New Jersey community residents who said they attended between 31 and 60 religious services in the year 1958 were the most prejudiced, and those reporting 61 or more services, the least prejudiced among his subjects, including those who claimed no attendance at all. The University of Texas subjects of Robert Young, William Benson, and Wayne Holtzman (1960) who attended church weekly were more tolerant of blacks than those reporting attendance once or twice a month, though they remained less tolerant than nonattenders. Hoge and Carroll (1973) found that members of eight Protestant churches showed the curvilinear pattern, though it was somewhat erratic, especially for members of the Southern churches. Dean Kilpatrick, Louis Sutker, and Patricia Sutker (1970) report the curvilinear trend in their Southern university data, but only for the female students.

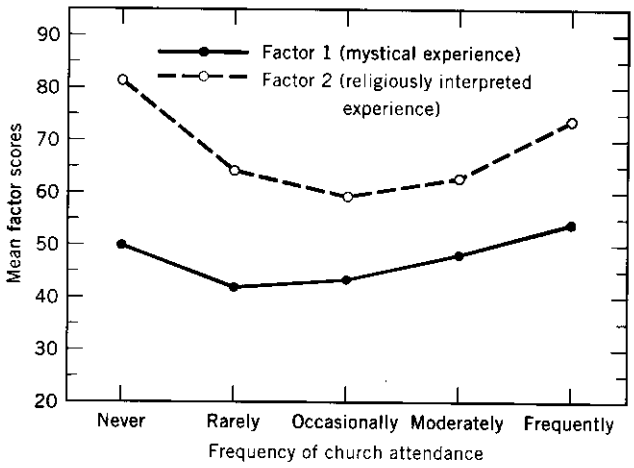


Figure 6.2. Relation of church attendance and mysticism factor scores. (The number of subjects was 34 for each category of attendance; analysis of variance was significant for each factor, $p < .01$.) Source: Based on Table 2 in Hood, 1976, page 1131.

Although the pattern has been found chiefly in studies of prejudice and church attendance, it has also appeared in research on religious belief and experience. In an experimental study of obedience to destructive commands modeled after Stanley Milgram's (1974) classic series of investigations, David Bock and Neil Warren (1972) found that subjects with moderate religious beliefs administered significantly higher levels of electric shock when the experimenter's confederate failed at his task than did the religiously extreme groups. Hood (1976) discovered that, among undergraduates who completed his Mysticism Scale, the more extreme the attendance frequency—high or low—the more likely the report both of mystical experience (Factor I on his scale) and of religiously interpreted experience (Factor II; see Figure 6.2).

Studies on well-being have reported the pattern as well. In a random sample of women who responded to a *Redbook* magazine questionnaire, Phillip Shaver, Michael Lenauer, and Susan Sadd (1980) found that reported physical and mental symptoms as well as overall unhappiness showed a striking curvilinear relation to self-reported degree of religiousness, with the slightly religious reporting the most symptoms and unhappiness; for both variables, the relation is significant at the .001 level (see Figure 6.3). Similarly striking is the curvilinear relation that Anette Dörr (1987, 1992) found between a religious experience scale (consisting of five items from Boos-Nünning's [1972] General Religiosity factor) and self-rated depressivity. Among her 162 subjects, a third of whom were patients being treated for depression, those falling in the middle group on the religious experience scale had the highest average level of depressivity. Those in the highest religiosity group were lowest on the depression scale.

Unfortunately, many studies do not scale the religiosity dimensions finely enough to allow the curvilinear pattern to appear. In addition, when significance tests have been applied to the data, they are usually not sensitive to curvilinearity. Even so, of the 25 studies they reviewed, Richard Gorsuch and Daniel Aleshire (1974) judge that the findings of 20 are consistent with this trend. Evidence is inconclusive on the question of which extreme has the lower prejudice scores.

Had Struening cut off the high extreme of his church-attendance scale at eight

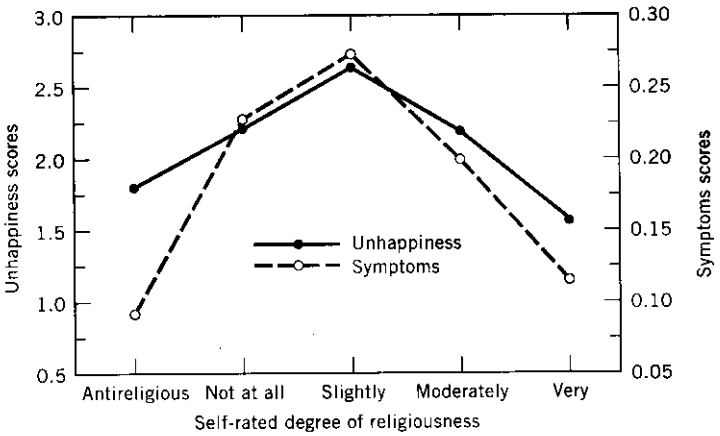


Figure 6.3. Relation of religiousness to unhappiness and symptoms scores. Source: Based on Table 1 and Figure 1 in Shaver, Lenauer, and Sadd, 1980, page 1567. Modifications made in consultation with the first author.

to ten times per month, we would probably conclude that authoritarianism also bears a curvilinear relation to piety. Such a conclusion would have been consistent with the finding of most other researchers that the various measures of conservative social attitudes tend to be highly intercorrelated (Allport and Ross, 1967; Maranell, 1967; O'Neil and Levinson, 1954; Rokeach, 1960; Weima, 1965). Prothro and Jenson (1950) see evidence in these correlations for a general tolerance factor, which Wilson (1973), seeking a more nearly neutral term, argues for calling conservatism. Yet Struening's results, and those of Frank Knöpfelmacher and Douglas Armstrong (1963), suggest that lumping these dimensions together may be premature. Knöpfelmacher and Armstrong discovered that the positive correlation between ethnocentrism and authoritarianism held for all their religious groups of adolescents *except* the Catholics. As in some other studies, the Catholic subjects, on the average, had higher authoritarian scores than any other group. Yet in their responses to the Bogardus Social Distance Scale, they were also significantly less ethnocentric than all other denominational groups. The Catholics in Stark's (1971) study showed a similar trend.

The Role of Acquiescence and Conformity: It is crucial to note that the items on the F Scale are so worded that agreement with them yields a high authoritarian score, and disagreement a low one (see Table 6.3). Such a scale is thereby open to the response set of ACQUIESCENCE, the tendency to answer "true" or "yes" whatever the item content. Indeed, Dean Peabody (1961) found that agreement with items on the F Scale can frequently be attributed to a general set to agree, whereas disagreement seems to be indicative of *anti*authoritarian attitudes. Thus scores on the F Scale might be expected to correlate with those on any other measure that entails a significant element of acquiescence—including, it would appear, most common scales of religiousness (Fisher, 1964) as well as other indicators of conservative attitudes. Struening's subjects who attended church two times a month or less often may have been individuals whose motives for conformity were strong enough to prompt occasional attendance yet insufficient to compel acquiescence to the church's teachings against discrimination. His other subjects may have had greater needs for conformity, leading to elevated F scores, frequent attendance, and relatively low prejudice scores. When the church itself tends to encourage

Table 6.3 **SELECTED ITEMS FROM THE CALIFORNIA F SCALE AND THE ROKEACH DOGMATISM SCALE**

California F Scale[a]	Rokeach Dogmatism Scale
1. Obedience and respect for authority are the most important virtues children should learn.	1. The United States and Russia have just about nothing in common.
2. Science has its place, but there are many important things that can never possibly be understood by the human mind.	2. Man on his own is a helpless and miserable creature.
3. Every person should have complete faith in some supernatural power whose decisions he obeys without question.	3. I'd like it if I could find someone who would tell me how to solve my personal problems.
4. Homosexuals are hardly better than criminals and ought to be severely punished.	4. It is only natural for a person to be rather fearful of the future.
5. The businessman and the manufacturer are much more important to society than the artist and the professor.	5. There is so much to be done and so little time to do it.
6. When a person has a problem or worry, it is best for him not to think about it, but to keep busy with more cheerful things.	6. Once I get wound up in a heated discussion I just can't stop.
7. Some day it will probably be shown that astrology can explain a lot of things.	7. If given the chance I would do something of great benefit to the world.
8. People can be divided into two distinct classes: the weak and the strong.	8. Of all the different philosophies which exist in this world there is probably only one which is correct.
9. Human nature being what it is, there will always be war and conflict.	9. A group which tolerates too much difference of opinion among its own members cannot exist for long.
10. Most people don't realize how much our lives are controlled by plots hatched in secret places.	10. Most people just don't know what's good for them.

[a] F Scale items are from Forms 40 and 45, containing 30 items, and Dogmatism Scale items are from Form E, containing 40 items. All items on both scales are scored in the positive direction.
Source: California F Scale items from Adorno et al., 1950, pages 255–256; Rokeach Dogmatism Scale items from Rokeach, 1960, pages 73–80.

prejudicial attitudes—a tendency that Glock and Stark (1966; Stark, et al., 1971) argue is inherent in Christian orthodoxy's claim to being the sole possessor of truth—all measures of conformity, including church attendance, would bear a linear relationship to prejudice.

Yet the conformity said to be one of the chief characteristics of the authoritarian personality—whose putative fondness for authority, distrust of other human beings, and inflexible patterns of thinking have long been familiar in the literature of the social sciences—is much more than the simple tendency to acquiesce. In fact, the conformity that is part of this personality syndrome appears to be separable from the acquiescent response set (Kirscht and Dillehay, 1967). Some prejudice is undoubtedly the result of unreflective and "innocent" acquiescence to the opinions of associates. In contrast, the prejudice of the authoritarian personality is thought to be highly motivated, a product of the suspiciousness, even hatred, of all persons

who are not part of one's own group. Yet prejudice is not the only response to the feeling of threat and the underlying sense of life's precariousness; churchgoing is one as well (Dittes, 1973b).

A Balanced Authoritarianism Scale: One obvious way to reduce the effects of a response set would be to prepare a new scale with half the items reversed—that is, written in the "contrait" direction—so that on these items agreement would *lower* the total score rather than raise it. Thus a strong tendency to agree would yield moderate rather than extreme scores. Canadian psychologist Bob Altemeyer (1988) took precisely this precaution in painstakingly developing a new 30-item Right-Wing Authoritarianism (RWA) Scale. Doubtful that we really know what the California F Scale measures, or what the Dogmatism Scale (see Table 6.3) that Milton Rokeach (1956) developed as an ideology-free alternative is about, Altemeyer found that the RWA is more internally consistent and unidimensional than these scales. It is also better at predicting other attitudes and behavior.

During 20 years of highly original research with one or another version of the RWA and a variety of measures of religiosity, most of which he also developed, Altemeyer found a rather straightforward pattern of relationships.

> The findings are really rather simple. Authoritarians in my samples tend to be religious persons, and vice versa. High RWAs usually have tightly wound religious ideologies. They appear to be under appreciable pressure to believe truly, and they keep doubts to themselves, split off and tucked away.
>
> One finds other evidence of balkanizing. Vast, complicated religious material such as the Bible is "lined up" to support authoritarian submission, aggression, and conventionalism. Contradictory material exists alongside the selected interpretation but is disconnected. Or Highs may say they agree with Jesus' admonition not to judge and condemn others, but this "agreement" has no apparent effect on their behavior. Their belief system appears self-confirming, enduring, and closed. Really, the beliefs could be anything, and hostilities based on them appear highly resistant to change (pp. 230–231).

Altemeyer emphasizes that not all his religious subjects were high-scoring RWAs, nor was it impossible to find high-scoring atheists. Yet overall these trends were consistent and strong, with the highest intercorrelations appearing with the Christian Orthodoxy Scale—"particularly ironic," he notes, "for the Gospels largely portray Jesus of Nazareth as tolerant, forgiving, and preaching a message of universal love" (p. 201). The highest scorers on the RWA were the Fundamentalists and Mennonites whereas United Church members, Anglicans, and Jews scored lower, with Catholics and Lutherans falling in between. The lowest scorers of all were those with no religious affiliation.

The Authoritarian's Positive Side: Like most other researchers, Altemeyer considers authoritarianism a distinctly undesirable trait. Thus he notes with real concern the more or less steady increase in RWA scores among University of Manitoba students between 1973 and 1987. He has subsequently undertaken studies to see if prejudice among right-wing authoritarians can be reduced through personal-value confrontation (Altemeyer, 1994).

Experimental psychologist and former chaplain A. T. Welford (1971), in contrast, doubts that the authoritarian personality "deserves all the scorn that has been heaped upon it."

> The "authoritarian" has been regarded as an inflexible, unimaginative, intolerant, over-conforming, prejudiced disciplinarian. Yet, if one looks closely at the questionnaire statements used to define him, it seems fair to describe him also as a person who

> is prepared to sacrifice some spontaneity for stability, some permissiveness for the sake of order and peace, some immediate pleasure for the pursuit of long-term aims, and some sentimental toleration for the sake of efficiency. He may, of course, go too far in these directions, and may not appear as a gay, interesting person with whom immediate easy friendship is possible. A substantial measure of his qualities is, however, essential for dealing responsibly with the world as it is, and even more for making ideals come true (p. 47).

Two investigations seem to underscore some positive qualities in the authoritarian or dogmatic personality. Employing a modified version of the F Scale as well as the Gough Adjective Check List, Mark Allen (1955) found that the "religious authoritarians" among his Mormon subjects described themselves as cooperative, idealistic, mannerly, praising, submissive, and forgetful; the nonauthoritarians tended to describe themselves as bitter, cold, cynical, egotistical, defensive, outspoken, prejudiced, self-centered, and stern. David Williams and Bruce Kremer (1974) also found that conservative attitudes need not always have negative correlates. To students enrolled in a secular counseling program, on the one hand, or in counseling programs in Roman Catholic or Protestant seminaries, on the other, Williams and Kremer administered three testing instruments: Rokeach's Dogmatism Scale, a measure of acceptance of traditional Christian doctrines, and a portion of a Test of Counselor Attitudes. Although the pastoral counseling students proved to have higher religiosity and dogmatism scores, as might have been predicted, they also showed more nondemanding acceptance of clients described to them than did the secular counselors.

A response set, here in the form of social desirability, may also have contributed to these results. Yet it is also true that what ostensible scales of authoritarianism and dogmatism actually measure is far from clear (Kirscht and Dillehay, 1967; McKinney, 1973). Without doubt these and related scales have yielded predictable results consistent with the theories proposed to explain the origins and dynamics of the personality characteristics in question. As is true of all other isms, however, there is a danger of reification and overgeneralization. No psychological assessment device should be taken at face value. The care with which Altemeyer has developed and critically examined his own authoritarianism scale, while continuing over the years to monitor changes in its psychometric characteristics, is exemplary in this regard.

The Role of Other Nonreligious Factors: Of special importance in interpreting scales measuring authoritarianism, intolerance, or conservatism are the factors of intelligence, education, socioeconomic status, and national region. Correlations (chiefly negative) between each of these variables and conservatism have appeared in disparate contexts and at various times. These factors also happen to be related to measures of religiousness. Thus, when statistical methods have been used to separate out their contribution to the relation between piety and prejudice, the result is frequently a much-reduced correlation. By taking into account education's negative correlation with prejudice, for example, Joe Feagin (1964) reduced religious orthodoxy's correlation with prejudice from .35 to .23 and thus their common variance from 12 to 5 percent. Bagley and Boshier (1975) found that the correlations of .20 between Roman Catholic affiliation and a measure of prejudice dropped to an insignificant .07 when they controlled for social class and affiliation-without-attendance. And Rob Eisinga, Ruben Konig, and Peer Scheepers (1995) found that, for their Dutch subjects, the introduction of such nonreligious variables as age,

region, political allegiance, and social class substantially reduced the effects of orthodox religious beliefs on secular anti-Semitism.

The mediating role of intelligence may be more fundamental. Reinterpreting authoritarianism chiefly in cognitive rather than in motivational terms, Chris Brand (1981) maintains that authoritarians are best characterized not by their fondness for authority but by "some simple-minded way in which the world has been divided up for them" (p. 23)—most basically, according to species, race, gender, and age. The conflicting needs of these groups loom large in the lives of authoritarians, whose prejudices naturally favor their own positions and interests. Education, especially in the liberal arts, counteracts so crude a world-view, Brand observes. Thus he attributes the negative correlations between authoritarianism and intelligence test scores (which average around $-.50$) primarily to "crystallized intelligence," the enduring intellective capacity that is shaped by education and experience. The correlation between right-wing authoritarianism and education ($r = -.28$, $p < .01$) reported by Altemeyer and Hunsberger (1992) supports this interpretation.

Religiousness, especially religious orthodoxy, has also been found to be negatively correlated with intelligence and irreligiousness to be positively related with it. Thurstone (1934) found intelligence to be positively correlated with liberal or radical attitudes, including agnostic or atheistic religious views. The same trend is reported by Dean Hoge (1974), who found social and religious liberalism to be positively correlated with verbal scores on the Scholastic Aptitude Test (SAT). The religious skeptics among Poythress's (1975) subjects scored significantly higher than religious believers on both the verbal and the quantitative portions of the SAT, a trend likewise observed by Heist and Yonge (1968). Poythress's skeptics, it should be noted, scored significantly lower on authoritarianism.

Brand speculates that religiousness is related chiefly to "fluid intelligence," the capacity to solve new problems. He observes that piety is consistently linked to age: young children and the elderly, more than other age groups, tend to show high levels of religious belief and involvement (see Argyle and Beit-Hallahmi, 1975). Not coincidentally, says Brand, fluid intelligence shows a mirror-image pattern: it rises through childhood and declines from middle age onward. Thus he concludes that to be religious is to be unable or unwilling to think independently (p. 27). Although fluid and crystallized intelligence are positively correlated, Brand intimates that their relation to one another and to the various facets of authoritarianism and piety is too complex to allow simple generalizations regarding the association of religious and authoritarian trends. Only when the level of fluid intelligence forecloses a liberalizing education and thus sharply limits crystallized intelligence, says Brand, will authoritarianism and religiosity appear to merge.

Two Types of Piety: Intrinsic and Extrinsic

Thought provoking and worthy of testing though Brand's generalizations are, they overlook the fact that people are religious in different ways. It is these differences, some researchers argue, that provide the key to understanding the religion–prejudice connection. Over the last several decades, one distinction in particular has stood out. Theodor Adorno and his associates (1950) reported in their classic study of the authoritarian personality that the conventionally religious—those disposed "to view religion as a means instead of an end" (p. 733) and to attend church in order to be classed with normal or even privileged people—are the ones who show ethnocentric attitudes. In contrast, persons who "take religion seriously," for whom

religion is "a system of more internalized, genuine experiences and values" (p. 310), are likely to oppose ethnocentrism.

Measuring the Two Types: Struck by these findings, Allport (1954) at first identified the contrasting religious outlooks as "institutionalized" and "interiorized." Later, Allport (1959) chose the terms *extrinsic* and *intrinsic,* which, in spite of conceptual and psychometric difficulties, have won widespread acceptance. The extrinsic orientation is characteristic of those who

> are disposed to use religion for their own ends. The term is borrowed from axiology, to designate an interest that is held because it serves other, more ultimate interests. Extrinsic values are always instrumental and utilitarian. Persons with this orientation may find religion useful in a variety of ways—to provide security and solace, sociability and distraction, status and self-justification. The embraced creed is lightly held or else selectively shaped to fit more primary needs. In theological terms the extrinsic type turns to God, but without turning away from self.

The intrinsic orientation characterizes those who

> find their master motive in religion. Other needs, strong as they may be, are regarded as of less ultimate significance, and they are, so far as possible, brought into harmony with the religious beliefs and prescriptions. Having embraced a creed the individual endeavors to internalize it and follow it fully. It is in this sense that he *lives* his religion (Allport and Ross, 1967, p. 434).

Allport and his students set about to develop measures of these two orientations. After an initial effort by Cody Wilson (1960), who sought to assess only the extrinsic trend, a new, 21-item scale that included both (see Table 6.4) was prepared. By factor-analyzing the responses to it of a group of Southern Baptists, Feagin (1964) discovered that the two types of items formed separate, virtually unrelated scales. This result disconfirmed Allport's assumption that the two orientations lie at opposite ends of the single dimension. On the basis of the factor loadings, Feagin formed a 12-item Intrinsic/Extrinsic Scale, with six items measuring each orientation. Allport and Michael Ross (1967), apparently considering the longer scale more adequate, used all but one of the original 21 items in their later study. Their 20-item Religious Orientation Scale, which Hood (1971) found to correlate relatively highly (.78) with Feagin's shorter form, is the one most widely employed today.

The Two Types and Social Attitudes: Among his Southern Fundamentalists, Feagin discovered that only the Extrinsic scale correlated consistently with an antiblack scale. For one group alone—the college students—did the pattern reverse itself, with the Intrinsic scale showing the significant relationship. In every case, however, the correlation was in the predicted direction. An extrinsic orientation tended to go along with prejudice, and an intrinsic one with rejection of it. Allport and Ross found much the same pattern in their denominationally more diverse sample. In most instances, those scoring high on the Extrinsic scale were more prejudiced on all five measures than high scorers on the Intrinsic scale. The most prejudiced of all, however, were the "indiscriminately proreligious," who agreed with both intrinsic and extrinsic items. Allport and Ross suggest that this response pattern reflects an undifferentiated cognitive style with roots deep in the personality structure. Such a style would also account for the stereotypy of prejudice.

For the record, it should be observed that in two of Allport and Ross's samples—the 53 Pennsylvania Presbyterians and the 35 Tennessee Methodists—the intrinsic types were *more* antiblack than the extrinsic types. Among the Roman Catholic, Lutheran, and Nazarene subjects, the indiscriminately proreligious appeared to be

Table 6.4 THE ALLPORT–ROSS RELIGIOUS ORIENTATION SCALE[a]

Intrinsic Items	Extrinsic Items
2. I try hard to carry my religion over into all my other dealings in life.* (.57/–.03)	1. What religion offers me most is comfort when sorrows and misfortune strike.* (.00/–.72)
7. Quite often I have been keenly aware of the presence of God or the Divine Being. (.36/.17)	3. Religion helps to keep my life balanced and steady in exactly the same way as my citizenship, friendships, and other memberships do.* (–.04/–.64) (Omitted from the Allport and Ross scale)
8. My religious beliefs are what really lie behind my whole approach to life.* (.59/.00)	4. One reason for my being a church member is that such membership helps to establish a person in the community.* (.03/–.49)
9. The prayers I say when I am alone carry as much meaning and personal emotion as those said by me during services.* (.56/–.08)	5. The purpose of prayer is to secure a happy and peaceful life.* (.03/–.72)
13. If not prevented by unavoidable circumstances, I attend Church (a) more than once a week, (b) about once a week, (c) two or three times a month, (d) less than once a month.* (.61/–.03)	6. It doesn't matter so much what I believe so long as I lead a moral life. (.10/–.05)
14. If I were to join a church group I would prefer to join (1) a Bible Study group or (2) a social fellowship. (.36/–.08)	10. Although I am a religious person I refuse to let religious considerations influence my everyday affairs. (.21/–.31)
16. Religion is especially important to me because it answers many questions about the meaning of life. (.40/.11)	11. The Church is most important as a place to formulate good social relationships.* (–.21/–.51)
18. I read literature about my faith (or church) (a) frequently, (b) occasionally, (c) rarely, (d) never.* (.68/–.02)	12. Although I believe in my religion, I feel there are many more important things in life. (.03/–.16)
20. It is important to me to spend periods of time in private religious thought and meditation.* (.60/.03)	15. I pray chiefly because I have been taught to pray. (–.14/–.26)
	17. A primary reason for my interest in religion is that my church is a congenial social activity. (.01/–.05)
	19. Occasionally I find it necessary to compromise my religious beliefs in order to protect my social and economic well-being. (.32/.10)
	21. The primary purpose of prayer is to gain relief and protection.* (.10/–.63)

[a] These items, revised by the members of a Harvard University seminar in 1963, are worded according to Allport and Ross (1967), who provided four alternative responses to each statement (see items 13 and 18 for examples). The numbering here follows Feagin (1964). Each item is followed by its loading on Feagin's two factors (intrinsic and extrinsic, respectively). An asterisk indicates that Feagin assigned the item to his corresponding six-item subscale. (See also Hunt and King, 1971.)
Source: Adapted from Table 1 in Feagin, 1964.

slightly *less* prejudiced than the extrinsic type. Unfortunately, it is not unusual for statistical trends inexplicably to reverse themselves. Yet, given the complexity of causation in the social–psychological realm and the relative crudeness of the measures, it is hardly surprising that findings are not always consistent. Buried in every sample, after all, are the anonymous individuals who deviate from the trends that the group shows as a whole. Recurrent patterns are the most the correlationists can

hope for, and these abound in research employing the Allport–Ross and Feagin scales.

Because the Intrinsic and Extrinsic scales have proved with most samples to be only slightly (and negatively) correlated with each other (Donahue, 1985), generalizations about the correlates of one scale do not predictably apply in reverse to the other. The Extrinsic scale, but not the Intrinsic, has shown persistent correlation with the conservative attitudes we have been discussing, including prejudice (Amón, 1969; Brannon, 1970; Hoge and Carroll, 1973; Matlock, 1973; Morris, Hood, and Watson, 1989; Ponton and Gorsuch, 1988; and studies cited earlier) dogmatism (Hoge and Carroll, 1973; Kahoe and Dunn, 1975; Strickland and Weddell, 1972; Thompson, 1974); authoritarianism (Kahoe, 1974, 1975); and ethnocentrism (Dicker, 1977).

The Intrinsic scale does occasionally correlate negatively with such social attitudes, as we have already seen, and it has also been found to be positively related to volunteering to help others (Bernt, 1989; Benson, et al., 1980; Hunsberger and Platonow, 1986). Yet in some other studies it has proved to be *positively* related to conservative or prejudicial attitudes. For the Roman Catholics, Protestants, and Jews whom A. Lange (1971) studied in Amsterdam, intrinsic religiosity consistently was correlated positively with authoritarianism, rigidity, and dogmatism (with coefficients ranging from .30 to .46). Similarly, in a group of predominantly Baptist college students, Kahoe (1975) discovered significant positive correlations between the Intrinsic scale and 11 of the 30 items composing the F Scale. Whereas the more numerous F Scale items correlating with the Extrinsic scale were scattered throughout the nine item cateogories of Adorno and his associates (1950), those related to the Intrinsic scale—including foremost items 2 and 3 in Table 6.3—are limited largely to four of these clusters, Conventionalism, Superstition and Stereotypy, Authoritarian Aggression, and, above all, Authoritarian Submission. Correlation with the last two qualities is confirmed by Bock's (1973) discovery that it was the intrinsic subjects, not the extrinsic, who tended to obey the commands of a scientific authority to administer an electric shock to an experimental confederate. And Altemeyer (1988, pp. 210, 218) reports that his Right-Wing Authoritarianism Scale correlated positively with all but one of the 11 Intrinsic scale items, yielding an overall correlation of .36 for students and .41 for their parents. The relation with the items of the Extrinsic scale was much less consistent, with corresponding correlations of $-.10$ and $-.09$.

As Glenn Griffin, Richard Gorsuch, and Andrea Davis (1987) remark on their finding that, among Seventh-Day Adventists on St. Croix, intrinsic religiosity was positively correlated with prejudice against Rastafarians, the relation between religious orientation and social attitudes is mediated by the particular religious norms in terms of which the person is oriented. If a tradition itself tends to foster prejudice against another group, as seemed to be the case on St. Croix, then intrinsically oriented persons within the tradition will likely share this prejudice. An intrinsic orientation in itself is thus no guarantee of positive social attitudes.

Because the Allport–Ross Intrinsic scale tends to correlate positively with orthodox Christian views (e.g., Johnson, George, and Saine, 1993), high scorers tend to share the social attitudes and values traditionally associated with conservative Christian groups. Thus, in comparison to low scorers, those high on the Intrinsic scale are likely to be less sexually permissive (Haerich, 1992; Reed and Meyers, 1991, Wann, 1993; Woodroof, 1985); to prefer nonrevealing clothing (Edmonds and Cahoon, 1993); to express nonfeminist attitudes (McClain, 1979), including,

among women respondents, opposition to egalitarianism (Kahoe, 1974b) and an emphasis on family over career (Jones and McNamara, 1991); and to be prejudiced against gay men and lesbians (Herek, 1987). Similarly reflecting traditional Christian values, high Intrinsic scorers tend to score higher on measures of responsibility (Kahoe, 1974a), altruism (Chau, et al., 1990), empathy (Watson, Hood, Morris, and Hall, 1984), and social interest (in the Adlerian sense of feeling connected to others; Leak, 1992). High scorers have also shown a tendency to score higher on Glock's religious practice, feeling, and knowledge dimensions (Dodrill, Bean, and Bostrom, 1973), to report more religious or mystical expereinces (Hood, 1970, 1973; Weima, 1986), and, in relation to Lawrence Kohlberg's (1981) levels of moral development, to show "principled" rather than "conventional" moral thinking when the individual's denomination itself operates at the higher level (Ernsberger and Manaster, 1981).[1]

The Indiscriminate Types: A few studies have sought to take into account the numerous subjects who agree with both intrinsic and extrinsic items, the "indiscriminately proreligious," as well as the indeterminate number who disagree across the board, the "indiscriminately antireligious" or "nonreligious" (Allport and Ross, 1967). Andrew Thompson (1974) found that the adolescents in Roman Catholic religious education classes who were *by comparison* indiscriminately antireligious had the lowest dogmatism scores of the four groups, although their scores were not significantly lower than those of the next lowest, the intrinsic category. (If we were certain that the acquiescent response set played a role in these data—and it well may have—this is precisely the ordering we would expect.) The adolescents' parents, on the other hand, gave responses that are consistent with Allport's predictions: the indiscriminately proreligious are the most dogmatic, the extrinsic and indiscriminately antireligious follow, and the intrinsic are the least dogmatic. Sanderson's (1974) study of student subjects yielded results similar to Thompson's. Right-wing political views—including superpatriotism, opposition to civil liberties, ethnic bigotry, and cultural intolerance—were most conspicuous among indiscriminately proreligious subjects, followed by the extrinsic subjects, the intrinsics, and the indiscriminately antireligious. Although intrinsically oriented subjects were third in rank, the Intrinsic scale nevertheless showed a low positive correlation with right-wing attitudes.

In Eugene Tate and Gerald Miller's (1971) study of United Methodist group members, the value Equality—found by Rokeach (1973) to be the one most closely related to racial tolerance of the 36 in his Value Survey—was ranked third in personal importance by the intrinsic group, fifth by the extrinsic, and ninth by the indiscriminately proreligious (Allport's order once again). The indiscriminately antireligious, unrepresented in Allport's data, ranked Equality sixth. The character of the "antireligious" is more clearly expressed in their having given the values Freedom, Mature Love, Self-Respect, Independent, and Intellectual noticeably

[1] This relation did not appear with the Unitarian-Universalists, for some of whom the theistic assumptions of the Religious Orientation Scale were undoubtedly problematical. Otherwise, Ernsberger and Manaster's findings cast doubt on Kohlberg's (1981, p. 303) claim that religious association plays no significant role in moral development. Kohlberg's scheme itself is challenged by Evangelical scholar Donald Joy (1983) for its strictly empirical character. Psychologist Carol Gilligan (1982) points out its exclusive emphasis on male development and the impersonal, masculine logic of fairness, which brings with it a corresponding neglect of the logic of relationship that is typical of women. The masculine disposition of the Evangelical Christian tradition is reflected in Joy's agreement with Kohlberg in elevating justice over care.

higher rankings than all the religious groups, and a strikingly lower ranking to Salvation, A World at Peace, Courageous, and Cheerful.

Problems with the Scales and the Typology: Although it is not difficult to find statistically significant relations between the Intrinsic–Extrinsic scales and a variety of other measures, some researchers challenge the import of such findings. Hoge and Carroll (1973) reduced to insignificance the correlation between Feagin's Extrinsic scale and two of three prejudice measures by controlling for social status and dogmatism. Feagin's subscale, they conclude, "lacks a clear definition of what is being measured; we know only that it is not tapping extrinsic religious motivation" (p. 189). The Intrinsic scale has been found in some studies to be positively related to measures of social desirability. Daniel Batson, Stephen Naifeh, and Suzanne Pate (1978) interpret this trend as evidence that correlations between this scale and positive social attitudes may be largely a product of wanting to look good. The findings of Gary Leak and Stanley Fish (1989) suggest that the Intrinsic scale's relation to desirable responding may be attributed to both conscious impression management and unconscious self-deception.

Content analysis of Allport's definitions and his items reveals that both are conceptually more complex than the simple intrinsic–extrinsic distinction would suggest (Hunt and King, 1971). Statistical analyses have led to the same conclusion. Factor analyses of the original Allport–Ross scales and of Spanish, German, and Norwegian versions have yielded as few as two factors and as many as six (Amón and Yela, 1968; Gorsuch and McPherson, 1989; Kaldestad, 1991; Kirkpatrick, 1989; Reed and Meyers, 1991; Zwingmann, 1994). The most common solution consists of three factors, an intrinsic one and two extrinsic ones, which Lee Kirkpatrick (1989) labels Social-Extrinsic and Personal-Extrinsic. Factor analyzed with other questionnaire items, however, the IE scales have proved to be less cohesive. When King and Hunt (1972b) tried to include the intrinsic–extrinsic dimension in their questionnaire, the Extrinsic scale lost some of its items and the Intrinsic scale disappeared altogether, its statements scattered among several other factors. Altemeyer (1988, p. 210), on the other hand, found the Intrinsic scale to be distinctly more coherent and reliable than the Extrinsic scale.

Given that (1) the evidence is overwhelmingly against the notion of a single intrinsic–extrinsic dimension; (2) the two dimensions themselves are complex and inadequately defined in operational terms; and (3) what is being pursued may in reality be on the level of general personality variables that are not religion-specific, Hunt and King (1971) urge the abandonment of the general labels in favor of distinctions and measures of greater specificity. Kirkpatrick and Hood (1990) likewise recommend leaving the simplistic IE framework behind in favor of more sophisticated approaches. Affirming Hunt and King's criticisms, they cite evidence that the Intrinsic scale is little more than a religious commitment scale. The Extrinsic scale, on the other hand, may not be a measure of religion at all, but a reflection of more general cognitive or personality characteristics. Furthermore, because these scales can be used meaningfully only with religious subjects, correlations with other variables are inevitably restricted. As Kirkpatrick and Hood point out, other measures of religion are better predictors, as in the case of Fundamentalism's higher correlation with prejudice than either I or E (McFarland, 1989). Among the other theoretical and methodological problems that Kirkpatrick and Hood note is the implicit, value-ladened distinction between good (intrinsic) and bad (extrinsic) religion. Dittes (1971) earlier addressed the same issue when he observed that the intrinsic–extrinsic typology has served its

users' prophetlike concern for the purity of religion far better than the scientist's quest for understanding. In spite of such criticisms, however, the typology and the Allport–Ross scales remain highly popular among contemporary researchers.

The Search for Alternatives

The search for more adequate alternatives also continues. One noteworthy proposal is offered by Bernard Spilka and his associates. In place of the intrinsic–extrinsic distinction, Russell Allen and Spilka (1967) posited two different cognitive orientations, a *committed* one said to be characterized by an abstract, complex, open, and yet coherent and serviceable way of thinking, and a contrasting *consensual* one, which tends to be literal-minded, dualistic, vague, closed, and removed from everyday activities. Subjects who, on the basis of interviews, were classified as *committed* in orientation proved to attend church more frequently and to consider themselves more religious than the *consensual* in orientation. They were also less likely to be prejudiced and more likely to be "worldminded." In their selected sample of relatively religious Protestant students, Allen and Spilka found that church attendance itself bore no significant relation to level of prejudice, nor did Wilson's (1960) extrinsic scale appear related to the committed-consensual distinction. A factor analysis of the Religious Viewpoints Scale issuing out of Allen and Spilka's research largely confirmed its two-factor structure. It was also demonstrated that persons scoring high on Rokeach's Dogmatism Scale were more likely to be consensual in orientation than committed (Raschke, 1973).

Yet because the Religious Viewpoints Scale—a factor-analytic distillation of 38 different measures—incorporates a substantial part of the Allport–Ross Religious Orientation Scale, it is doubtful that it measures something distinctive. The committed and intrinsic subscales have seven items in common, and correlations between them range from .62 to .88. Such an overlap is sufficient to consider them as measures of the same dimension. Similarly, the Consensual and Extrinsic subscales, although they share but one item, correlate around .50 (Minton and Spilka, 1976; Spilka et al., 1977). Thus in their effort to operationalize the two cognitive orientations, these researchers have inadvertently returned to Allport's original distinction, though with labels that may be more accurately descriptive. Had they succeeded in measuring the two points of view, which correspond to the contrasting dualistic and contextual–relativistic frameworks that emerged in William Perry's (1970) study of intellectual development in college students, they would have enriched the literature with a valuable research instrument.

Religion as Quest: Daniel Batson and his associates offer a more distinctive alternative. For his Good Samaritan experiment, recounted in Chapter 5, and in his continuing research on the attitudinal and behavioral correlates of religiosity, Batson fashioned his own questionnaire, the Religious Life Inventory. Two of its three scales were written to augment the Allport–Ross Religious Orientation Scale: the External scale as a measure of the extrinsic, means orientation[2] and the Internal

[2] Batson's External scale measures something other than extrinsic orientation. Five of its six items assess the degree to which personal or organizational influences have shaped the respondent's religious outlook rather than the degree of its present utilitarian character. Contrary to Batson's initial assumption that these two variables would be related, the External scale proved to be rather highly correlated with the Intrinsic, Internal, and Orthodoxy scales but not with the Extrinsic scale (Batson, Schoenrade, and Ventis, 1993, p. 172).

scale as a measure of the intrinsic, end orientation. The third, the Quest scale[3]—and the crucial one from Batson's standpoint—was designed to measure the quest orientation, "the degree to which an individual's religion involves an open-ended, responsive dialogue with existential questions raised by the contradictions and tragedies of life" (Batson, Schoenrade, and Ventis, 1993, p. 169). When Batson factor-analyzed the Religious Life Inventory in combination with the Allport–Ross scales and a doctrinal orthodoxy scale patterned after Glock and Stark's (1966), he found three factors: Religion as Means, Religion as End, and Religion as Quest.

Batson, who completed doctoral studies in both theology and psychology, had argued in an earlier, theological work (Batson, Beker, and Clark, 1973) that an experience may be identified as religious, not on the basis of its referent or intentional object (e.g., the divine), as the phenomenologists represented in Chapter 12 maintain, but to the degree it is generated by conflicts that "challenge one's reality at its core, e.g., 'What is the meaning of life, given death?' " (p. 192). Christian experience in particular is said to entail a radical perceptual shift, a fundamental reorientation outward, away from self and toward the needs of others. Furthermore, Christians are called upon to manifest this changed outlook by committing themselves to responsible action, in spite of acknowledged limitations in their view of reality, and by showing a corresponding nondoctrinaire openness to further change in accord with later experience. "The Christian hope," Batson writes, "is not in a given concept or event [e.g., Jesus' incarnation or resurrection] but in a direction, an opening movement outward in loving concern for one's neighbor" (p. 59). Hence Batson's interest in the parable of the Good Samaritan and his valorization of the reflectively critical quest orientation over the other two.

As we saw in the last chapter, Batson concludes from his Good Samaritan data that persons for whom religion is a quest are sensitive to the wishes of persons in need, whereas intrinsically religious persons seemingly act out of an insistent internal need to be helpful, pressing themselves on others even when their aid is not wanted. Similarly, when Batson, Naifeh, and Pate (1978) supplemented the anti-black scale used by Allport and Ross with a more subtle, behavioral measure of prejudice—choosing a black or white interviewer—in an effort to minimize the effects of social desirability, they found that only the Quest factor and scale were consistent in correlating negatively with both measures of prejudice. Although the End factor and Intrinsic scale were also negatively related to the questionnaire measure of prejudice, they proved to be positively (though insignificantly) related to the behavioral measure. Thus on the more subtle indicator of prejudice, the Means and End factors were indistinguishable, whereas the Quest factor correlated significantly more negatively with prejudice than either of the other two orientations. A generalized replication using a different behavioral measure of prejudice produced much the same pattern of results (Batson, Flink, Schoenrade, Fultz, and Pych, 1986).

Seeking more direct validation of their scales, Batson and Lynn Raynor-Prince (1983) found that only the Quest factor showed a significant, positive relation to cognitive complexity in the religious domain, as measured by a paragraph completion test employing sentence stems suggesting existential conflict, for example, "When I consider my own death. . . ." Both the End factor and the related Intrinsic scale showed negative—albeit, again, statistically insignificant—correlations.

[3] The Quest scale was originally called the Interactional scale, but for ease of communication, other researchers—and now Batson himself—use the shorter, more meaningful label.

Batson argues that his quest concept reintroduces three aspects of Allport's concept of mature religion that are overlooked by the notion of intrinsic religiosity: complexity, doubt, and tentativeness. He incorporated this way of being religious, he writes, "to tap what I considered a more mature, flexible type of religiosity than the other two" (Batson, 1976, p. 207).[4] Batson has found support for this early conviction in the research that he and his associates have subsequently carried out. The generally well-regarded intrinsic orientation, conclude Batson, Schoenrade, and Ventis (1993), is marked by a freedom from worry and guilt that is bought at the high price of uncritical bondage to one's religious beliefs; by thinking that tends to be simplistic and rigid; and by a self-centered desire to *appear* tolerant, sensitive, and loving. In contrast, the skeptical quest orientation appears to be associated with freedom from the bondage of doctrine but not from existential concerns; with flexible open-mindedness, competent self-reliance, and self-actualization; and, at the crucial level of social consequences, with tolerance of others and sensitivity to their needs. If one were to judge which of these orientations constitutes "true religion," Batson and his associates intimate, it would be quest, not the intrinsic orientation (pp. 375, 198, 288, 364, 189).

Assessing the Quest Orientation: Batson's critics, who are often advocates of the intrinsic orientation, have challenged his work on both methodological and conceptual grounds. They have noted with concern, for example, the low internal consistency of the original, six-item Quest scale. This problem has now been addressed by the development of a more reliable, twelve-item scale, which incorporates five of the original items (see Table 6.5). With an internal-consistency reliability hovering around .80 and a correlation with the original scale of about .86, the new scale appears to be measuring the same thing but much more reliably (Batson and Schoenrade, 1991a, 1991b).

Factor analysis confirms what is apparent from the item content: the Quest scale is not unitary, but consists of two or more factors, including doubt. In one series of analyses, the Doubt factor proved to be negatively correlated with the Intrinsic scale and positively correlated with the Extrinsic scale. It also showed a very slight tendency to be associated with maladjustment (Watson, Morris, and Hood, 1989). Spilka and his associates have found the Quest scale to be related negatively to a standard measure of social desirability and positively related to measures of trait anxiety and religious questioning (designated "religious conflict") (Spilka, Kojetin, and McIntosh, 1985; Kojetin, McIntosh, Bridges, and Spilka, 1987). They infer that high scores on quest reflect personal distress, a conclusion that accords with Mariano Moraleda's (1977) finding that, among Spanish adolescents, religious ambivalence and doubt are associated with higher levels of anxiety.

Doubt is one of the facets of Allport's conceptualization of mature religion that Batson and his colleagues say have been neglected by the Intrinsic scale. Yet as Michael Donahue (1985) notes, Allport ascribes to doubt an interim role in the maturation of the religious sentiment, a factor that gradually fades as increasing commitment and its fruits strengthen the individual's faith. The instability of the

[4] Batson and Ventis (1985, p. 400) more recently deny that they represent the quest outlook as more mature. To others, however, their valorization of quest is unmistakable. Derks and Lans (1986) wonder whether "Batson's *psychological* theory of religious attitudes [isn't] essentially a *theological* theory of true Christianity" (p. 204). They identify Batson's theory with social gospel theology, which we know from Chapter 2 was a major factor in the rise of the psychology of religion at the beginning of the twentieth century.

Table 6.5 **BATSON'S TWELVE-ITEM QUEST SCALE, ARRANGED ACCORDING TO THE SCALE'S INTENDED THREE FACETS**

	Complexity
1.	I was not very interested in religion until I began to ask questions about the meaning and purpose of my life.
2.	*I have been driven to ask religious questions out of a growing awareness of the tensions in my world and in my relation to my world.
3.	My life experiences have led me to rethink my religious convictions.
4.	*God wasn't very important for me until I began to ask questions about the meaning of my own life.
	Doubt
5.	*It might be said that I value my religious doubts and uncertainties.
6.	For me, doubting is an important part of what it means to be religious.
7.	I find religious doubts upsetting. (rating reversed in scoring)
8.	*Questions are far more central to my religious experience than are answers.
	Tentativeness
9.	As I grow and change, I expect my religion also to grow and change.
10.	I am constantly questioning my religious beliefs.
11.	*I do not expect my religious convictions to change in the next few years. (rating reversed in scoring)
12.	There are many religious issues on which my views are still changing.

*Items taken from the six-item Quest scale
Source: Adapted from Table 1 in Batson and Schoenrade, 1991b.

attitude of doubt is underscored in a classic study by Robert Thouless (1935). When Thouless asked subjects to indicate the degree of certainty with which they believed or disbelieved 40 statements on religious and other subjects, he found a marked tendency either to hold or to reject the belief in question with a considerable degree of conviction. This "tendency to certainty" was significantly greater for religious propositions than for nonreligious ones. Doubt or uncertainty, Thouless concludes, is an uncommon or at least a short-lived state for most people. Rarer still, he notes, is the capacity to adopt and sustain the laudable attitudes of tentativeness or of partial assent, which Batson's scale purports to measure but in fact may not.

Hood and Morris (1985) propose that the quest and intrinsic orientations be understood as stages of religious faith. The quest perspective, they say, is characteristic of persons still "in process" whereas the intrinsic orientation describes persons who have found satisfying answers to the existential questions. Consistent with this interpretation, the Quest scores of a group of 205 Presbyterians ranging in age from 11 to 83 tended to decline with age ($r = -.19$, $p < .01$), whereas scores on the Intrinsic scale generally increased ($r = .29$, $p < .001$) (Watson, Howard, Hood, and Morris, 1988). The men in Marvin Acklin's (1985) study showed a similar decline in Quest scores over the years (age and quest correlated $-.30$, $p < .01$; see Figure 6.4). Furthermore, their Quest scores proved to be negatively correlated ($r = -.31$, $p < .01$) with a short version of the Washington University Sentence Completion Test of Ego Development, a measure of emotional and cognitive maturity. The comparable correlations for the women were slightly positive but insignificant. Acklin concludes that Batson's Quest factor measures a "precommitment or transitional religious outlook"; contrary to what Batson had

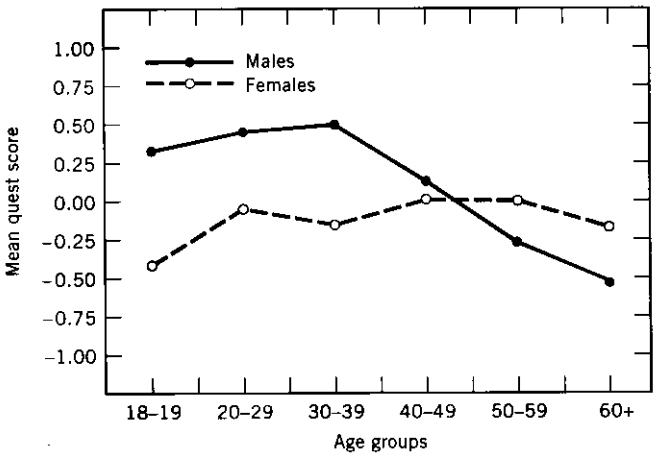


Figure 6.4. Quest factor scores by age group (N = 20 for each of the six groups.)

suggested, it "does not appear to be a measure of religious maturity at all" (pp. 53, 60).

There are even doubts that the quest orientation, at least as it is represented by Batson's Quest scale, is properly called religious. As Donahue (1985) points out, the Quest scale tends not to correlate with other measures of religiousness. This is not surprising when such measures are limtied to literal belief and other conventional expressions of faith, as has often been the case. But it is odd that Batson's scale showed no distinctive relation to Hutsebaut's (1994) historical relativism factor, which is partially defined in terms of quest. And whereas Altemeyer and Hunsberger's (1992) Quest scale shows a high negative correlation (−.79) with their Religious Fundamentalism scale, the Batson Quest scale as augmented by McFarland (1989) yielded a far more modest one (−.36).

Although certainly instances of the quest attitude can be found in the world's religious traditions, by itself it is not likely to foster deep religious experience. We may recall that, of the seven worship services that Goodwin Watson (1929) presented for evaluation by his young subjects, the one that focused on the quest for meaning in a vast universe received the lowest ratings. Without some sense of an objective reality, James Pratt (1920) would say, a worshiper is not likely to experience subjective effects. Nor is the quest orientation likely to inspire the saintly virtues by which James (1902) found religion to be vindicated, as Batson and Raynor-Prince (1983) themselves appear to acknowledge. For Thouless, the attitude of tentativeness is not a religious orientation but an intellectual one, the attainment of which he sees as one of the goals of a liberal education.

Whatever the Quest scale measures, other investigators are likewise combining it with the IE scales to study social attitudes. Their findings are often similar to what Batson and his associates report. In one study, for example, the Quest scale was the only one of the three to be significantly and positively related to level of moral judgment (Sapp and Jones, 1986). In another investigation, which compared rankings on the Rokeach Value Survey with Batson's three ways of being religious, the quest orientation correlated more highly with the ranking of Equality—the value most revealing of racial tolerance and such liberal attitudes as approval of church involvement in social and political issues—than with any of the 17 other

terminal values. Neither the means (extrinsic) nor the ends (intrinsic) orientation showed any relation to the ranking of Equality (Luce, Briggs, and Francis, 1987). Of the seven other values to which the quest orientation showed a significant relation, four are reported by Rokeach (1973, p. 101) to be either positively or negatively related to racist attitudes. For all four, the correlation with the quest orientation is in the direction of antiracism. In contrast, two of these four values—Salvation and Family Security—correlate in the racist direction with the intrinsic orientation. Once again, then, a less transparent indicator of tolerance and concern for others proved to be positively related to quest but either unrelated or negatively related to end or intrinsic religiosity.

In a study by Sam McFarland (1989), an augmented Quest scale proved to be a uniquely consistent predictor of nondiscrimination among university undergraduates. McFarland correlated his general measure of discrimination and its four components—against blacks, women, communists, and gay and lesbian persons—with the Extrinsic, Intrinsic, and Quest scales as well as a Fundamentalism scale and the extrinsic subscales identified by Kirkpatrick (1989). He found that for males the extrinsic measures tended to be positively related to discrimination of various kinds and the Intrinsic scale to be unrelated. For females the Intrinsic scale was *positively* related to all forms of discrimination except toward blacks, whereas the extrinsic subscales tended to be *negatively* related. When these correlations were recalculated after removing the specific contribution of Fundamentalism—which was positively related to every form of discrimination—all but one of the intrinsic correlations were reduced to insignificance. The extrinsic correlations remained largely unchanged. The quest measure, on the other hand, was significantly and negatively correlated with all measures of discrimination, for males and females, before and after the effects of Fundamentalism were partialed out. Altemeyer and Hunsberger (1992) found that both McFarland's Quest scale and a 16-item Revised Quest Scale of their own were negatively correlated with right-wing authoritarianism and four measures of prejudice. Fundamentalism, in contrast, was positively associated with all five of these variables.

Concluding Thoughts on Quest: Whatever the present conceptual and psychometric shortcomings of the Batson Quest scale and its relatives, these measures are finally creating a place in the correlational literature for the liberal religious outlook. Nearly half a century ago, Pratt (1950) argued that, for the modern, honest, and thinking individual, worship can only be understood as "an active search," "an ONTOLOGICAL venture"; religion, he said in further anticipation of Batson, "has always to do with the practical question of destiny," of personal and collective fate, and it "almost inevitably gives one a larger horizon than one would otherwise have" as well as "a certain *depth*." Pratt, too, discerns at the heart of Christian experience "the impulse to helpfulness and to service" (pp. 59, 62, 66, 73, 122).

In like spirit, British educator Ronald Goldman (1967) observes that the general temper of our time is one of "search and questioning." Religious education in particular must become an open-ended "exploration of experience in depth" rather than the teaching of infallible truths. It must be an education for uncertainty and personal choice (pp. 7, 14). Goldman's (1964, 1965) own research on religious thinking and his subsequent recommendations for religious education suggest that, as a sustained posture of tentativeness that allows for a life of commitment, the quest orientation is indeed a more mature stage than the settled orthodoxy tapped

by the Intrinsic scale. William Perry's (1970) research on the intellectual development of college students and James Fowler's (1981) on the stages of faith development, compel the same conclusion. As Thouless and Fowler found, however, such a posture is rarely attained. We may suspect that it is not well represented by Batson's questionnaire items, none of which suggests the movement toward a goal that is implied in the metaphor of quest (Wulff, 1992). It may be, however, that the ideas of a path and destination are incompatible with today pluralistic religious searching (Woodward, 1994). Clarifying and assessing the quest orientation is surely one of the psychology of religion's most urgent tasks.

RELIGION AND MENTAL HEALTH

The Intrinsic, Extrinsic, and Quest scales were all initially developed to clarify religion's relation to social attitudes. Today they are more commonly used for exploring piety's connection to personal adjustment. This redirection of effort reflects a general shift in the research on the consequences of piety, from an interest in its social correlates to an emphasis on its implications for health, both physical and mental. Physical health is easily defined and assessed, usually in terms of longevity, the absence of disease, or self-ratings of overall health. "Mental health," a widely used borrowing from the medical realm, has a range of connotations that suggest not merely the absence of psychiatric symptoms but also the presence of such positive qualities as self-acceptance, autonomy, personal integration, and self-actualization (Jahoda, 1958). It is correspondingly more difficult to operationalize.

Correlates with Physical Health

Correlational evidence indicating that religiousness in one form or another is related to physical health dates back to the early nineteenth century. The great bulk of this research, however, consists of epidemiological studies in which religiosity was included as just one variable among many and measured by the simplest of indicators, usually religious affiliation. Moreover, significant findings in relation to religion were commonly left undiscussed. Thus until recently these rather numerous findings were largely unknown (Levin and Schiller, 1987).

In their review of well over 200 such studies, Jeffrey Levin and Preston Schiller (1987) report a number of interesting trends. The prevalence of certain diseases varies significantly from one religious group to another, sometimes the result of genetic inheritance or of particular health or dietary practices, such as the vegetarianism of Seventh-Day Adventists. And contrary to Galton's report, various studies of religious professionals show some health advantages for certain classes of them, though at least one study subsequent to Galton's confirms a higher mortality rate for missionaries than for lay people. In some studies but not others, frequency of religious attendance, self-rated religiosity, and other such measures are positively related to self-rated health and decreased mortality from various specific causes.

Levin and Schiller conclude their review of this extremely diverse literature by calling for the development of a paradigm for guiding new, more adequate research on how and why religiousness affects physical health. They also urge the thoughtful adoption of more subtle measures of religiosity, especially indicators of Allport's intrinsic type. To promote such research, Levin and his associates are working to establish a new field that they call the epidemiology of religion (Levin, 1994a; Levin, 1994b; Levin and Vanderpool, 1987). Meanwhile, Herbert Benson (1996)

and his associates continue to document the positive effects on physical health of the "faith factor"—the combination of the relaxation response with "remembered wellness," Benson's preferred term for what others call the placebo effect.

Mental Health and Religion: A Complex Relation

Research on religion and mental health, on the other hand, needs no such advocacy, for it is today widely pursued using a variety of measures and research techniques. Given that both variables—mental health as well as religion—are conceived as multidimensional, the possibilities for interrelationship seem virtually inexhaustible. In an early effort to assay the literature on religion and mental health, Bernard Spilka and Paul Werme (1971) conclude that the inconsistencies they found in the empirical findings testify not to the absence of a relationship, as some had inferred, but to a complex association that requires thoughtful redefinition and assessment of each term. They suggest four different ways that religion may be related to mental health: (1) It may become an expressive outlet for existing mental disturbance. (2) It may suppress symptoms and resocialize the individual, encouraging more conventional and socially acceptable forms of thought and behavior. (3) It may provide a refuge or haven from the life stresses that might otherwise precipitate mental disorder. (4) And it may provide resources for the development of broader perspectives and the fuller realization of individual capabilities.

Ecclesiogenic Neuroses

Religious association and teachings may also present hazards, as Spilka and Werme acknowledge. Struck by how many of his patients with sexual disorders came from conservative Christian circles where sex was a forbidden topic surrounded by secrecy, prohibitions, and threats, Berlin gynecologist Eberhard Schaetzing (1955) coined for these disturbances the term "ecclesiogenic neuroses"—that is, church-caused disorders. The term was soon adopted by others, who expanded its meaning to encompass a variety of other disturbances likewise thought to have roots in religious teachings and practices. In his handbook on the prevention of suicide, Berlin physician and theologian Klaus Thomas (1964) estimates that about 10 percent of all neuroses are ecclesiogenic. Judging from his own work with Christian patients, psychoanalyst Heinrich von Knorre (1991) thinks this estimate is rather high, but he does confirm that sexual disorders are common among Christians. Psychiatrist Samuel Pfeifer (1993), on the other hand, objects to the term ecclesiogenic because of the oversimplified causal model he says it suggests. Piety alone, he asserts, does not make anyone sick. In every case of religion-related disorder, he claims, religious elements encounter a "neurotic" personality existing in the broader context of the stresses of human life (p. 110).

Certainly religious elements interact with personal disposition. The unusually high incidence of schizophrenia among the followers of Bratslav Hasidism, for example, reflects in part the appeal that this messianic, ultraorthodox sect in Israel and New York has for isolated and paranoid individuals. But it is participation in the sect's unusual practices, such as praying all night at the tombs of spiritual leaders, that finally brings about the psychotic break (Witzum, Greenberg, and Buchbinder, 1990).

Patriarchal Religion and Child Abuse: The ecclesiogenic label, with its emphasis on pathological influences in religious tradition, may nevertheless be appropriate on occasion, as recent work on religion and child abuse suggests. Historian Philip Greven (1991) documents, through religious books on child discipline and

the autobiographies of notable religious personalities, that devoutly Christian parents have long physically abused their children under the conviction that the child's will must be broken if the child is to live in conformity with the superior will of God. Recent research indicates that Fundamentalists above all others are today likely to use corporal punishment (Ellison, 1996; Grasmick, Bursik, and Kimpel, 1991; Wiehe, 1990), and they are also likely to express views on childrearing similar to those of parents with a history of abusing their children (Neufeld, 1979). The apocalyptic religious world-view that is common among those who advocate such corporal punishment may be understood, says Greven, as an expression of "the nuclear core of rage, resentment, and aggression" that remains from their own childhoods of abuse (p. 206). It is perhaps not surprising, then, that a fundamentalist upbringing is a common factor in patients with multiple personality disorder (Bowman, 1989).

It would likewise be inappropriate to ascribe to individual neurosis the stress that countless women have suffered in growing up in religious cultures dominated by patriarchal imagery and relationships. According to an interview study carried out in the Netherlands by Annie Imbens and Ineke Jonker (1985), various elements in the Christian tradition in particular have contributed to the occurrence of child sexual abuse, first serving to justify incestuous acts in the eyes of the perpetrators and then making it difficult for the victims to seek and obtain help. These elements include a stern patriarchal image of God and the presumption by men that they are the ones who are called by God. Women are correspondingly subordinated, derogated, suppressed, and silenced, both in biblical texts and religious practice. Adhering to the norms prescribed for them by the Christian tradition, girls become easy prey for sexual abuse by male family members. They are then blamed for the transgression and held responsible for disrupting the family if the violation becomes known. Traumatized both sexually and religiously, all 19 women in Imbens and Jonker's subject group eventually left the church. The increased risk for sexual abuse that is associated with conservative Christian beliefs and the high rates of religious defection among women who suffer such abuse in childhood have also been observed by others (D. Elliot, 1994; Russell, 1986; Taylor and Fontes, 1995). Growing awareness of these trends is challenging counselors, religious and secular, to find ways to respond effectively to abuse in religious families (Horton and Williamson, 1988).

In many Christian churches today, traditional patriarchal language and values are being gradually rethought and systematically replaced. Meanwhile, women are coping in a variety of ways, commonly by noting improvements over the past or contrasting their situations with less favorable ones in other religious traditions and secular contexts. Their strong emphasis on relational values rather than cognitive and hierarchical ones likewise serves to minimize the effects of inequality (Ozorak, 1996). While seeking a larger role in the functioning of religious organizations, women are also claiming the prerogative of renaming the sacred, which they find not in some distant space but in the palpable realities of everyday life (Gray, 1988). Largely excluded from the interreligious dialogue that has sprung up in the face of the new religious pluralism, as religious leaders and scholars explore ways of maintaining their own cherished traditions while yet respecting those of others, women have undertaken a gender-sensitive dialogue of their own (O'Neill, 1990). And as they work to redefine the issues of spirituality for women, many are leaving the familiar traditions behind in their quest for spiritual resources—including the image of the Goddess—that are powerful and transforming for them (Christ, 1995;

Women ritually share milk and honey at a 1993 conference in Minneapolis dedicated to "re-imagining" traditional Christian concepts in terms of women's experience. Drawing 2000 women from major Protestant denominations, the conference created a furor within the United Methodist and Presbyterian (U.S.A.) churches, some of whose officers were among the conference planners (Steinfels, 1994).

King, 1993; Pirani, 1991). Healing—physical as well as spiritual, including relief from the pain suffered in growing up in a patriarchal world—is assumed in feminist spiritual thought to be needed by all women (Eller, 1993, p. 109).

Conservative Religion and Mental Health: Acting to counter such changes is a worldwide resurgence of militant fundamentalism, a response to perceived threats in modern, secular culture, including the challenges of the feminist movement (Hawley, 1994). There has been a corresponding surge in scholarly investigation, much of it under the sponsorship of The Fundamentalism Project at the University of Chicago (Marty and Appleby, 1991, 1992). Empirical research on fundamentalist and other conservative religious groups is also growing, some of it aimed at the question of the relation of conservative religious views to mental health.

One study, for example, directly tests Schaetzing's claim that "dogmatism of the church" causes neurosis. Hartmut Spring and his associates (1993) assessed the level of religious anxiety and depressivity—the indicators of ecclesiogenic neurosis proposed by Thomas (1964)—in two Catholic congregations, one long situated in a heavily Catholic milieu and the other recently established in an area with a minority of Catholics. The assumption that the first congregation, designated "traditional," was higher in dogmatism of the church than the second, "pluralistic" congregation was supported by the finding of higher mean scores on orthodoxy and authoritarian submissiveness among members of the first group. As predicted, members of the traditional congregation proved to be significantly higher on religious anxiety (e.g., in relation to hell, the Devil, feelings of guilt, punishment from God, and so on) than members of the pluralistic group. Although members of the

two congregations did not differ from each other in depressivity, both groups scored distinctly higher in comparison to the normative population. The heightened depressivity scores thus also lend support to the ecclesiogenic hypothesis. On the other hand, as Spring and his colleagues acknowledge, one could account for these data another way: persons already high in depressivity may be drawn to the Church for help.

From his own experience as a former conservative evangelical, the eminent biblical scholar James Barr (1980) affirms a causal relation in accord with the ecclesiogenic hypothesis. He maintains that Fundamentalist sentiments induce in open and happy young people the strained and suspicious outlook of the authoritarian mind. Historian and psychoanalyst Charles Strozier (1994), in contrast, found that all of the Fundamentalists he interviewed had histories of trauma, which they eventually learned to talk about in the rhetoric of literal Christian belief. The "broken narratives" for which they sought healing in born-again experience became for Strozier the defining characteristic of Fundamentalists. Although they found consolation and hope in the dramatic mythic imagery of the anticipated apocalypse, they remained extraordinarily preoccupied with personal guilt and shame. Themes of real or symbolic death were also conspicuous in their thinking, whether they were reflecting on their personal pasts or anticipating the imminent violent destruction that they believed is necessitated by human evil. Life in the present is for them frightening and fragile, a constant struggle.

Strozier maintains that Fundamentalism "can only be understood as a kind of collective illness in our contemporary culture"; there is, he adds, "generally something unsteady" about its followers (p. 3). Yet he declines to describe the lives of Fundamentalists as pathological. Categories of pathology are meaningless, he says, for understanding a mass movement that involves roughly a quarter of the United States population. Of Strozier's respondents, all of whom seemed to be typical members of the churches in which he worked, none was perceived by others as odd or disturbed.

Empirical researchers agree that Fundamentalists do not score higher than other groups on measures of pathology, on the average, as long as the measures are religiously neutral and social class is held constant (Hood, Morris, and Watson, 1986; Stark, 1971). In regard to other measures of mental health, however, the findings are less consistent. On the positive side, for example, Christian, Jewish, and Muslim fundamentalists proved as a group to be much more optimistic than members of more liberal Christian and Jewish groups, in accord with the greater hopefulness of the religious messages they had heard and read (Sethi and Seligman, 1993, 1994). Fundamentalism has also been found to be a positive indicator of marital happiness (Hansen, 1992).

On the negative side, a group of certified family therapists rated Fundamentalist families as significantly less healthy than nonfundamentalist families on three of eight factors on a Family Health Scale and more healthy on only one. Fundamentalist families had clearer expectations of how family members were to behave in relation to each other but they were also less emotionally close, less flexible in dealing with change, less likely to encourage members to assume responsibility and exercise their own judgment, and less likely to show caring without smothering (Denton and Denton, 1992). Fundamentalism has also been identified as a common feature of families with adolescent members exhibiting "conversion," which here refers not to religious conversion but to a disabling physical disorder for which no organic cause can be found (Seltzer, 1984).

Insignificant differences have been found as well. Born-again Christian students in one study were no more likely to report "great stress" than others (Schafer and King, 1990). In another study, Fundamentalists from the Southern Baptist Church proved not to differ in level of ego development from nonfundamentalist members of the United Methodist Church (Weaver, Berry, and Pittel, 1994).

Religious Orientation and Mental Health

As we noted earlier, inconsistencies of this sort are characteristic of the entire literature on religion and mental health. In a meta-analysis of the literature, Allen Bergin (1983) found that 47 percent of the studies he reviewed reported a positive relationship, 23 percent a negative one, and 30 percent no relation at all. In a later review of the literature, John Gartner, Dave Larson, and George Allen (1991; Gartner, 1996) suggest that some of these discrepancies can be accounted for in terms of the types of measures that have been used. Most studies reporting positive associations assess the mental health variables in terms of directly observable behavior (e.g., using or not using drugs). In contrast, the preponderance of studies reporting negative relationships assess mental health by using questionnaires purporting to measure one or another hypothetical intrapsychic construct (e.g., rigidity or self-actualization). These questionnaires, Gartner and his colleagues suggest, are not only less reliable and valid but also more likely to be value biased. Weighing in on the side of behavioral measures, of mental health as well as religion, they conclude that the bulk of the evidence suggests that religion is associated with positive mental health.

Most researchers remain committed to the more subtle and broadly significant constructs, including the dimensions of religiosity. Much of the research today on the mental-health correlates of religion uses the Allport-Ross Religious Orientation Scale, with the expectation that the Intrinsic scale will be positively associated with mental health and the Extrinsic scale negatively associated. Findings have tended to support these predictions. The Intrinsic scale, for example, has proved to be positively associated with life satisfaction (Zwingmann, 1991; Zwingmann, Moosbrugger, and Frank, 1991), psychological adjustment (Watson, Morris, and Hood, 1994), self-control and better personality functioning (Bergin, Masters, and Richards, 1987), self-esteem (Nelson, 1990; Ryan, Rigby, and King, 1993), an internal locus of control (Jackson and Coursey, 1988; Kahoe, 1974a; Stewin, 1976), purpose-in-life scores (Bolt, 1975; Crandall and Rasmussen, 1975), spiritual well-being (Mickley, Soeken, and Belcher, 1992), adjustment and morale in the elderly (Koenig, Kvale, and Ferrel, 1988; Van Haitsma, 1986), but also proneness to guilt (Chau, et al., 1990; Richards, 1991). Negative associations with the Intrinsic scale have been found for anxiety and death anxiety in particular (Bergin, Masters, and Richards, 1987; Powell and Thorson, 1991), neuroticism (Chau et al., 1990), depression (Dörr, 1987, 1992; Genia, 1996; Nelson, 1989), impulsivity (Robinson, 1990), and maladaptive narcissism (Watson, Morris, Hood, and Biderman, 1990). When correlations with the Extrinsic scale are significant, they tend to show the opposite pattern, such as positive correlations with anxiety (Bergin, Masters, and Richards, 1987), depression (Genia and Shaw, 1991), and fear of death (Bolt, 1977; Kraft, Litwin, and Barber, 1987) and negative ones with autonomy (Tisdale, 1966) and internal locus of control (Kahoe, 1974a; Park, Cohen, and Herb, 1990; Strickland and Shaffer, 1971).

The Quest scale is less often used in studies of mental health than the IE scales. It also less frequently yields significant findings. If any trends are emerging, one of

them may be positive correlations with measures of personal distress, a relation we earlier noted in the research of Spilka and his associates. Like these researchers, Vicky Genia (1996) found a significant negative correlation between quest scores and social desirability. Using Batson's 12-item Quest scale, she also found a low positive correlation with depression (.16) and a negative one with a measure of self-esteem (−.16), though both correlations were reduced to insignificance when she controlled for social desirability. Richard Ryan, Scott Rigby, and Kristi King (1993) found comparably low positive correlations between the quest orientation and measures of anxiety and depression, but in their study none was large enough to be statistically significant. Consistently significant and positive correlations have been found, on the other hand, between the quest orientation and fear of death or negative perspectives on death, on the one hand, and with openness and cognitive complexity on the other (Batson, Schoenrade, and Ventis, 1993, pp. 282–285). Yet even these correlations are modest.

Religion as a Way of Coping

While many researchers remain enthusiastic about the IE distinction if not also quest (Gorsuch, 1988), others are developing more complex and potentially fruitful theoretical frameworks. One of the most promising is offered by Kenneth Pargament and his associates (Pargament, 1990, 1996; Pargament and Park, 1995). Rather than conceiving of religion as either a means or an end, as in the traditional IE construction, Pargament (1992) proposes that religion be conceived as "a general disposition to use particular means to attain particular ends in living" (p. 211). A means-*and*-ends analysis of religion, he says, brings out aspects of the religious life that are overlooked in the literature dominated by the three religious orientations, including the social aspects that address the need for intimate association and the diverse feelings, beliefs, and practices that are religion's content. For such an approach, Pargament has adopted the term *coping* as it is used in the clinical and social-psychological literatures. The coping model emphasizes the constructive role that religion can play within the complex, ongoing process by which people try to comprehend and deal with the various personal and situational problems that come into their lives.

Each of the elements in the coping process, says Pargament (1990), requires careful assessment. The problematic event or situation, which may be anything from a minor irritation to a major life change, must be understood according to its subjective meaning for the individual. How a person experiences and appraises a situation, and then estimates his or her own ability to handle it, will naturally play a large role in shaping the outcome. The coping activities themselves may take a variety of forms, some directed at the situation and others aimed at the person's emotional response. Coping efforts may be genuinely constructive and rational, but they can also be ineffective or even self-defeating. Whatever its forms, coping can lead in time to reappraisals of the precipitating event. The possible outcomes of coping are similarly diverse: they may be situational, psychological, social, or physiological in nature; they may be essentially positive or negative, or a mixture of gains and losses; and they may range from immediate or short-term outcomes to much longer-term ones, perhaps of a very different character.

Religion may enter into the coping process in a variety of ways. The *critical event* may itself be religious, such as a conversion or mystical experience or some insight or realization from reading sacred scripture; or the event may be religiously framed, as in the case of an interpersonal experience within one's congregation or a life

transition that is marked by a religious ceremony, such as confirmation, bar or bas matzvah, a wedding, or a funeral. *Appraisals* of the situation may also be religious in character. When good or bad things inexplicably happen, they may be interpreted as rewards or punishments from God, or as divinely ordained opportunities for learning certain virtues. They may also be thought to have occurred without God's consent, and to be an occasion of grief in the heavenly realm as well. The *coping activities* set in motion may likewise be religious. One may seek advice or consolation from a member of the clergy or some other religious associate, if not also from God. One may show a variety of cognitive, emotional, and behavioral responses, such as deciding what lesson God intended through this event, having certain feelings toward God, trying to lead a more virtuous life, or promoting justice in the wider world. The *outcomes*, finally, can be religious, including changes in religious beliefs, feelings, and participation if not also in the overall direction of one's spiritual life. Pargament (1990) notes that association with a religious tradition may multiply the number of available resources for coping, but various traditional teachings and practices may also impose serious constraints.

In underscoring the complexity of the process of coping, Pargament's theoretical framework offers a major challenge to other researchers on religion and mental health. In place of the simple correlational procedures that many use, his model calls for the inclusion of multiple variables, including moderating variables that determine when and how the others are related; the use of statistical procedures sensitive to curvilinear relationships; the employment of qualitative techniques for assessing the individual's subjective construction of events and idiosyncratic efforts at coping; and a recognition that the complex outcomes of the coping process may entail tradeoffs, a combination of positive and negative gains. Associations with a strict religious group, for example, may provide feelings of intimacy and self-worth at the same time that it discourages the development of certain skills and fosters intolerance of others' views.

While thus recognizing the dark side of religion, Pargament's coping model is fundamentally a positive one, for coping implies some degree of success in adjusting to life's circumstances. The model has the great virtue of viewing religion within its living context: the ongoing lives of individuals as they search for significance in their day-to-day experiences. Multifaceted as this model is, it provides the foundations for a broad and sustained research program. Like other emerging theoretical frameworks in the psychology of religion, such as the perspectives of attachment theory (Kirkpatrick, 1992, 1995a; Oksanen, 1994) and general attribution theory (Proudfoot and Shaver, 1975; Spilka, Shaver, and Kirkpatrick, 1985; Spilka and McIntosh, 1995), it offers a new focus and impetus for the correlational approach and the prospect of a more vital and dynamic field.

EVALUATION OF THE EXPERIMENTAL AND CORRELATIONAL APPROACHES

Experimental and correlational techniques tend to inspire one or the other of two extreme responses. Some persons, including many of the researchers whose work we have reviewed in the last two chapters, assume that measurement, laboratory controls, and statistical analysis are the only sure means for gaining scientific knowledge of human experience and behavior. They are thus disdainful of the more subjective, interpretive methods used by the humanistic disciplines. Humanistic

scholars, on the other hand, are commonly mystified and alienated by the application of quantitative and experimental procedures in the human sphere. Peculiarly sensitive to the nuances and ambiguities of human language and accustomed to using textual and other historical materials, these scholars are deeply skeptical about the reduction of any human experience to rating scales and averages.

Wisdom would seem to lie in a more moderate position. There is no question that correlational and experimental procedures can be—and frequently have been—misused, yielding results that have little meaning or relevance. Thoughtfully applied, however, objective research methods can be highly serviceable. As Marie Jahoda (1977, p. 155) argues, experiments "present the strictest test of assumptions that has yet been devised. With all their pitfalls, for this purpose no better technique exists, as long as the phenomena under investigation are not simplified out of existence by the experimental manipulation." Galton, we read earlier, considered statistics an extraordinarily powerful tool for disentangling complex relationships—as long as these methods are "delicately handled" and "warily interpreted."

CAN RELIGION BE BROUGHT INTO THE LABORATORY?

At various points in the last two chapters we have noted some of the pitfalls and problems of the methods of the objective approach, especially as they are applied to religion. It is always difficult to bring real-world phenomena into the experimental laboratory. Phenomena as delicate and complex as religious ones are even more problematical. The most dramatic success in the laboratory approach to religious experience has been achieved with the use of psychedelic drugs. In addition to the legal and ethical complications, however, there is still the question of how authentic drug-induced experiences are. The debate has been vigorous (see Osmond and Smythies, 1959; H. Smith, 1964; Zaehner, 1957, 1972), but particularly informative are reports from those who have experienced both states. They claim that drug-induced experiences are not as profound and meaningful as those achieved through traditional religious means (Jordan, 1971). Moreover, it is doubtful that, in comparison to religious experiences without drugs, psychedelic states have as much "faith-filled carryover." Indeed, Houston Smith suspects that psychedelic experiences may as readily abort a religious quest as further it (Smith, 1967, p. 144; 1976).

Apart from psychedelic drugs, which obviously raise complex issues and are at any rate almost impossible to obtain under current law, is true experimentation possible in the psychology of religion? Huxley (1961) cites research on fasting and sleep deprivation as evidence that ascetic practices, like drugs, are a means of changing body chemistry toward the production of transformed religious awareness. Might ascetic practices, along with other modes of bodily manipulation, also serve as a means of continuing experimentation? Contrary to Huxley's report, the research findings of Ancel Keys and his colleagues (1950), summarized in Chapter 3, included no "visionary experiences." Yet the Keys study was not designed to elicit them. Given a different set and setting, we can easily imagine a different outcome. It is doubtful, however, that so dangerous and personally costly an experiment would be undertaken for the purposes we are considering.

The hypnotic investigations of Bernard Aaronson and the quasi-hypnotic explorations of Houston and Masters, on the other hand, seem genuinely to offer promising and practical techniques for more carefully controlled experimentation, at least with the minority of subjects capable of "deep trance" phenomena (Gib-

bons and DeJarnette, 1972). Less dramatic, perhaps, but already more or less in accord with accepted research practice are the numerous laboratory studies of meditation. If there are doubts abroad about experimentation in the psychology of religion, some researchers obviously do not share them.

According to Wilhelm Koepp (1920), a fundamental paradox undermines any experimental approach to religion undertaken for scientific purposes. Religious experience, he argues, can only be called forth for its own sake. When it is asked to serve a scientific aim, it is inevitably transformed. The phenomenon loses its essentially religious character, leaving only aesthetic and other secondary aspects. Virginia Hine (1969) illustrates this principle in her own observations of glossolalia. The strongly positive emotional states usually accompanying the experience of speaking in tongues, and often persisting long after the utterance itself, are likely to be wholly absent when the Pentecostal subject is working with a scientific observer. It is understandable, therefore, that many profoundly religious persons refuse to cooperate with the would-be experimenter. Yet, turning to the "irreligious" is no solution, according to Koepp, for they are incapable of meeting "the first requirement of a religious experiment, actual entry into the religious sphere" (p. 58).

Agreeing that a strictly experimental psychology of religion is impossible, Batson (1977) holds out for quasi-experimental methods such as he himself has employed. The same label is applicable to Harold Burtt and Don Falkenberg's (1941) study, in which they demonstrate the persuasive effect that knowledge of majority or expert opinion can have on religious attitudes and beliefs, as well as to Jean-Pierre Deconchy's (1971, 1980) similarly conceived but more highly elaborated investigations in France. Quasi-experimental techniques do not entail the same degree of control that characterizes the ideal laboratory experiment. Yet they may still elicit the critic's doubts about the genuineness of the religious phenomena under evaluation, the adequacy of the various measures employed, and the ethics of artificial intervention and deception.

CAN RELIGIOUS FAITH BE MEANINGFULLY MEASURED?

Experimental and correlation psychologies of religion are founded on the assumption that fundamental aspects of religious faith are measurable. Those who disagree do not deny that some expressions of faith—attendance at religious services, most obviously—can be quantified. What they do doubt is that an individual's religiousness can be evaluated numerically or categorically without gravely misrepresenting it.

The most common means used to assess piety is the questionnaire. Early forms of this device almost always required respondents to write lengthy answers. The problems were legion. To answer most of the questions adequately required an exceptionally high level of cooperation, memory, introspective capacity, self-knowledge, vocabulary, and precision in description. Ever present were the problems of suggesting answers in the very asking of the questions; the unknown principles by which the subject decided what and how much to include; the tendency to draw on conventional or orthodox expressions; the probability that the formulation of a response produced an artificial clarity and order; the nearly universal fact of selective return; and the fragmentary and superficial quality the results often show as a whole (Coe, 1916; Pratt, 1920; Siegmund, 1942; Spranger, 1924; Stählin, 1912; Uren, 1928).

Today's questionnaire typically requires the respondent merely to indicate

agreement or disagreement with a series of statements, or to select one alternative out of a cluster of three or four. The radical standardization and simplification of questionnaire responses greatly facilitate the statistical analyses that are the hallmark of objective psychology. Yet thus far it has proved extraordinarily difficult—some would say impossible—for the researcher to write alternatives that will accommodate all the respondents to their own satisfaction, and then to win the approval of the historian of religion and the psychometrician. Users of these scales have often been appropriately circumspect in their claims for them, and they have made their items and procedures public so that others might judge their adequacy or usefulness for themselves. Nonetheless, the conclusions that are drawn are sometimes subtly but seriously misleading.

Misleading Generalizations

The category of the "nonreligious" will illustrate the point. To be classified as nonreligious, people must either refuse to identify themselves with some traditional religious group or score relatively low on some scale, probably a measure of orthodoxy or ORTHOPRAXY. The "nones" and the "nonreligious" have surprised researchers by consistently proving themselves least prejudiced, least likely to yield to authority, and most likely to report mystical experiences. Gorsuch and Aleshire (1974) recognize that subjects of this type may actually view themselves as "religious." Some may even believe in God and attend church regularly (Vernon, 1968).

Whereas common sense may bring to mind the delinquent or criminal as the most obvious example of the nonreligious type, research has found that the lawbreaker frequently reports a religious affiliation and positive religious attitudes (e.g., Payne, 1972; Yochelson and Samenow, 1976). On the other hand, "Those who break away from religion," declares Cohen (1946, p. 347), "are often among the most high-minded members of the community." "It is a fact," agrees Ignace Lepp (1963, p. 134), "that a number of superior people refuse the Christian religion—and indeed all religion—because they do not find it noble enough, because its ideal does not satisfy them, because its exigencies do not strike them as 'being up to the mark.' " These "atheists in the name of value," however, do not form a homogeneous group that some new scale can be written to identify.

If "nonreligious" is seriously misleading, then the category "indiscriminately antireligious" (Allport and Ross, 1967) surely needs reconsideration too, especially when the subjects included therein—perhaps selected from religious environments, as Andrew Thompson's (1974) were—have only relatively lower scores than the others. The whole enterprise of measuring religion in terms of "objective" referents—beliefs and practices, foremost—as well as conventionally pious sentiments, needs careful rethinking. At the very least, the names applied to these measures—and thus also the generalizations that are based on them—should be replaced with more modest labels commensurate with the character of the questionnaire items. Gorsuch and Aleshire's (1974) use of "Christian faith" instead of "religion," although the term is only relatively more precise, is a clear step forward.

Also potentially misleading are the generalizations regarding trends that have been found through the use of these measurement devices. The many "statistically significant" correlations and group differences that have been reported in this literature are typically rather small. Put in more exact, technical terms, the amount of "explained variance" is rarely more than half the total variance, and frequently much less, and the distribution of scores of various subject groups always overlap considerably. Leaving behind the individuals who make up these groups obscures

the fact that many of their scores contradict the "statistically significant" differences finally reported. Moreover, the important but exceedingly difficult question of "How much of a difference makes a difference?" cannot really be answered apart from some specified practical concern. In other words, *statistical* significance does not automatically guarantee *human* significance (Bakan, 1966c). If applied indiscriminately, Ana-Maria Rizzuto (1979, p. 5) warns, the conclusions of the experimental and correlational psychologists may be harmful in the clinical context.

THE EXCLUSION OF SUBJECTIVITY

The problems we have discussed have long been recognized by thoughtful proponents of objective methods. Less often appreciated are the costs of rejecting other methods, particularly those designed to be sensitive to human subjectivity and the holistic patterns of individuality. "Anyone who thinks that he must concentrate and rely on only one type of approach, such as laboratory experiment, for his knowledge of man," writes H. P. Rickman (1979, p. 89) in his study of Dilthey, "is like an observer who peeps through the key-hole when the door is wide open."

Science Through a Keyhole

For objective psychologists of religion, the keyhole is typically a questionnaire, and what they glimpse through it is the personal world of human experience. The doorway is constituted by the investigators' own subjectivity as well as the various expressions—from spontaneous gestures to written documents—by which a person can come to know the subjectivity of others. To state the paradox more directly: though formally denying themselves access to the inner, subjective world, the objective psychologists ask wholly untrained and methodologically naive subjects to plunge into it themselves, but usually under the constraint of an extremely narrow range of possible responses. Then, on so uncertain a foundation, these researchers attempt to build a science that is intended to comprehend that underlying subjectivity. In later chapters we will see what happens when investigators find some means by which to pass through the doorway themselves.

Two Contrasting Lives: A brief example will suggest how a subjective method—in this instance, the case study, based on personal documents and biographical sources—can provide a fuller view of individual piety and its relation to personality trends. Drawing on research from her doctoral dissertation on religion in the lives of well-to-do Colonial Americans, Susan Kwilecki (1986) offers two contrasting examples: Robert Carter (1728–1804), who represents the intrinsic–committed orientation, and John Hancock (1737–1793), who typifies the extrinsic–consensual orientation.

Carter, a recluse interested in law and music but distinctly not in the financial and political responsibilities that fell to persons of his station, was for a time a rationalistic deist. Later, during a period of illness, he became a convert to evangelical Christianity and soon played a major role in promoting the Baptist Church in Virginia. A few years later, after discovering the writings of Emanuel Swedenborg and corresponding with the Baltimore Swedenborgian Society, he moved to Baltimore and became the group's leader. When it later disbanded, he associated himself with a variety of other religious assemblies until his death a few years later. Although the religious and ethical writings he left behind lack sophistication, they reveal a conceptually differentiated outlook well integrated with other areas of life, suggesting the intrinsic–committed orientation.

In contrast, Hancock, a wealthy and charming extravert who had an exceptional need to be in the limelight, rarely referred to religion in his writings. The few, brief allusions are either politically expedient formulas or incidental expressions of early-learned traditional doctrine lacking any real conviction. Even the death of his only son, though the occasion for profound grief, did not inspire any genuine religious feeling. The church to which he was a generous patron was but one more arena in which to be noticed and to exercise power. Hancock, says Kwilecki, is a textbook model of the extrinsic–consensual orientation.

Beyond giving reality and texture to objective psychology's abstractions, these cases illustrate how religious orientation is inextricably bound up with personal disposition and need. Although correlational psychologists have reported a large number of correlations between religion and personality variables, only idiographic case studies bring these variables to life and show their complex interrelations. Yet the relation is reciprocal, for only correlational research, notes Kwilecki, can reveal the generality of the trends that appear so strikingly in these two examples.

Closing the Door to Subjectivity

The exclusion of subjectivity is more radical still in the experiments on meditation that employ physiological measures as the dependent variables and then infer differences in mental states from them. According to Indologist Fritz Staal (1975), "Such methods tell us how meditation affects the body, not how it affects the mind. . . . A balanced study of mental states can only result from fully taking into account not only behavioral and physiological, but also experiential or subjective data" (p. 103). Although Staal finds some physiological investigations—notably those of Thérèse Brosse (1963)—both carefully done and suggestive, he concludes that they tell us little.

> Despite the impressive amount of careful research that has gone into the study of Yoga, and of mysticism generally, through EEG, ECG, and similar methods, one cannot help feeling that it is like studying art through films of the eye movements of art viewers. Even if the results are valid, their significance is minute (p. 104).

In contrast to Brosse's measured conclusions, those of Robert Wallace are judged to be unreflectively dogmatic and largely promotional. "The kind of advertisement for transcendental meditation that Wallace indulges in, and that acts instead on many as a deterrent, throws little light on transcendental meditation, or on meditation in general" (pp. 106–107). Thomas Mulholland (1972) similarly regrets the pseudoscientific "alpha cult" that is another by-product of EEG research. After three decades of research, he says, scientists have found no clearly defined psychological process associated with alpha. Even its physiological origin is a matter of uncertainty.

As meditators themselves, many of these researchers know from first-hand experience that what they gain is not well represented in these physiological measures. One such researcher, Michael West (1987, p. 193), finds his own experience "vastly more fruitful" and the psychological research "vastly more frustrating" in comparison to each other, a discrepancy that he compares to a Zen koan. Why does he keep on meditating when empirical evidence suggests that the practice reduces arousal no more than ordinary rest? "The simple answer would be to say that I trust my experience more than my science. But the more complex anwer is that my reading of the research and theoretical literature in this area has not given me reason to discontinue my practice and has offered some reason for sustaining it (if any more reason beyond my subjective experience were needed)." West might

find a solution to his koan if he employed a method that allowed him to take his experience more fully and systematically into account, much as Carl Albrecht (1951) does in his own, phenomenological study of meditative–mystical experience (see Chapter 12). By comparison, the "phenomenological" studies reviewed in West's book hardly deserve the name, for the observers are apparently untrained in phenomenological investigation, and the whole enterprise continues to be framed with assumptions that violate the phenomenological attitude.

In an age that is suspicious or disdainful of the "merely subjective," it is easy to understand why most meditation researchers emphasize objective methods. Advocates of TM realize that many people will take transformed states of consciousness seriously only if they can be shown to be related to palpable physical reality. Yet selling these states on the basis of their calming effect rather misses the point. In its traditional contexts, meditation has always been directed toward some higher state of awareness or illumination, not mere relaxation or reduction of stress. In separating the practice of meditation from this broader purpose and then reducing its significance to objectively observable bodily changes, its modern-day proponents risk trivializing what has been one of the most powerful instruments of the spiritual quest.

THE NEGLECT OF RELIGIOUS CONTENT

The methods employed by the objective psychologists lend themselves far more readily to the study of religious *persons* than to religious *content.* They are thus also distinctly ahistorical. The corresponding neglect of the inexhaustibly rich world of myth, symbol, and ritual is one major reason why many scholars in religious studies are less than enthusiastic about the objective approach. In sharp contrast, the depth psychologies provide a variety of interpretive principles for uncovering the underlying meaning and coherence of these universal yet highly variable expressions of the religious spirit. There are profound differences among these principles, and they too require critical assessment. Yet they at least have the virtue of addressing religion in its totality.

The investigation of religious content is not utterly foreclosed to objective psychologists, as a study of the ancient Chinese Yin–Yang symbol demonstrates (Craddick, Thumin, and Barclay, 1971). These researchers asked 242 university students in psychology and English classes to rate a projected green-and-blue version of this circular symbol using the SEMANTIC DIFFERENTIAL, a series of 30 seven-point scales defined in terms of bipolar adjective pairs, such as good–bad, sacred–profane, hard–soft, and agitated–quiet. The results show that the Yin–Yang diagram had essentially positive connotations for these students, who tended to perceive it as good, beautiful, kind, pleasant, happy, and clean. Its intended representation of integration and harmony was reflected in relatively high ratings on relaxed and quiet, whereas its potency was also suggested by associations with strong, powerful, deep, and agile. Its ancient association with the conjunction of male and female (among other opposites) was appropriately reflected in a mean rating precisely at the midpoint for the masculine–feminine scale. These researchers conclude that their data confirm the traditional interpretations of this symbol, if not also aspects of Jung's psychology.

Once again, we have a study of a religious element removed from its traditional context. Using subjects who most probably know little if anything about the Taoist tradition, with which this symbol is usually associated, these investigators made no

effort to discover its connotative meaning for the Eastern mind. Rather, they sought to explore the "inherent"—thus presumably universal—meanings of the symbol. Although the semantic differential's bipolar structure would seem peculiarly appropriate for studying an abstract symbol that is thought to represent the conjunction of life's basic oppositions, it is impossible to know whether a middle rating reflects the dynamic balance traditionally implied by the symbol or merely a judgment that the symbol's qualities are in that respect indeterminate. Furthermore, even in the few cases in which the *mean* rating does fall close to the midpoint, there is considerable *individual* variation around that mean. If the Yin–Yang symbol does have inherent qualities in line with the Taoist interpretation, it would seem that many Western students do not clearly perceive them. Whereas the harvest from this study is thus rather modest, the undertaking at least suggests the possibility of exploring religious content by objective procedures.

We should be cautious, of course, in concluding that the problems and limitations of any particular study are intrinsic to its method or approach. Objective studies frequently suffer more from inadequate conceptualizing than from the inherent constraints of quantitative techniques (Kirkpatrick, 1995b). As we will see in subsequent chapters, objective procedures can serve a highly useful function in testing out hypotheses issuing out of other psychological perspectives. They are not sufficient in themselves for a comprehensive psychology of religion, but it would seem foolish to reject them out of hand on the basis of what has been accomplished with them thus far. Their potential, we might suspect, is far from fully realized.