THE NEUROSCIENTIFIC STUDY OF RELIGIOUS AND SPIRITUAL PHENOMENA: OR WHY GOD DOESN’T USE BIOSTATISTICS

by Andrew B. Newberg and Bruce Y. Lee

Abstract. With the rapidly expanding field of neuroscience research exploring religious and spiritual phenomena, there have been many perspectives as to the validity, importance, relevance, and need for such research. In this essay we review the studies that have contributed to our current understanding of the neuropsychology of religious phenomena. We focus on methodological issues to determine which areas have been weaknesses and strengths in the current studies. This area of research also poses important theological and epistemological questions that require careful consideration if both the religious and scientific elements are to be appropriately respected. The best way to evaluate this field is to determine the methodological issues that currently affect the field and explore how best to address such issues so that future investigations can be as robust as possible and can become more mainstream in both the religious and the scientific arenas.

Keywords: health; methodology; religion; spirituality.

With the rapidly expanding field of research exploring religious and spiritual phenomena, there have been many perspectives as to the validity, importance, relevance, and need for such research. There is also the ultimate issue of how such research should be interpreted with regard to epistemological questions. The best way to evaluate this field is to determine the methodological issues that currently affect the field and explore how best

Andrew B. Newberg, M.D., is assistant professor of radiology and psychiatry at the University of Pennsylvania Health System, 110 Donner Building, H.U.P., 3400 Spruce Street, Philadelphia, PA 19104; e-mail andrew.newberg@uphs.upenn.edu. Bruce Y. Lee, M.D., is a physician-scientist Fellow, Division of General Internal Medicine, University of Pennsylvania.

[Zygon, vol. 40, no. 2 (June 2005).] © 2005 by the Joint Publication Board of Zygon. ISSN 0591-2385
to address such issues so that future investigations can be as robust as possible and make this body of research more mainstream.

Are scientific investigations of religiousness of divine origin? In the biblical book of Daniel we read:

“Please test your servants for ten days, and let us be given some vegetables to eat and water to drink. Then let our appearance be observed in your presence and the appearance of the youths who are eating the king’s choice food; and deal with your servants according to what you see.” So he listened to them in this matter and tested them for ten days. At the end of ten days their appearance seemed better and they were fatter than all the youths who had been eating the king’s choice food. (Daniel 1:12–15 NKJV)

Thus, even in early religious texts there was a notion that there could be some way of evaluating the effects of religiousness on the human person. This example may well be one of the first descriptions of a controlled trial. However, biomedical research obviously has advanced significantly since biblical times, even though the study of religious phenomena is difficult.

In our companion article (Lee and Newberg 2005; see pp. 443–68 in this issue) we address many of the methodological issues related to the study of spirituality and health. In the present article we focus more on the physiological and neurobiological studies that have been performed and the potential issues associated with such studies. In some sense, this research builds upon the clinical work, because it is helpful to understand the ultimate expression of these phenomena as they affect a person’s life and health. However, physiological studies are crucial for understanding how the clinical results may come about. Furthermore, physiological studies examine the specific nature of spirituality and its effect on the body.

In this essay we review four dimensions of this area of research with a critical perspective on methodology and statistical analysis. The four dimensions as they relate to the neuroscientific study of religious and spiritual phenomena are (1) appropriate measures and definitions; (2) subject selection and comparison groups; (3) study design and biostatistics; and (4) theological and epistemological implications.

**MEASUREMENT AND DEFINITION OF SPIRITUALITY AND RELIGIOUSNESS**

One of the most important issues related to the measurement of religious and spiritual phenomena has to do with correlating subjective and objective measures. For example, if a particular type of meditation reduces blood pressure or is associated with changes in cerebral metabolism, it is critical to know what was actually experienced by the individual.

**Subjective Measures of Spirituality.** In some sense, the most important measures of religious and spiritual phenomena are those that pertain to the subjective nature of the experience. When a person has a religious
or spiritual experience, he or she usually can try to describe it in terms of various cognitive, behavioral, and emotional parameters. A person will usually define the experience as “spiritual,” which distinguishes the experience from others that are regarded as “nonspiritual.” The issue of measuring the subjective nature of these phenomena is akin to opening the mysterious “black box” in which something is happening, but it is not immediately observable by an outside investigator. The problem becomes more difficult when trying to compare experiences between individuals and across cultures. A spiritual experience for a Jew may be vastly different than a spiritual experience for a Hindu. Furthermore, there is likely to be a continuum of experiences ranging from barely perceptible to absolutely mystical (d’Aquili and Newberg 1993). The question for any researcher is how to get some handle on the subjective component of such experiences. Is there a way to quantify and compare the subjective feelings and thoughts individuals have regarding their spiritual experiences? It is difficult to develop adequate scales to measure spirituality and religiousness and often even more difficult to find them. Such scales are difficult to find in the literature especially when they are reported in nonscientific journals that are not typically cited or referenced in literature reviews (Larson, Swyers, and McCullough 1998).

A number of attempts have been reported in the literature to develop a self-reporting scale that measures the subjective nature of a particular religious or spiritual phenomenon. The book *Measures of Religiosity* (Hill and Hood 1999) provides fertile ground for various scales and questionnaires that assess everything from a person's feeling of commitment to awe to hope to the direct apprehension of God. Some have been assessed for validity and reliability, which is critical if these scales are to have any use in future research studies. Testing the *validity* implies that the results return information about what the scale is supposed to measure (Patten 2000). For example, a valid scale of a feeling of hopefulness would ask questions regarding the amount of hope a person has. If this scale did not address hope but rather happy emotional responses, it would not be a valid measure of hope. *Reliability* assesses whether the scale when given to the same person at different times yields roughly the same results (Patten 2000).

Although it is important to assess the reliability and validity of scales, this is particularly problematic with regard to religious and spiritual phenomena. The reason for this difficulty is the problem with defining these terms. If someone defines *spiritual* as a feeling of awe, and another defines it as a feeling of oneness, what types of questions should be used to assess spirituality? A questionnaire that asks about feelings of awe might not really be measuring spirituality. Therefore, until clear and operational definitions of spirituality and religiousness can be determined, there will always be the potential problem of developing valid scales. Reliability is also a problem, because spirituality and religiousness can be very consistent or
widely variable within an individual. He or she might subjectively feel different at different times, and therefore the reliability of any scale with the intention to measure spirituality is always problematic.

Another problem with individual scales is the issue of usefulness across traditions and cultures. For example, many of the scales that are referenced in *Measures of Religiosity* are Christian-based and may not be useful for evaluating Jewish or Buddhist perspectives. Other scales have a more universal quality or at least can be modified to accommodate other perspectives. However, this might bring into question the validity and reliability of such scales in different contexts.

There is another interesting problem with scales that attempt to measure the subjective nature of spiritual or religious phenomena. This arises from the fact that most scales of spirituality and religiousness require the individual to respond in terms of psychological, affective, or cognitive processes. Thus, questions are phrased: How did it make you feel? What sensory experiences did you have? What did you think about your experience? Such measures are very valuable to individuals interested in exploring the neural correlates of such experiences, because psychological, affective, and cognitive elements usually can be related to specific brain structures or functions. The problem with phrasing questions in this way is that one never escapes the neurocognitive perspective to get at something that might be "truly" spiritual. It might be suggested that the only way in which an investigator can reach something that is truly spiritual would be through a process of elimination in which all other factors—cognitive, emotional, and sensory—are eliminated, leaving only the spiritual components of the experience. The most interesting result from a brain scan of someone in prayer would be to find no significant change in the brain during the time that the individual has the most profound spiritual experience.

As already mentioned, a problem with developing adequate measures is ensuring that they measure what they claim to measure. A subjective scale designed to measure the degree of an individual's religiosity needs to focus on the things that make someone religious. However, this first requires a clear definition of *religiousness* and *spirituality*. Furthermore, these definitions must be operationalized so that any measure or study can have a firm enough grasp to actually measure something (Koenig 1998; Koenig, McCullough, and Larson 2001). To that end, it is important to avoid narrow definitions that might impede research and also to avoid broad definitions that cannot be measured. For example, definitions of religion that pertain to a single God would eliminate almost two billion Hindu and Buddhist individuals from analysis. A definition of religiousness that is too broad, however, might include many bizarre experiences and practices such as cults or devil worship.

One approach to defining religiousness and spirituality offered by a consensus conference of scientists interested in studying spirituality and health
created criteria for each definition (Larson, Swyers, and McCullough 1998). The criteria for spirituality included the subjective feelings, thoughts, experiences, and behaviors that arise from a search or quest for the sacred. The term search referred to attempts to identify, articulate, maintain, or transform, and the term sacred referred to what the individual perceived as a divine being, ultimate reality, or ultimate truth. The criteria for religion or religiousness included the criteria for spirituality and/or a search for nonsacred goals (such as identity, belonging, meaning, health, or wellness) in the context of spiritual criteria. Furthermore, religiousness was associated with “means and methods of the search that received general validation and support from within an identifiable group of people” (1998, 21). Such definitions or criteria meet the need for an operationalized approach but also have the potential to exclude certain elements of religiousness and spirituality. Such a definition also may be problematic from theological or philosophical perspectives. Therefore, any definition of religion and spirituality will benefit from being considered dynamic and able to adapt and change according to future findings and analyses.

Another issue related to problems with definitions is that there are so many approaches to religious and spiritual phenomena that often it is difficult to generalize from one study to another. Some scholars have pointed out that one type of meditation practice or one type of experience might be substantially different from other types (Andresen 2000; Andresen and Forman 2000). It is certainly critical to ensure that any study clearly states the specific practices, subpractices, and traditions involved. Furthermore, changes in the brain associated with one type of meditative practice may not be specifically related to a different type of practice. Of course, the dynamic nature of this body of research may provide new ways of categorizing certain practices or experiences so that one can address the question regarding whether different types of meditation are truly different or are only experienced to be so.

**Objective Measures of Spirituality.** Objective measures of religious and spiritual phenomena that pertain to the neurosciences include a variety of physiological and neurophysiological measures. Recent advances in fields such as psychoneuroendocrinology and psychoneuroimmunology address the important interrelationship between the brain and body. Any thoughts or feelings perceived in the brain ultimately have effects on functions throughout the body. While this can complicate measures as well as introduce confounding factors, this integrated approach allows for a more thorough analysis of religious and spiritual phenomena (Newberg and Iversen 2003). Several types of measures already reported in the literature include measures of autonomic nervous system activity. These are the most common approaches to specific religious and spiritual practices such as meditation or prayer. A number of studies have revealed changes in blood pressure and heart rate associated with such practices (Sudsuang, Chentanez,
and Veluvan 1991; Jevning, Wallace, and Beidebach 1992; Koenig, McCullough, and Larson 2001). It is interesting that the actual changes may be quite complex, involving both a relaxation and an arousal response. Early work by E. Gellhorn and W. F. Kiely (1972) developed a model of the physiological processes involved in meditation based almost exclusively on autonomic nervous system (ANS) activity, which, while somewhat limited, indicated the importance of the ANS during such experiences. These authors suggested that intense stimulation of either the sympathetic or parasympathetic system, if continued, could ultimately result in simultaneous discharge of both systems (what might be considered a “breakthrough” of the other system). Several studies have demonstrated predominant parasympathetic activity during meditation associated with decreased heart rate and blood pressure, decreased respiratory rate, and decreased oxygen metabolism (Sudsuang, Chentanez, and Veluvan 1991; Jevning, Wallace, and Beidebach 1992; Travis 2001). However, one study of two separate meditative techniques suggested a mutual activation of parasympathetic and sympathetic systems by demonstrating an increase in the variability of heart rate during meditation (Peng, Mietus, Liu, et al. 1999). The increased variation in heart rate was hypothesized to reflect activation of both arms of the autonomic nervous system. This notion fits the characteristic description of meditative states in which there is a sense of overwhelming calmness as well as significant alertness. Also, the notion of mutual activation of both arms of the ANS is consistent with recent developments in the study of autonomic interactions (Hugdahl 1996).

Measures of hormone and immune function have been explored as an adjunct measure to various clinical outcomes (O’Halloran et al. 1985; Walton et al. 1995; Tooley et al. 2000; Infante et al. 2001). Thus, if a hypothetical study showed that the practice of meditation results in reductions in cancer rates, it might be valuable to measure the immunological and hormonal status of the individuals to determine the physiological basis of the effect. Certain cancers are related to abnormalities in immune (such as leukemia or lymphoma) or hormonal function (breast and prostate cancer). Also, alterations in various hormones and immune functions may be related to specific changes in brain function, and this interaction can be bidirectional. Thus, certain brain states may enhance hormonal status, but these hormonal states may in turn affect brain function. This can be observed in women with premenstrual syndrome and in other circumstances in which various neurohormones alter emotional, cognitive, and behavioral states.

Neurophysiological changes associated with religious and spiritual states can be obtained through a number of techniques, each with its own advantages and disadvantages. In general, the primary requirement is that the methodology evaluate functional changes in the brain. However, there are many ways of measuring such functional changes. Early studies of
meditation practices made substantial use of electroencephalography (EEG), which measures electrical activity in the brain (Banquet 1973; Hirai 1974; Hebert and Lehmann 1977; Corby et al. 1978). EEG is valuable because it is relatively noninvasive and has very good temporal resolution: the instant an individual achieves a certain state, the EEG should change accordingly. For this reason it has continued to be useful in the evaluation of specific meditation states (Lehmann et al. 2001; Aftanas and Golochekikine 2002; Travis and Arenander 2004). The major problem with EEG is that spatial resolution is very low, so any change can be localized only over very broad areas of the brain. Another problem is that analysis can be difficult because of the extensive number of recordings that are made during any session. However, EEG may be particularly valuable to include in studies employing functional imaging techniques, because the EEG may help to signal certain states or at the very least determine whether the individual being studied has fallen asleep.

Functional neuroimaging studies of religious and spiritual phenomena have utilized positron emission tomography (PET), single photon emission computed tomography (SPECT), and functional magnetic resonance imaging (fMRI). In general, such techniques can measure functional changes in the brain in pathological conditions, in response to pharmacological interventions, and during various activation states. Activation states have included sensory stimulation (visual, auditory, and so forth), motor function and coordination, language, and higher cognitive functions (such as concentration) (Newberg and Alavi 1996). The changes that can be measured include general physiological processes such as cerebral blood flow and metabolism in addition to many aspects of the neurotransmitter systems. The serotonin, dopamine, opiate, benzodiazepine, glutamate, and acetylcholine systems all have been evaluated in a number of brain states (Newberg and Alavi 2003; Warwick 2004; Kennedy and Zubieta 2004).

Although functional neuroimaging studies have contributed greatly to the understanding of the human brain, each has its advantages and limitations with respect to evaluating religious and spiritual phenomena. Functional MRI primarily measures changes in cerebral blood flow. In general, this is a valid method for measuring cerebral activity, because a brain region that is activated during a specific task will experience a concomitant increase in blood flow. This coupling of blood flow and activity provides a method for observing which parts of the brain have increased activity (increased blood flow) and decreased activity (decreased blood flow). Functional MRI has several advantages. It has very good spatial resolution and can be coregistered with an anatomical MRI scan that can be obtained in the same imaging session. This allows for a very accurate determination of the specific areas of the brain that are activated. It also has very good temporal resolution, so that many images can be obtained over very short periods of time, as short as one second. This allows for a brain response
that may occur very quickly to be observed. Thus, if a subject is asked to pray ten different prayers sequentially while in the MRI, the differences in blood flow can be detected in each of those ten prayer states. Also, fMRI does not involve any radioactive exposure. The disadvantages are that images must be obtained while the subject is in the scanner, and the scanner can make up to one hundred decibels of noise, which can be very distracting when individuals are performing spiritual practices such as meditation or prayer. However, several investigators have successfully utilized fMRI and have performed the study by having subjects practice their meditation technique at home while listening to a tape of the fMRI noise so that they become acclimated to the environment (Lazar et al. 2000). Also, the MRI noise can affect brain activity, particularly in the auditory cortex. In addition, fMRI relies on a tight coupling between cerebral blood flow and actual brain activity, which, while a reasonable assumption, is not true in all cases. Well-known examples in which brain activity and blood flow are not coupled include stroke, head injury, and pharmacological interventions (Newberg and Alavi 2003). A detailed evaluation of this coupling in all brain states has not been made. Another disadvantage is that at present fMRI cannot be used to evaluate individual neurotransmitter systems.

PET and SPECT imaging also have advantages and disadvantages. The advantages include relatively good spatial resolution for PET (comparable to fMRI) and slightly worse for SPECT imaging. PET and SPECT images can also be coregistered with anatomical MRI, but this must be obtained during a separate session, and therefore matching the scans is more difficult. PET and SPECT both require the injection of a radioactive tracer, so radioactivity is involved, although usually this is fairly low. Depending on the radioactive tracer used, a variety of functional parameters can be measured including blood flow, metabolism (which more accurately depicts cerebral activity), and many different neurotransmitter components. The ability to measure these neurotransmitter systems is unique to PET and SPECT imaging. Such tracers can measure either state or trait responses. It should also be mentioned that some of the more common radioactive materials such as fluorodeoxyglucose (which measures glucose metabolism) can be injected through an existing intravenous catheter when the subject is not in the scanner. This allows for a more conducive environment for performing practices such as meditation and prayer. This tracer becomes “locked” in the brain during the injection period, and the person can then be scanned after he or she has completed the practice, but the scan will still measure changes associated with the practice (Herzog et al. 1990–1991; Newberg et al. 2001). A major drawback to PET and SPECT imaging, in addition to the radioactive exposure, is that these techniques have generally poor temporal resolution. Depending on the tracer, this resolution can be as good as several minutes and as bad as several hours or even days. PET or SPECT would be very difficult to use to study ten different
prayer states in the same session. However, two or three states might be measured in the same imaging session if the appropriate radiopharmaceutical were used (Lou et al. 1999). So, depending on the goals of the study, the various neuroimaging techniques are better or worse.

There are other more global problems that affect the ability to interpret the results of all functional imaging studies. The most important of these is how to be certain what is actually being measured physiologically and how it compares to various subjective experiences. There are already potential problems addressing what a particular scan finding means in terms of the actual activity state of the brain. For example, it is not clear what will be observed if there is increased activity in a group of inhibitory neurons. Would that result in increased or decreased cerebral activity as measured by PET or fMRI? A bigger problem is trying to compare the observed physiological changes to the subjective state. With regard to religious and spiritual experience, it is not possible to intervene at some “peak” experience to ask the person what he or she is feeling. Therefore, if a person undergoes fMRI during a meditation session and has a peak experience, how will the researcher know which scan findings it relates to? In addition, there are typically a number of changes in the brain with varying degrees of strength. It is not clear what degree of change should be considered a relevant change—10 percent? 20 percent? From a statistical perspective, analyzing images has a number of problems including how to compare images across subjects and conditions and how to take into account the problems of multiple comparisons in terms of both activation states and individual brain regions. A program called statistical parametric mapping (SPM) can be used to evaluate various images; it works by normalizing the images, coregistering the images, and then analyzing them pixel by pixel for significant changes (Friston et al. 1995). This is a very conservative statistical approach because of the problem with multiple comparisons, and therefore subtle changes may be missed. Of course, there is still the question of whether changes observed that are not significant in SPM are still clinically relevant. Furthermore, in the study of religious and spiritual states it may be important to evaluate subjects on an individual basis because such states may be highly variable phenomenologically across subjects.

In spite of these limitations, neuroimaging studies have been used successfully to evaluate specific spiritual and meditative practices. Six studies that we know of have spanned the different neuroimaging techniques (Herzog et al. 1990–1991; Lou et al. 1999; Lazar et al. 2000; Newberg et al. 2001; Kjaer et al. 2002; Newberg et al. 2003). There is some coherence in their findings, with the frontal lobes, parietal lobes, thalamus, and limbic system appearing to be connected in a network associated with such practices. It may be that the various types of practices activate a network of brain structures in relatively similar ways. It is interesting also that there
do seem to be some differences that correlate well with the variations among
the approaches. One study also measured changes in the dopamine system
and found increased activity during meditation-related practices (Kjaer et
al. 2002). The level of complexity of our understanding continues to im-
prove as more studies are performed. Future studies are necessary to more
thoroughly evaluate the neurophysiological changes that occur in the brain
during various religious and spiritual phenomena.

SUBJECT SELECTION AND COMPARISON GROUPS

Another interesting methodological issue in the study of religious and spir-
tual phenomena is to determine who are the most appropriate subjects to
study and who should represent the comparison group(s). The issue of
who to study with regard to religious and spiritual phenomena depends
somewhat on the definition of the phenomena. Obviously, if a researcher
wanted to evaluate physiological changes during meditation, there would
be thousands of different possible groups to consider studying. But which
elements of a particular practice or experience are of most interest? The
more specific a researcher wants to be in terms of the phenomena, the
more focused will be the subject group. If a researcher wanted to study the
physiological effects of the rosary, the group would have to consist of those
individuals who practice the rosary. If the focus were on feelings of unity,
many different practices could be chosen, or perhaps the study group would
consist of individuals from many different backgrounds.

An important issue in this regard is the level of expertise or proficiency
of the individuals being studied. In the case of meditative practices, there
could be very different results between novice, experienced, and master-
level individuals. The differences could be related to different physiologi-
cal processes that occur at the different levels, or the differences could be
related to whether more novice individuals can perform the practice in a
manner that is similar to their usual level of practice while under the scruti-
ny of the researcher. The noise of an MRI scanner may result in a novice's
not being able to meditate adequately, while a more experienced individual
may have less of a problem with the distraction. Thus, the difference ob-
served might relate to the fact that one of them successfully performed the
practice rather than to actual differences between the practices. It also is
important to select individuals who are of similar socioeconomic and health
backgrounds. If Franciscan nuns are less likely to smoke, their brain scans
might differ from a group of other individuals who do smoke.

The other major issue in terms of subject selection relates to the com-
parison or control groups. One possibility frequently employed in func-
tional neuroimaging studies is that the individual acts as his or her own
comparison. Studies of various meditative practices typically compare the
meditation state to the subject's own baseline waking state. Others have
suggested that a more appropriate comparison would be a state in which
individuals are doing a task that is similar but has no specific spiritual meaning. One study explored whether different mantras (some spiritual, some not) have different effects on the brain electrical activity during meditation (Telles, Nagarathna, and Nagendra 1998). The result of this study was that there was such a distinction. Another issue with regard to using subjects as their own comparison involves excluding other factors that are part of the practice. Thus, a practice that involves burning incense would be better compared to a baseline state in which incense is also used; otherwise, the primary change observed may be in the olfactory regions of the brain and may have nothing to do with the spiritual practice. Similarly, if a practice requires the eyes to be open—reading prayers, for example—the baseline state should have the subject with eyes open or possibly even reading nonreligious texts. Some studies have looked at such differences and found distinctions in cerebral activity depending on whether a subject was reading a religious or nonreligious text (Azari et al. 2001). Other types of practices might also be used as comparisons including artistic and creative practices, athletic events, or cognitive and visuo-spatial tasks. Comparison groups could be other practitioners in the same tradition but with different levels of expertise or practitioners in other traditions in which similar practices are performed. These groups might help to determine longitudinal effects of various spiritual practices, but factors such as age and health might interfere with the interpretation of such studies.

Placebo groups are another important problem with the study of religious and spiritual phenomena. It is not clear what a placebo would represent in many cases, because most people who are spiritual know whether or not they are actually performing their spiritual practice. Placebo groups in this case more likely will represent other tasks that resemble the spiritual one but lack the specific spiritual content. Thus, if reading a prayer is going to be studied, reading a nonreligious text would represent a reasonable comparison.

One other question with regard to the subject selection of studies is whether or not any of these studies actually measures God. This is more of an interpretive component at present, but many scholars have tried to use such studies to either prove or disprove religion or God. At the moment, science does not appear to be capable of aiding in the making of such claims. If an intercessory-prayer study works, it is not clear what the actual mechanism is. God actually helping the sick or human consciousness affecting people at a distance are two possibilities. If an intercessory-prayer study does not work, does it mean that God has been disproved?

STUDY DESIGN AND BIOSTATISTICAL ANALYSIS

Based on the above review of the existing literature and the proposed operational definition of spiritual experience, there are at least seven neuroscientific paradigms that can readily contribute to the initial
operationalization of spiritual experience (Larson, Swyers, and McCullough 1998). These seven paradigms include (1) the neurophysiology of spiritual interventions, (2) spiritual interventions associated with psychopharmacological agents, (3) drug-induced spiritual experiences, (4) neuropathologic and psychopathologic spiritual experiences, (5) spiritual experiential development in infants, children, and adolescents, (6) physical and psychological therapeutic interventions, and (7) neurophysiology of the sense of certainty. We consider these study designs and then review the biostatistical issues with such studies.

_The Neurophysiology of Spiritual Interventions._ The first paradigm involves an experimental spiritual intervention such as prayer or meditation with concomitant measures of a psychological- and spiritual-dependent evaluation. This will help to define and delineate the nature of the spiritual intervention. These psychological and spiritual measures can then be compared to simultaneously derived neurobiological parameters, such as electroencephalographic activity, cerebral blood flow, cerebral metabolism, and neurotransmitter activity. Such measures can be performed with state-of-the-art imaging techniques including PET, SPECT, and MRI. Body physiological scalar parameters such as blood pressure, body temperature, heart rate, and galvanic skin responses (which measure autonomic nervous systems activity) can also be measured. Other body physiological parameters such as immunological assessments, hormonal concentrations, and autonomic activity should be included to complete a thorough analysis of the effects of spiritual interventions.

_Altering Spiritual Interventions._ The second paradigm that might be employed to investigate spiritual experience from a neuroscientific approach uses pharmacological agents or other interventions in an attempt to alter spiritual interventions. For example, studies might attempt to show the effects of an opiate antagonist on the strength of the subjective experience of meditation or prayer. Preliminary studies (on one or a few subjects) of this type have shown no effect on EEG patterns during meditation when subjects were given either an opiate or benzodiazepine antagonist (Sim and Tsoi 1992). The effects of transcranial magnetic stimulation, other pharmacological agents, or even surgical procedures (performed for other purposes) could be evaluated. However, it is clear that more extensive studies measuring a number of neurophysiological parameters are required. Other agonists and antagonists may be used to determine their ability to augment or diminish spiritual experiences. In addition, the exploration of various pharmacological agents on spiritual interventions may help to delineate the role of different neurotransmitter systems. Such studies also offer the possibility of measuring dose responses in terms of spiritual interventions.
**Drug-Induced Spiritual Experiences.** A third paradigm involves persons whose use of hallucinogenic agents has resulted in intensive spiritual experiences. Since it has long been observed that drugs such as opiates, LSD, and stimulants can sometimes induce spiritual experiences, careful studies of the types and characteristics of drug-induced spiritual experiences, perhaps using modern imaging techniques, may help elucidate which neurobiological mechanisms are involved in more “naturally derived” spiritual experiences. Some studies related to the use of such hallucinogenic agents have already been performed (Vollenweider et al. 1997; 1999; 2000), but a more extensive study of such agents, particularly in relation to religious and spiritual experiences, is required. Comparing this paradigm to naturally occurring spiritual phenomena may allow for a better distinction of pathologic and nonpathologic spiritual experiences.

There are obvious ethical and legal considerations with studies such as these. However, subjects who have had pharmacologically induced spiritual experiences can be studied using radioactive analogues of such agents as a means of determining the concentration of receptors and their agonists. Another related approach would be to study the effects of drug withdrawal on spiritual experience. However, we have found no reports in the literature of such findings.

**Neuropathologic and Psychopathologic Spiritual Experiences.** A fourth paradigm involves patients with various known neuropathologic and psychopathologic conditions. Neurological conditions including seizure disorders (particularly in the temporal lobes), brain tumors, and stroke have been associated with spiritual experiences or alterations in religious beliefs. Temporal lobe epilepsy has been associated with hyperreligiousity and religious conversion (Bear 1979; Bear and Fedio 1977). Psychiatric disorders such as schizophrenia and mania also have been associated with spiritual experiences and religious conversion. Delineating the type of pathology and the location of that pathology will aid in determining the neurobiological substrate of spiritual experience. Thus, neuropsychiatric disorders can be an effective resource for the neuroscience of spiritual experience.

Research on pathological conditions has classically been used to elucidate the normal functions of biological systems. Studying spiritual experiences in psychiatric and neurological disorders may be central to the identification of largely nascent neurobiological systems that subserve “normal” spiritual experience. This presents a crucial distinction to the historic psychiatric implication that spiritual experience is an expression of psycho- or neuropathology. It provides a framework in which normal spiritual experience can occur in pathological and normal conditions and pathologic spiritual episodes can occur in individuals with or without psychopathological disorders. However, care must be taken to avoid referring to spiritual experience only in pathological terms or as associated with
pathological conditions and also to avoid reducing spiritual experiences only to neurophysiological mechanisms.

**Spiritual Experiential Development.** There is a fairly extensive literature regarding the developmental aspects of religion and spiritual experience (Fowler 1981; Tamminen 1994; Oser 1991). These reports consider the overall development of spiritual experience from infancy through adolescence and into adulthood. They also consider the neurocognitive developments necessary for spiritual experience to arise. A more primitive form of undifferentiated faith may occur in infancy, while the more complex aspects of spiritual experience that include cognitive, cultural, and affective components usually require growth into adulthood (Fowler 1981). Most of these analyses of spiritual experiential development are grounded in psychology. However, neuroscience may be able to use these findings and compare them to the development of various brain structures and neurocognitive processes. This may help elucidate which brain structures and functions are required for various components of spiritual experience. The developmental approach can also be viewed from the end-of-life perspective. Alterations in spiritual functions may be associated with diffuse neuropathological conditions such as dementia. Furthermore, it may be useful to study alteration in spiritual functions that are associated with decrements both in neurocognitive functions and in physical health.

**Physical and Psychological Therapeutic Interventions.** There are a large number of ongoing studies exploring the therapeutic effects of meditation, stress management, prayer, and other related interventions for various psychological and physical disorders including anxiety disorders, hypertension, coronary artery disease, cancer, and the human immunodeficiency virus (Kabat-Zinn 1992; Carson 1993; Levin and Vanderpool 1989; Levin 1994; Miller, Fletcher, and Kabat-Zinn 1995; Leserman et al. 1989; Zamarra et al. 1996; Massion et al. 1995; Schneider et al. 1995). While these studies focus on the effects of the intervention on various disease parameters, it may be possible to “piggyback” on these studies to include measures of spiritual experience and well-being. Using measurement scales already available in the literature, it may be possible to determine the relationship of spiritual experiences and well-being to the intervention as well as to the progression of the disorder. Performing high-quality studies is essential to demonstrating the relationship between spirituality and health especially in light of criticism regarding the methodology of these early studies (Sloan, Bagiella, and Powell 1999; Sloan and Bagiella 2002).

**Neurophysiology of the Sense of Certainty.** This seventh paradigm considers one of the essential characteristics of spiritual experience—a sense of certainty that the experience represents reality. This is what helps differentiate dream states from spiritual experiences that may have similar compo-
nents. The spiritual experiences are invariably perceived as being real. This sense of certainty is also an integral part of the concept of “faith.” There is no literature that can be used to begin a search for the neurophysiological substrate of the sense of certainty. However, one avenue would be the study of memory and memory loss currently being pursued by several groups (Nyberg et al. 1996; Gevins et al. 1996; Owen et al. 1996; Blaxton et al. 1996). Can a distinction be made between those answers that subjects give that they are certain about and the answers about which they harbor doubt? The use of brain imaging or EEG studies may help elucidate the brain structures involved in such a determination.

**Statistical Analysis Issues.** In terms of statistical analysis, several issues arise in the study of religious and spiritual phenomena. Such phenomena are frequently very complex and have many different components. As mentioned above, these components are both subjective and objective. In order to account for the variety of components, a number of variables must be factored into the statistical analysis. Thus, simple statistical comparisons may oversimplify the findings and miss important covariates that may have significant contributions to the findings. Every effort should be made to perform statistical analyses in studies of religious and spiritual phenomena with the same rigor and complexity as in other biomedical studies. To this end, it is imperative that well-qualified statisticians evaluate data from these studies in order to ensure a high quality of research.

Another problem that may be somewhat unique to religious phenomena is the interindividual differences that may be beneficial to evaluate. For example, in our research study of Tibetan Buddhist meditators, we asked each participant to practice the same type of meditation for the same amount of time. In this way, the data were easier to pool for group analysis. However, we may have missed important interindividual differences related to the strength and depth of the meditation practice, the specific experiences individuals may have had, and whatever unique techniques they used in their practices. Statistical analysis is limited in evaluating interindividual differences, especially when the focus is on physiological measures in the brain or body. Future development of analyses that can better explore such interindividual differences will benefit this field.

**Theological and Epistemological Implications**

One of the most ancient problems of philosophy is how to tell whether the external world corresponds, at least partially, to our mental representation of it. The question of what is “really real” has been considered, with various answers, since the time of the presocratic Greek philosophers in the West. Preoccupation with this question is even older in Eastern religio-philosophical traditions. In considering the neuroscientific approach to religious and spiritual phenomena, we might ponder whether epistemological issues can actually be addressed.
A number of researchers claim that because there is a neurobiological correlate of a religious phenomenon, there is nothing more to that phenomenon. While this interpretation may ultimately be accurate, that a neurobiological correlate exists does not specifically refer to the causal mechanism of such phenomena. That is, if the brain activity changes during a mystical communion with God, it is not clear whether the brain activity caused that experience or responded to that experience. Even situations in which religious states are induced by pharmacological agents does not necessarily detract from the spiritual nature of these states. Shamanic practices in which various substances are ingested to aid in the spiritual journey are not viewed as less real or less spiritual because of the use of these substances. Use of such substances alone, however, does not typically result in profound religious experiences. It is clear that the specific context in which various practices and experiences arise is crucial to the spiritual nature of these phenomena.

In reconsidering the epistemological question from a neuroscientific perspective, sometimes referred to as neuroepistemology, how reality is experienced in the brain results in a complex paradox (Newberg, d’Aquili, and Rause 2001). The three most common criteria given for judging what is real are (1) the subjective vivid sense of reality, (2) duration through time, and (3) agreement intersubjectively as to what is real. Each of these can be related to specific brain functions. But it may be demonstrated that all three of these criteria determining what is real can be reduced to the first—the vivid sense of reality. The sense of duration through time depends on the structuring of time in baseline reality. It appears that the ability to have a sense of time, or more properly duration, is structured by the brain. Alteration of the function of parts of the brain that subserve temporal ordering, for any reason, results in a significant distortion of the perception of time in a number of ways. Most dramatically, during certain spiritual practices and states there is no sense of time or duration while the person is in that state. It becomes obvious that time and duration are not absolutes; they derive their perceived qualities from brain structuring. Hence, it begs the question to derive the reality of baseline reality from one of the qualia, in this case time, which is itself structured by baseline reality (the brain). This same critique applies to any appeal for the reality of objects that depend on characteristics of baseline reality the perception of which is known to be structured by the brain. The third criterion for the reality of entities, intersubjective validation, again arises from begging the question. The “subjects” who agree or disagree about entities being real are themselves only images or representations within the sensoricognitive field of the analyzing subject-philosopher. Thus, any person analyzing his or her own experience must start out, at least, as a naive solipsist. In fact, we are satisfied that every criterion of the reality of entities collapses into the first, the vivid sense of reality.
If we conclude that reality is ultimately reducible to the vivid sense of reality, what are we to make of religious and spiritual states that appear to the experiencing subject to be more real than baseline reality, even when they are recalled from within baseline reality? If we take baseline reality as our point of reference, it seems that there are some states that appear to be inferior to baseline reality and some states that appear to be superior when these states are recalled in baseline reality. And this is the crucial point. These different experiences of reality appear more real than baseline reality when recalled from baseline reality. Thus, individuals almost always refer to dreams as inferior to baseline reality when they are recalled and discussed within baseline reality. The same is true of psychotic hallucinations—after they are cured by phenothiazines or other psychotropic medications. A person having emerged from such a psychotic state will recall it as psychotic.

The same cannot be said of many religious and spiritual states, which appear to be more real than baseline reality and are vividly described as such by experiencers after they return to baseline reality. This is true of a number of such states including absolute unitary states (Newberg, d’Aquili, and Rause 2001), “cosmic consciousness” as described by R. M. Bucke (1961), certain trance states, hyperlucid visions (usually of religious figures, religious symbols, and dead persons), and near-death experiences (Newberg and d’Aquili 1994). So real do these experiences appear when recalled in baseline reality that they often alter the way the experiencers live their lives. Studies have been performed on this topic with near-death experiencers. Those who have had the core experience clearly behave more altruistically, more kindly, and with greater compassion toward other human beings than they showed before the experience (Ring 1980). Furthermore, there is a marked tendency for near-death experiencers not to fear death. And these beneficial changes persist not only for a short period of time but for years afterward. Enough time has not passed for us to say that they persist throughout the remainder of the experiencers’ lives, but the evidence is pointing in that direction.

If it is true that all of the proposed criteria by which reality is judged to be real can be reduced to the vivid sense of reality, we have no choice but to conclude that in some sense these states, especially absolute unitary states or pure consciousness, are in fact more real than the baseline reality of our everyday lives. The word real here is used not in a poetic or metaphorical sense but in the same sense as in the utterance that this rock, or this table, is real.

Suffice it to say that when one approaches questions of reality from a neuroscientific perspective, reality becomes a very slippery concept, often manifesting itself in profoundly counterintuitive ways to the scientist, philosopher, or mystic.
CONCLUSION

Although the neuroscientific evaluation of religious and spiritual phenomena has advanced considerably since some of the initial studies that were performed more than thirty years ago, this field of research is still in its early stages. Many unique methodological issues face this field in addition to the problematic barriers of funding and academic stature. However, pursuit of such projects may ultimately pay large dividends both for science and religion. From the religious perspective, the results of such studies may help toward a better understanding of the human experience of religion. These studies enhance human knowledge of how spiritual and religious pursuits affect the mind, brain, body, and behavior. From the scientific perspective, such research may shed new light on the complex workings of the human brain as well as the relationship between brain states and body physiology. Finally, addressing methodological and statistical issues can enhance both fields, because such issues may result in improved scientific and statistical techniques and also contribute to theological and philosophical dialogue. Overall, this integrated field of neuroscience and religion is an important area of scholarship for the twenty-first century and beyond and may lead to a better relationship between how human beings experience God and the biostatistics.

REFERENCES


