Are Religious Experiences Really Localized Within the Brain? The Promise, Challenges, and Prospects of Neurotheology

Paul F. Cunningham
Rivier College

This article provides a critical examination of a controversial issue that has theoretical and practical importance to a broad range of academic disciplines: Are religious experiences localized within the brain? Research into the neuroscience of religious experiences is reviewed and conceptual and methodological challenges accompanying the neurotheology project of localizing religious experiences within the brain are discussed. An alternative theory to current reductive and mechanistic explanations of observed mind–brain correlations is proposed — a mediation theory of cerebral action — that has the potential for addressing what Chalmers called the “hard problem” of consciousness.

Keywords: neurotheology, brain-localization of functions, mind and body

William James (1902/1936) defined religion as a person’s “total reaction upon life” (p. 35) and speculated about the possibility of a “science of religions” at the start of the twentieth century (pp. 478–480). At the beginning of the twenty-first century and with the advent of modern electronic imagining devices that have permitted noninvasive study of the brain at a level of detail and precision not possible in James’s time, a science of religion has announced itself in an apparent reconciliation of science and religion rooted in cognitive neuroscience and evolutionary approaches to religion (McNamara, 2006). The science of religions is called neurotheology and its declared purpose is to provide some scientific basis to humanity’s spiritual nature by investigating the observed correlations between brain activity and qualities of experience associated with particular kinds of religious behaviors in a laboratory setting (Alper, 2008; Alston,
As McNamara (2009) put it: “Religious experiences are realized via the brain in human beings, and knowing how the brain mediates religious experiences can tell us something about potential functions of religious experiences” (p. 11). Studies of brain functions are interpreted to show that it is the brain by which all religious experiences are derived.

The Neuroscience of Religious Experience

Neuroscientists have identified a set of widely distributed, functionally integrated and densely interconnected regions of the brain associated with different types of religious behaviors, such as contemplative prayer (Newberg, Pourdehnad, Alavi, and d’Aquili, 2003), glossolalia (Newberg, Wintering, Morgan, and Waldman, 2006), concentration meditation (Austin, 1999; Lehmann et al., 2001; Newberg and d’Aquili, 2001), reading scriptural texts (Azari et al., 2001), recollective techniques (Beauregard and O’Leary, 2007; Beauregard and Paquette, 2006), and mantra meditation (Stigby, Rodenberg, and Moth, 1981). According to McNamara (2009), “the most important regions of the brain for studies of religious expression appear to be a circuit linking up the orbital and dorsomedial prefrontal cortex, the right dorsolateral prefrontal cortex, the ascending serotonergic systems, the mesocortical (DA) [dopaminergic] system, the amygdala/hippocampus, and the right anterior temporal lobes” (p. 127). Hallucinogens such as peyote, soma, and ayahuasca used in religious and shamanic rituals have been found to stimulate dopaminergic activity and inhibit serotonergic activity in the frontotemporal regions of the brain (Borg, Andrée, Soderstrom, and Farde, 2003; Nichols and Chemel, 2006). The practice of meditation is reported to stimulate dopaminergic neurochemical activity in trained meditators (Austin, 1999; Kjaer et al., 2002). These and other studies have given rise to a number of neurological models that presume religious-type experiences to be localized in what appears to be sharply demarcated brain regions of the temporal lobes and associated limbic system structures (d’Aquili and Newberg, 1993; Newberg and d’Aquili, 2001, p. 33; Persinger, 1987; Persinger and Healey, 2002; Saver and Rabin, 1997; Trimble, 2007).

The presumption that religious experiences are not only localized but also generated by chemical and biological processes in the brain is so thoroughgoing and complete in some areas of psychology that all cognitions, emotions, and behaviors — religious or otherwise — are considered to be similarly organically conditioned. Thus human morality is said to reside in the frontal and temporal lobes of the brain (Rankin, 2007), feelings of empathy are formed and regulated by neurons in the anterior insular cortex and neural circuitry in the limbic system (Grattan, Bloomer, Archambault, and Eslinger, 1994), the sense of self depends on its representation at synaptic junctures in the anterior temporal and prefrontal
cortex of the right cerebral hemisphere (Kircher and David, 2003; Seeley and Sturm, 2007), belief in God is hardwired into our genes and temporal lobes (Hamer, 2004; Persinger, 2001), and feelings of free will are a function of the frontal lobes (David, Newen, and Vogeley, 2008). As Francis Crick (1994) bluntly put it: “You’re nothing but a pack of neurons” (p. 3). James (1902/1936, p. 15) speculated that since scientific theories are organically conditioned just as religious cognitions and emotions are organically conditioned, perhaps scientific theories can be localized within the brain in the same way.

A wealth of clinical and experimental evidence demonstrates a close connection between various aspects of religious experience and physiological processes in the brain. But does it necessarily demonstrate that such experiences are localized, generated, or even stored in the circuit of brain regions with which they are associated? The question is of sufficient theoretical and practical importance to a broad range of academic disciplines that it deserves further examination, especially in its bearing on what is traditionally called the mind–body problem. Velmans (2009, p. 4) points out that the modern consciousness–brain problem is not one problem but many. In the present context of neurotheology, there is the problem of whether religious experiences can be adequately defined and made accessible through verbal report in a manner that permits them to be validly associated with particular brain regions. Assuming that religious experiences are really localized in the brain, there is the problem of understanding their functional relationship with underlying neural correlates. If religious experiences are not generated or stored in localized regions within the brain, then the problem of where these experiences come from and where they go so they can be later recalled needs to be addressed. The question cannot be answered simply for there are many ramifications.

**Conceptual Challenges in Mind–Brain Localization Research**

*The problem of operationalization.* Localization of any experience in the brain requires a very clear definition of the experience for which a locus can be sought. Given the varieties of religious experience chronicled by James (1902/1936), it is important to state exactly what is meant by “religious experience” and to distinguish it from other types of experiences so that the to-be-localized psychological components can be precisely identified and described. Religious experiences, like all other psychological experiences which do not take up space and cannot be physically observed, require something else which can be observed to represent or “stand in for” them (i.e., an operational definition). This ordinarily takes the form of experimental and control tasks that are used to operationalize the presence or absence of the phenomenon being studied. Recitation of a biblical text vs. reading children’s nursery rhymes (Azari et al., 2001), recollection of a prior intense religious experience vs. recollection of a prior intense interpersonal relationship
(Beauregard and Paquette, 2006), glossolalia vs. singing with eyes closed (Newberg et al., 2003), Yoga meditative relaxation vs. resting state (Kjaer et al., 2002) are some of the ways that the presence and absence of the independent variable “religious experience” has been operationalized in the laboratory setting. The operationalization of nonphysical experience into physical behaviors that can be studied in laboratory settings has resulted in a certain artificial shrinking and scaling down of what constitutes religious experience to those aspects that can be defined in operational terms, studied in an exterior fashion, and demonstrated by experiment. James (1909/1947, pp. 60, 68, 106) called this “the vice of intellectualism” — excluding from the reality of the phenomenon that is being defined everything that is not included in the concept’s operational definition. The religious experiences operationalized in the laboratory setting bear little resemblance to the “gold standard” of religious experience described in classic and contemporary accounts of mystic experience that occur spontaneously in natural field settings (Foster, 1985; Happold, 1963/1970; Hixon, 1989; May, 1991; Roberts, 1985; Woods, 1980; Yogananda, 1946/1974). There are considerable challenges involved in describing and defining, much less controlling and predicting, “the great mystic achievement [in which] we both become one with the Absolute and we become aware of our oneness” (James, 1902/1936, p. 321), especially experiences that in more natural contexts are reported to be transient, noetic, and ineffable (i.e., inaccessible for introspection and unavailable for verbal reports) [Uttal, 2000b].

How adequately the operational definition represents the construct under investigation will be always open to question. Given the variety of operational definitions of “religious experience,” it is a challenge to accumulate a conceptually coherent body of knowledge that logically relates one type of religious experience to another. The unavoidable speculative nature of religious experience and the fact that brain states change with introspection would seem to make the prospect of localizing these experiences in the brain extremely difficult, no matter now precisely the coordinates of the brain regions themselves may be mapped. As Uttal (2001) put it: “The inadequate and nebulous definition of psychological constructs poses the principle problem for localization research” (p. 92). Reserving the term “religious experience” for phenomena with specific kinds of experiential qualities (qualia) and adopting such a convention would be one way to avoid confusion and make communication about the topic easier (Forman, 1998; Hollenback, 2000, pp. 40–119; James, 1902/1936, pp. 371–372; Stace, 1960/1987). An empirically grounded phenomenological cartography of religious experiences that provides a detailed and orderly taxonomy of the full range of religious experiences would be very helpful in reaching reasonable conclusions about the prospects of localizing distinguishable religious experiences in particular regions of the brain (Fodor, 1983; James, 1902/1936, pp. 370–420; Kelly et al., 2007, p. 520).
The problem of indeterminacy. Part of the expectation that religious experiences exist as distinguishable psychological processes that can be localized within the brain comes from the finding of limited modularity of sensory/perceptual and linguistic processes, and the clear association between particular sensory receptors and specific neural structures. Religious experience, however, cannot be anchored either to a particular physical stimulus or to a specific sensory mechanism that can be traced to an anatomically separate area of the brain. Similar stimuli (sensory and nonsensory) can produce very different religious experiences and similar religious experiences can be produced by different stimuli. Stimuli capable of evoking religious-type experiences, for instance, range from single words and musical sounds to simple nature scenes or the rituals of a religious service, while the person is active or at rest, under the sway of alcohol or sober, in personal crisis or recovering from physical illness (Grinspoon and Bakalar, 1979/1997, p. 36; Wulff, 2000, p. 410). Religious cognitions, emotions, and behavior elicited by the same stimulus vary among individuals and do not represent either a single invariant response to a standard trigger or pure experiential patterns, but interact in nonsystematic ways with other variables such as expectation and setting. The sheer number of heterogeneous “triggers” of religious experience suggests a non-specific catalyst operates that stimulates or activates a general capacity of the human organism to attain religious states of consciousness which can be reached by multiple and diverse biopsychosocial pathways (Kelly et al., 2007, p. 552). In the absence of a common set of environmental triggers to anchor ideas about what makes an experience “religious,” it is difficult to assure that any laboratory experiment is actually manipulating, controlling, and measuring what it is supposed to be manipulating, controlling, and measuring (van Lommel, 2010, pp. 181–182).

Nor is there a unique neurobiological structure separate from all others that has been identified as a specific “organ of religious perception” (Saver and Rabin, 1997, p. 499). Religious cognition, emotion, and behavior are mediated by the same dorsolateral and orbital frontal cortices, the same parietal multimodal association areas, the same limbic and subcortical networks, and the same sensory/motor systems of the brain as non-religious cognition, emotion, and behavior. As James (1902/1936) put it: “Religious melancholy, whatever peculiarities it may have qua religious, is at any rate melancholy. Religious happiness is happiness. Religious trance is trance” (p. 25). If religious cognition, emotion, and behavior are mediated by the same neural systems as ordinary experience, then it would appear that what is distinctive to religious experience is not to be found in the brain but somewhere else.

The problem of modularity. Some may think that modularity is a “straw man” argument and that no cognitive neuroscientist seriously believes that religious experiences are objectively analyzable into isolated information-processing functions localizable in particular areas of the brain. Yet such an a priori assumption
is widespread in cognitive psychology today (Uttal, 2001, pp. 89–146). Visual memory, semantic information processing, short-term memory storage, executive processes, visual pattern discrimination, face perception, single word processing, directed attention, and working memory are but a few of the large number of cognitive processes that have been the object of the localization quest and associated with particular regions of the brain by PET and MRI imaging techniques (Gazzaniga, Ivry, and Mangun, 2009).

Paradoxically, cognitive neuroscientists will acknowledge that the brain is a dynamic, functionally integrated, and highly interdependent system of complex synaptic-neural networks that interact in non-linear ways (Farah, 1994; Fodor, 1983; Friston et al., 1997; Gratton and Fabiani, 1998; Shallice, 1988; Van Essen, Anderson, and Felleman, 1992). The brain's structural, organizational, and functional complexity is reflected in its high interconnectivity and redundant coding (Shepherd, 1994), feedback and feedforward processes (Stone, Vanhoy, and Van Orden, 1997), massive parallelism (Rumelhart and McClelland, 1986), interplay of excitatory and inhibitory nerve functions (Petersen, Fox, Posner, Mintun, and Raichle, 1988), and presence of multimodal neurons scattered in large and broadly distributed regions within the brain and body (Gashghaei, Hilgetag, and Barbas, 2007; Jiang, Haxby, Martin, Ungerleider, and Parasuraman, 2000; Köhler, Moscovitch, Winocur, Houle, and McIntosh, 1998). There is no solitary brain action, in other words. Brain actions may appear separate, but they are all part of other action whose integrated functioning is “designed to mediate naturally occurring behaviors” (Shepherd, 1994, p. 9).

Brain regions that appear discontinuous (e.g., sensory-motor cortex) actually merge with other regions (Stein and Meredith, 1993) with a gradient of probability function moderated by psychological factors such as attention (Posner and Petersen, 1990; Sanes, 1993; Teder-Sälejärvi and Hillyard, 1998). Areas that seem functionally demarcated (e.g., speech areas) substantially overlap with one another to a considerable degree (Lenneberg, 1974). Brain areas long considered to have a single function are now known to serve multiple functions. For example, the cerebellum, once considered to be involved solely in motor coordination, is now known have a substantial cognitive function (Thompson, 1990). A principle of organization called multiplexing describes how the same neural network can perform different functions and different neurons can perform similar functions (Shepherd, 1994, p. 434). The phenomenon of recovery of function demonstrates that dynamic changes in localization do occur and points to the capacity of brain functions to reorganize themselves and create new neural networks and brain locales after an injury or new experiences (Kaas, 1991; Kolb, 1989). Brain functions are plastic and changeful, in other words. The apparent anatomical boundaries are functional boundaries, and functions may move from one cortical area to another (Posner and DiGirolamo, 1999). As a result, it will always be uncertain what the observed activity at a particular brain region really means.
The problem of inference. Clinical and experimental observation that damage caused by trauma, tumors, infection, or ablation to a particular brain region results in an identifiable behavioral or cognitive deficit (e.g., loss of consciousness) is often used as evidence that the damaged area is the locus of the deficit (Van Orden, Jansen op de Haar, and Bosman, 1997). The logic of this conclusion, however, is problematic for at least four reasons. First, “since every cortical field is probably anatomically connected to between 10 and 20 other cortical areas (Felleman and Van Essen, 1991) one cannot from the mere localization and intensity of the activated field deduce what type of information transformation underlies the activation” (Roland, Kawashima, Gulyas, and O’Sullivan, 1995, p. 783). In complex highly interconnected systems such as the brain, the functioning of any brain region that is dependent on another region may be affected by damage to that other region, even if that other region plays no direct role in the function (Wood, 1978, p. 590).

Second, when the brain localization project is framed as a logical argument — “If a religious experience is localized within the brain (antecedent), then neural activity will occur in that area of the brain at the same time as the religious experience happens (consequent)” — an important logical fallacy is exposed. To claim that a religious experience is localized in the brain (concluding the antecedent is true) simply on the basis that neural activity is observed to accompany the experience (affirming the consequent) commits the classic logical fallacy called affirming the consequent. What this means is that the way the localization project is set up makes it impossible by the rules of logic implicit in the experiment itself to prove the hypothesis necessarily true. Even the strategy of falsification (negating or denying the consequent) will not work unless all possible variables and limiting conditions have been identified and controlled in order to perform the crucial test so that no other explanation is possible for the failure of experiment except the falsity of the localization hypothesis. This kind of control is practically and theoretically impossible using the methods of empirical science (Lakatos, 1970).

Third, alternative explanations for the observed results can always be proposed. Just because changes to the brain co-vary with changes in experience (i.e., are correlated) does not mean that some other, as yet unidentified, variable could not be present as the real cause. The brain could be a filter, a conductor, or some other mediatory mechanism that conveys rather than contains or causes the religious experience. The many “solutions” to the mind–body problem proposed over the centuries — dualist-interactionism, dual aspectism, idealism, reductive materialism, phenomenalism, emergentism, functionalism — show how there are, in principle, an unlimited number of possible explanations for the observed association of neural activity and mental activity with no one of them proven true by either logical argument or empirical experiment (Velmins, 2009).
The fact that manipulation of the brain’s electromagnetic field by electrical stimulation of local neuronal networks evokes conscious experience in some brain surgery patients may suggest that experience is retained in electrically coded data in the body’s cells (Penfield, 1958). About 10% of the 520 temporal lobe epilepsy patients studied by Penfield and Perot (1963) reported artificially induced experiential phenomena (e.g., memories, flashes of light, sounds, a sense of detachment from the body) in response to local electrical stimulation of neuronal networks. None reported any experience that resembled religious states of consciousness — a finding replicated by subsequent investigators (Gloor, 1990; Gloor, Olivier, Quesney, Andermann, and Horowitz, 1982; Halgren, 1982; Halgren, Walter, Cherlow, and Crandall, 1978). The observation that conscious experiences may be occasionally elicited by brain stimulation does not necessarily mean that the hypothetical electrically coded data must be physically contained or localized within the material portion of the cell. Alternative explanations exist. The idea that the body has a protophysical counterpart composed of “subtle” energetic systems that are not material and in which subjectively felt experiences exist as distinct, patterned series of electrical impulses may sound quite esoteric, yet is a concept basic to almost all Eastern physiology associated with Yoga and acupuncture (Benor, 2001, 2004; Evans, 1986; Gerber, 2001; Swanson, 2011) and has been used to explain certain unsolved biological problems (e.g., how morphological forms of organisms are determined, why the amputated limbs of certain species are able to regenerate) [Becker and Selden, 1985; Sheldrake, 1981, 1990].

Lastly, phenomena of psychophysiological influence (e.g., placebo response, biofeedback, hypnosis, multiple personality disorders, and stigmata) show that the mind is capable of changing the anatomy and function of the body and brain. Brain imaging studies (fMRI, PET) have documented permanent changes in the distribution of brain activity in the cerebral cortex of patients undergoing traditional cognitive behavioral therapy, mindfulness meditation therapy, and placebo treatment for a variety of psychological ailments (Davidson, Kabat–Zinn, and Schumacher, 2003; Mayberg et al., 2002). Placebo treatments have also been observed to produce changes in immune system response for a variety of physical ailments, including Parkinson’s disease (Benedetti, Mayberg, Wager, Stohler, and Zubieta, 2005; Wager et al., 2004) and cancer (Klopfer, 1957). Top–down interactions of this kind are difficult to explain using the chain of logic employed in usual bottom–up explanations (Kelly et al., 2007, pp. 117–239; Murphy, 1992, part 2).

For example, suppose experienced meditators motivated to seek “enlightenment” through the practice of meditation are asked to engage in concentration meditation in the laboratory and their brain states recorded. The mind-to-brain causal chain would proceed from (1) an alteration in the cognitive and emotional information state of the mind of the meditator to (2) the presence of a neural
representation of the meditative activity in the synaptic-neural network to (3) increased neurophysiological and metabolic activity to (4) electronic imaging of the “hot spot.” A reversal of the causal chain would require that the informational state of the network of neurons and its neural representation cause the cognitive and emotional state of the meditator — as if it is the brain, and not the person, that is the originating source of the inner motivation to seek enlightenment through meditation. The neural representation formed at the synaptic-neural network level would either be different from or similar to the representation formed in the mind of the meditator. If similar, then some form of isomorphism would seem to be required that assumed the organization of the neural representation to have the same structure as the mediator’s mental representation (see, for example, Koffka, 1935, pp. 56–67). If different, then causal explanations of how brain affects mind and how mind affects brain would appear to address two different mind–body problems. They would be two different orders of events and to confuse them would be to commit the logical error that Aristotle called a “category mistake.”

Methodological Challenges in Mind–Brain Localization Research

How we measure determines what is measured. If neuroscientists know that the brain is a distributed, dynamic, and complexly interconnected system, then why does the quest continue to locate particular forms of religious expression within particular regions of the brain? One reason is that it may not be logically possible to proceed in any different way because of the use of a basically modular research strategy. The scientific method by its nature is reductive and deterministic, and favors breaking down systems into its constituent parts, examining these parts in detail, and naming and classifying phenomenon which may not, in psychobiological fact, be separable (Slife and Williams, 1995, pp. 127–166). It is a cognitively economic and practically convenient explanatory strategy for organizing what may otherwise be an enormously complex and fundamentally wholistic and indeterminate phenomenon.

One important methodological challenge for mind–brain localization research lies in the recognition that “how we measure in large part determines what we measure — or, more precisely, what we think we are measuring” (Uttal, 2001, p. 91). Heisenberg put the idea this way: “What we see is not nature, but nature exposed to our method of questioning” (quoted in Gowan, 1975, p. vi).

The different types of memory, for instance, depend in large part on the methods used to study the retention of information. As new methods are developed, other types of memory with different characteristics are likely to emerge. The same applies to the neuroscience of religious experience. When different brain imaging techniques are used, religious experiences previously assigned to one part of the brain may be assigned to another brain region. For example, in many
experiments there have been discrepancies between the PET estimates of blood flow and fMRI measurements of oxygen use because an increase in blood flow does not occur simultaneously with an increase in oxygen consumption — the former lags behind the other (Reiman, Lane, Van Petten, and Bandettini, 2000). The delay between the two measures (with PET sensitive to blood flow and MRIs to oxygen use) gives rise to results that differ in their respective definitions of the regions that appear to be activated during any psychological experience — religious or otherwise (Greenberg, Hand, Sylvestro, and Reivich, 1979; Roland, 1993).

One illustration of how religious experiences previously assigned to one part of the brain can be assigned to another brain region using the same imaging technique is the way that criterion threshold cut-off limits are established in brain localization research. Sharp boundaries between brain activity versus no brain activity detected by various imaging technologies (e.g., PET, fMRI, SPECT) are a function of choices made by the experimenter about criterion threshold cut-off limits ($p < .05$, $p < .01$, $p < .001$) [i.e., decisions made about signals hidden in noise]. Regional levels of activation above a chosen threshold limit are deemed “significant” and thus present; levels of activation below a chosen cut-off limit are “zeroed out” and considered absent. Varying the threshold is going to have a major effect on what regions are shown to be active. With lower criterion levels, more and more regions are likely to be shown activated. A conservative criterion threshold value, for example, could hide localized activity and a liberal one suggest unique localization that is entirely artifactual. Once a decision about an acceptable threshold of activation is made, some active areas will be “missed” that might well have shown up otherwise using a different threshold value (Uttal, 2001, pp. 167–169).

**Cortical variability.** Another important methodological challenge in mind–brain localization is the many technical sources of “noise” (metabolic, neurological, electronic) contained in data obtained from imaging procedures. One source of noise in image data is cortical variability. Individual brains differ significantly in the details of their functional organization, and even gross anatomical structure may differ noticeably from one person to the next. Brain centers may vary from place to place, sometimes by amounts so large as to make images produced by a fMRI or PET look very different from subject to subject. Brain systems, in other words, are not organized the same way in every individual. This cortical variability is a source of individual differences among subjects that is artificially reduced through the extensive manipulation of data that occurs in all imaging technologies.

One common strategy to handle cortical variability and other sources of noise in imaging data is to “subtract” the image of brain activity occurring during a treatment condition (e.g., concentration meditation) from the image of brain activity occurring during a baseline or control condition (e.g., resting state) in
the same subject. The “difference image” or residual trace is presumed to indicate where the activity particular to the treatment condition is localized. While the procedure sounds logical enough, a string of assumptions underlie the validity of any conclusions that may justifiably be drawn from it (Uttal, 2001, pp. 186–195). For example, the baseline (control) and treatment (experimental) conditions are assumed to differ only with respect to the absence or presence of the critical independent variable being operationalized. Yet there is no independent assurance that the religious experience desired when the experimenter set up the experiment is actually occurring during the course of the measurements (Jack and Roepstorff, 2002, 2003).

It is known that dynamic changes in neural and metabolic activity at localized brain regions can be influenced by psychological factors such as attention. Attending or not attending to a stimulus can produce drastically different fMRI brain responses (Buechel and Friston, 1997; Friston et al., 1997; Uttal, 2000a). Emotional effects and muscular responses (e.g., teeth clenching) can also produce false localization by distorting portions of the image (Reiman et al., 2000). It is extremely difficult to isolate and control attention, emotion, and movement to assure that attention is not being paid to something else, that distracting emotions are not being felt, or that spontaneous muscle movements are not being made which distort measurements during baseline and treatment conditions.

The subtractive method also assumes that the two different conditions do not activate common brain mechanisms, and that the subtracted regions play little or no role in the experience being studied. “It is only when the subtraction is carried out that a more or less sharply demarcated response is observed. The assumption that only this difference is significant and that all of the subtracted activity is irrelevant is another weak link in a highly questionable chain of logic” (Uttal, 2001, p. 189). The widespread neurological activity that is observed to occur throughout the entire brain during both the baseline and treatment conditions before one image is subtracted from another argues against these assumptions.

A second strategy used to handle the noisy data of imaging studies is to simply “average” several different images pixel by pixel and then divide the sums by the number of images processed (Shallice, 1988). By averaging together the data obtained from a number of different subjects, a model of the brain is created that ignores the wide discrepancies which can occur between individuals. This strategy of pooling and then averaging numerical data is a statistical technique commonly employed in psychophysical experiments to estimate central tendencies and results in the creation of response functions that are totally unlike the score of any individual subject. When the values are plotted in brain imaging studies, they represent a kind of “average image” suggesting a more or less precise localization of the experience or cognitive process under study. The pooling of
a number of broad regions of activity from individuals can produce a fortuitous overlap that suggests a narrowly localized region where in fact none exists.

To reduce the variability among subjects even further, difference images are sometimes standardized in size and shape before the averaging is carried out. Elaborate and extensive standardizing and averaging often results in images still so noisy that even more sophisticated processing tools may be used (e.g., spatial frequency filters). One illustration of how elaborate and extensive this manipulation of data can become is provided by Wise et al. (1991) in an account of their procedure in a typical experiment.

The data from each subject were first standardized for brain size and shape and reconstructed parallel to the intercommissural line . . . . To increase the ratio of signal to noise and to account for the normal variability of the anatomy of the cerebral gyri and sulci between individuals, the reconstructed images were smoothed using a low pass filter of length 9 pixels on a side in the transaxial plane. As the study was designed to examine regional changes in blood flow across activation conditions, the data were first normalized for global flow differences by analysis of covariance, with measured global flow as the confounding covariate . . . and then averaged for each condition across the six subjects . . . . Subsequent statistical analysis of the data to detect significant areas of change between task and rest were performed by a planned comparison of means with a Bonferroni correction at a p level of 0.05 accounting for the effective number of independent pixel measurements by analysis of the autocorrelation function of the images. (p. 1806)

Roland, Kawashima, Gulyas, and O'Sullivan (1995) discuss the complex nature of these averages, transformations, and corrections performed in almost all currently active brain imaging laboratories. How much data validity is lost from these elaborate and extensive “data cleaning” methods is discussed by Pulvermüller (1999).

The New Phrenology

The above-mentioned conceptual and methodological challenges make the prospect of brain localization of religious experience problematic and the project requires further analysis and discussion (McNamara, 2009, pp. 10–14; Uttal, 2011; Wildman and McNamara, 2008). According to Uttal (2001), “although the ‘bumps on the skull’ idea is no longer with us, the idea that mental components exist and that they can be assigned to specific locations of the brain very much is. Indeed, the central problem facing cognitive neuroscience is how to deal with the unproven assumption that mental processes are as accessible, separable, and localizable as are the material aspects of the brain” (pp. 108–109). Even if it is granted that the brain is not homogeneous and is in fact divided into regions exhibiting functional and structural differences, the problem remains of deciding how these regions may be related to religious experiences and whether those experiences are analyzable into separable cognitive or affective
components that are localized in the brain (Fodor, 1983; Shallice, 1988; Van Orden and Paap, 1997).

Suppose it turns out that a particular sort of religious experience can be localized in a particular circuit of brain regions. “What exactly would that tell us about how our brain–mind actually accomplished their function? . . . . Imaging systems such as a PET scan can tell us only where something is happening, not what is happening there . . . . Even if we could find precise modular locations in the brain associated with well-defined psychological constructs, we still would not have solved the problem of how brain activity becomes mental activity” (Uttal, 2001, pp. 26, 70, 126). It is important to recognize that alternative assumptions and approaches are available that permit current reductionist and deterministic explanations to be seen as merely possible explanatory strategies rather than as necessary in any or all explanations of the observed correlation between religious experiences and brain activity (Slife and Williams, 1995).

The Mediating Brain

In light of the conceptual and methodological difficulties attending the neurotheology project of localizing religious experiences within the brain, alternative theoretical systems should be considered that can account for the empirical connection between the various aspects of consciousness involved in religious experiences and neurological processes in the brain in a nonreductive way. Other forms of functional dependence between mind (consciousness) and body (brain) besides a generative one are possible, including a permissive function (like the opening of a valve), a transmissive function (like a television receiver), and a mediating function (like a sensory receptor) that enables an exchange of information or energy between one domain or dimension of action and another that are functionally connected to one another but distinct (Bergson, 1912/2007, 1913; Broad, 1953, pp. 22–23; Huxley, 1954/1963, pp. 22–24; James, 1898/1956; Kelly et al., 2007, pp. 603–639; Myers, 1903/1961; Schiller, 1891/1894, pp. 293–295; Tart, 1989, 1993). As James (1898/1956) put it: “My thesis now is this: that, when we think of the law that thought is a function of the brain, we are not required to think of productive function only; we are entitled also to consider permissive or transmissive function. And this the ordinary psychophysicist leaves out of his account” (p. 15).

The brain, in other words, may not produce religious cognitions and emotions, but instead mediates them. The mediating function of the brain might be one of “straining, sifting, canalizing, limiting, and individualizing that larger mental reality existing behind the scenes” (Kelly et al., 2007, p. 22). The idea was expressed by Kant in his Critique of Pure Reason: “The body would thus be, not the cause of our thinking, but merely a condition restrictive thereof, and, although
essential to our sensuous and animal consciousness, it may be regarded as an impedier of our pure spiritual life” (quoted in James, 1898/1956, pp. 28–29). German–British philosopher F.C.S. Schiller (1891/1894) put the matter this way:

Materialism is . . . putting the cart before the horse, which may be rectified by just inverting the connection between Matter and Consciousness. Matter is not that which produces Consciousness, but that which limits it, and confines its intensity within certain limits: material organization does not construct consciousness out of the arrangement of atoms, but contracts its manifestation within the sphere which it permits . . . . For if a man loses consciousness as soon as his brain is injured, it is clearly as good an explanation to say the injury to the brain destroyed the mechanism by which the manifestation of that consciousness was rendered possible, as to say that it destroyed the seat of consciousness. (p. 293)

The mediatory brain hypothesis accounts for the close connection between various aspects of religious experience and physiological processes in the brain in non-reductive and non-materialistic terms.

The mediatory brain hypothesis has the additional advantage of advancing understanding of “anomalous” experiences reported in the psychological literature that are incompletely explained in purely materialistic and mechanistic terms (e.g., psi functioning, psychophysiological influence, automatism and secondary centers of consciousness, near death and related phenomena, genius, mystical experience, discarnate communication, out-of-body experiences, dreams) [see Kelly et al., 2007 for a review]. The mediatory brain hypothesis also holds the promise of adequately addressing in a non-reductive way what Chalmers (1997) called the “hard problem” of consciousness — explaining why conscious experience accompanies neurological brain functioning.

How the Brain “Mediates” Mind

The use of an analogy approach can help clarify understanding of the mediation theory of cerebral action. Consider, for example, the relationship between thought and language. In this analogy, the spoken or written words of a natural language (surface structure) represent the physical brain, and the meaning that is expressed or conveyed through the words (deep structure) represents mind or consciousness. Just as a thought or a feeling is structured and colored by the words used to express it, so is mind or consciousness structured and colored by the brain that transmits it. Words are used to transmit information, but the information (mind) and the words used to convey it (brain) are two different things. On the one hand, the words that are used to tell of an experience are not the experience that they attempt to describe, and so there must always be a gap between one’s thought (mind) and its expression (brain). On the other hand, human thought is so conditioned by language that it is difficult to conceive of a thought that is not verbally structured, as speakers of a second language are aware.
The relationship between thought (mind) and language (brain) also applies to the act of reading. The physical letters or words upon the page (brain) have the reality only of black marks on a white field. The letters that compose the words are symbols that have agreed upon meanings. The nonphysical information (mind) being transmitted is not an attribute of the letters or the words themselves. The information is not “contained” in the written letters any more than the thought or feeling is contained in the spoken phonemes. The printed word does not contain information — it transmits information. The physical letters or words are simply carriers of information; the information they convey is invisible. Where is the information that is being transmitted, if it is not contained in the letters or words upon the physical page? It resides within the self.

Thoughts, emotions, and dreams do not originate from the mind, but from the self who has them. It is more accurate to say that the self transmits thought to the mind and that the brain receives the thought which the mind transmits. The brain would be that very small portion of the mind that is projected into the physical dimension of basic reality and exists materially, being a part of the mind within matter. The brain would be used to express the mind, but the brain is neither the source of nor identical with the mind that it attempts to convey. It would be more accurate to say that the mind contains the brain than to say the brain contains the mind. On this view, the brain would be the physical counterpart of the mind and the means by which cognitive functions of the self (e.g., awareness and attention, memory and imagination, language comprehension and production, problem-solving and decision-making) are expressed both through the non-physical mind and the physical body. The functions of a basically non-physical psyche that is essentially independent of physical reality become physically real by means of the transductive function and filtering and focusing effects of the material brain. The mind would exist separate from and independent of the brain and would not be physically represented in the material brain, although its effects would appear within it. The mind would have its existence within the scope of the physical field but independent of it, in other words, expressing itself directly through the brain. The brain is that part of the mind that is more or less observable, reacts to physical stimuli, is manipulated with physical instruments, and measured with electronic imaging devices. Its function is to aid in the physical survival of the organism through the manipulation of material reality. The self’s thoughts, emotions, and dreams that are transmitted to the mind are structured through the brain’s actions and reactions, only a portion of which are detected by current imaging technology.

Mediation Theories of Cerebral Action

Kelly et al. (2007, pp. 607–639) present a “non-Cartesian dualist-interactionist” model that proposes a psychological filter theory of the Myers–James variety...
that builds upon the dualist models of Sherrington (1942), Eccles (1980, 1989, 1994), and Penfield (1975). In this model, “the psyche has the kind of internal organization and dynamics assigned to it by Myers and James . . . . able to function in some manner on its own” (2007, p. 608). Charles Tart (1989, 1993, 2009) presents a dualist model called “emergent interactionism” that incorporates psi as the mechanism which allows an intangible mind to interact with a tangible body. On Tart’s view, “Pure mind is something fundamentally different from the nature of the body and other physical things. What we ordinarily experience . . . is not mind by itself, ‘pure mind,’ or the body itself, but mind embodied” (1993, p. 126). “. . . Consciousness, as we experience it, is an emergent factor, a systems effect from the interaction of the brain [B] system and the mind/life [M/L] system . . . . the M/L and the B systems interact by psi” (1989, p. 210). The thought received by the mind from the self is telepathically transmitted to the brain which then transduces it into a form that is meaningful to various portions of the physically-oriented personality.

Both the Kelly at al. and Tart models incorporate the complex variable of human personality into the mind–body equation. Both models also highlight the importance of including a psychodynamic perspective into mediation theories of cerebral action that is psychologically sound and faithful to the underlying complexity of the phenomena of religious experience (Taylor, 2009). In the context of the multiplex personality theory of F.W.H. Myers (1889–1895/1976; 1903/1961, p. 27), for instance, the subliminal Self “located” in the inner psychological realm of the individual would serve the function of transmitter and the value climate of psychological reality would be the medium that takes the place of space. F.W.H. Myers’ subliminal Self is a hypothesized personality structure of extraordinary creativity, organization, and meaning — psychology’s nearest corollary to the soul (see also Roberto Assagioli’s 1965/1993, 1988/1991 concepts of the superconscious and the higher Self). Distinct, though not separate, from the outer ego of the personality, this inner self forms the personality’s larger identity, orders the intricate systems of the body, and makes available superior inner knowledge in dreams and in states of creative inspiration. Subconscious communication and adjustment is made from the mind to the body by this inner self who has connections with the entire physical organism. The inner self is the organizer of the subconscious, directs the movements of the physical body and those intimate survival mechanisms without which the body could not exist, and is that portion of human personality that survives the biological death of the body.

White Crows Abounding

Some may believe that it is more parsimonious to assume that cognitive neuroscience will provide an eventual explanation of religious experience (and
consciousness itself) in terms of cerebral anatomy and physiology given its remarkable progress than to posit a brain-independent source of religious cognitions and emotions. No one can predict with certainty what further progress neuroscience will make in understanding the actions of neural mechanisms that accompany religious-type experience. There is good reason to believe and sufficient information available at the present time, however, to indicate that a purely materialistic and mechanistic account of religious experience “is not only incomplete but false as a theory of mind” (Kelly et al., 2007, p. 27).

William James pointed out that “if you wish to upset the law that all crows are black, it is enough if you prove one single crow to be white” (Murphy and Ballou, 1973, p. 41). All one needs is a single solid finding where normal cognitive processes happen in the absence of requisite brain material or when awareness occurs in the absence of measurable brain functioning to change the way one thinks about the relationship between mind and body. Evidence exists that we have not one white crow but many, pointing to the distinctly real possibility that religious experiences may not in fact be localized in three-dimensional space within the material brain at all (Clarke, 1995; Forman, 1998, 2010).

For example, well-studied cases of individuals with clear anatomical brain damage caused by hydrocephalus (“water on the brain”) reveal a small percentage of subjects with 95% of their cranium filled with cerebrospinal fluid who have IQ’s greater than 100 (Lonton, 1979; Lorbor, 1983). Some individuals whose entire left hemisphere has been removed retain normal intelligence and linguistic functioning (Acosta, Montanez, and Leon–Sarmiento, 2002; Borgstein and Grootendorst, 2002; Smith and Sugar, 1975). Abundant senile plaques, neurofibrillary tangles, and other degenerative changes in brain regions most involved in Alzheimer disease are present upon autopsy in many healthy individuals who showed no cognitive impairment on neuropsychologic testing when alive (Davis, Schmitt, Wekstein, and Markesberg, 1999). Individuals with severe forms of neurological disorders, such as dementia and chronic schizophrenia, are known to experience brief episodes of lucid awareness shortly before death (“terminal lucidity”) [Nahm and Greyson, 2009]. Some cardiac arrest patients report experiencing lucid awareness during the temporary loss of all measurable cerebral and brain-stem activity (i.e., total absence of body reflexes, brain-stem reflexes, respiration, blood flow, electrical activity) [Greyson, 2003; Parnia, Waller, Yeates, and Fenwick, 2001; Sabom, 1998; Sartori, Badham, and Fenwick, 2006; van Lommel, Wees, Meyers, and Elfferich, 2001]. If the conventional hypothesis that conscious experience is generated by the brain were correct, then there could be no ordinary cognitive functions occurring in the absence of requisite brain matter or conscious awareness when the brain shows no measurable neural activity.

One reason why the search for the memory engram in the brain has proven so elusive is that it may not exist there in the first place. Braude (2006), Bursen (1978), and Heil (1978) argue that the hypothetical construct of the “memory
trace” is logically incoherent and empirically false for a number of reasons. The concept requires a persisting structural modification that “stores” a memory trace (or engram) that is context-dependent, accessible by multiple and diverse memory cues that were not present at the time the memory was originally encoded and can be quite different from what was put in memory, and capable of bridging temporal gaps ranging from a few seconds to a lifetime in a dynamic brain in which the fats and proteins that constitute the neurons’ cellular membrane undergo constant regeneration (Romijn, 1997). The structural features of the brain and its neuronal networks are simply incapable of performing the requisite tasks assigned to them under the memory-trace hypothesis on purely anatomical and functional grounds (Berkovich, 1993). While the brain may mediate the capacity to remember and serve as an important vehicle for the expression of memory, it is theoretically quite possible that the brain does not store memory any more than the musical instrument important for the expression of music stores the music that it plays, or the television set important for tuning into TV transmissions stores the programs that it displays on the TV screen (Grof, 1985, pp. 21–23; Sheldrake, 1990, pp. 90–93).

The proposition that religious experience (and other aspects of consciousness) cannot be adequately explained purely in terms of cerebral anatomy and physiology is further supported by research into “anomalous information transfer.” Ganzfeld studies of telepathy (Bem and Honorton, 1994; Bem, Palmer, and Broughton, 2001), research into remote viewing [clairvoyance] and precognition (Honorton and Ferrari, 1989; Tart, Puthoff, and Targ, 2002; Utts, 1996), and investigations of psychokinesis (Bengston and Krinsley, 2000; Jahn, Dunne, Nelson, Dobyns, and Bradish, 1997; Radin and Nelson, 2003) have been replicated in a number of different laboratories, by various experimenters, and across diverse cultures (Irwin, 1989; Radin, 1997; Rao, 2001; Tart, 2009). The primary implication of this research is that the human personality has inherent abilities which allow for perception without the mediation of sensory processes and the capacity for action upon matter without the mediation of recognized physical energies or mechanisms. A mediatory brain theory is compatible with psi phenomena with its implications that mind and consciousness are not restricted to the body and are of a different nature than physical matter, and that a basic independence of self and personal identity from matter operates.

Non-Cartesian Integral Dualism

If the mind is not to be found in some localized region in the brain, then how is Descartes’ classic distinction between the mind as nonextended substance and the brain as extended substance with its ensuing interactionism problem avoided? Descartes’ error was not that he recognized that the nonmaterial mind is not extended in space and that the material brain is. Thoughts, emotions,
and dreams do not exist extended in space, and do not basically exist in time, though they may be glimpsed through time. Many psychological experiences known to exist would seem not to exist, measured purely in terms of physical extension and temporal duration. Yet they do exist and to some extent manipulate physical events by their capacity to direct waking experience, maintain religious and political structures, and form human culture and civilizations (Brann, 1991). To deny the reality of what does not exist in space would be to deny much of humanity’s own heritage and abilities in the arts and sciences. One cannot deny one’s own psychological reality or the fact that thoughts exist without self-contradiction (e.g., Dennett, 1991). If thoughts seem occasionally insubstantial, it is because their substance is of a different quality. If psychological experiences appear less real than more obviously material realities, they also have at times a peculiar vividness and intensity, immediacy, and directness that is reflected in their strong and sometimes explosive emergence into the affairs of the human cultural world in a way that cannot be overlooked. Their nature is simply different from the nature of material objects that take up physical space.

While neurotheologists look for the mechanism of religious experiences within neurobiological structures and processes that take up physical space in time, transpersonal psychologists look for the mechanism in psychical structures and processes that, like the experiences they convey, do not exist extended in space or basically exist in time. According to James (1902/1936), “from the point of view of their psychological mechanism . . . [all religious experiences] spring from the same mental level, from that great subliminal or transmarginal region of which science is beginning to admit the existence, but of which so little is really known” (p. 417). That “great subliminal or transmarginal region” is not to be found in some localized region between two ears behind the forehead. The cross-cultural and historical ubiquity of religious experience and the pluralistically diverse experiential qualia of individual mystical experiences may reflect not a corresponding uniformity in organic conditioning, but instead a psyche freed from the limitations of its mediating physiological mechanism (James, 1902/1936, p. 58). Rather than confining oneself to a biological theory of mind, a nonreductive explanation of religious experience would take its grounding in the phenomena themselves and in pluralistic, subjective approaches to understanding them (Heron, 1998; Roth, 1987; Varela and Shear, 1999; Wulff, 2000, p. 430).

A modified non-Cartesian integral dualism would posit that mind and body are not entirely distinct “substances,” but instead two interweaving processes that are mental and physical at once. The body is as mental as the mind and the mind is as physical as the body, in other words. For practical purposes, there is an apparent division between mind and body. In basic reality there is no such division. No real boundaries exist, only diversity of function. Cognitive neuroscience demonstrates how one brain action can affect all others so intimately
that it is basically impossible to speak of one action in isolation. The history of psychodynamic personality theories documents how there are no sharp boundaries between conscious and subconscious portions of the whole self (Taylor, 2009). There are shadings and variations and that is all. Cutting-edge theories of cognitive science show how there can be no real division between the brain, the body, the world, and cognition (Chemero, 2009; Rockwell, 2007). The physical world rises up before our eyes, while those eyes are a part of the world they perceive. Subjective continuity is always a part of the objective reality that is experienced. Basic reality is participatory and deeply connective (Ferrer and Sherman, 2008).

**Descartes’ Creative Error**

Descartes’ error was not that he recognized a nonmaterial mind and a material body, but that he failed to recognize the material aspects of mind and the nonmaterial aspects of brain. In its modern-day incarnation, monism assumes that mind is all matter, while dualism assumes that mind has no matter, and both assert that matter has no mind — only mind experiences, not matter. Monism fails to fully acknowledge the nonmaterial reality of thought, while dualism fails to fully acknowledge thought’s physical reality. The problem with monism is not that it recognizes the unity of mind and body, but that it fails to properly acknowledge their differences. The problem with dualism is not that it recognizes the differences between mind and body, but that it fails to properly acknowledge their unity. The virtue of monism is its recognition of the physical reality of thought. The virtue of dualism is its recognition of the basically independent and separate nature of mind (or consciousness). The failure common to both monism and dualism is their inability to recognize that the body and its brain possess its own distinctly real and unique consciousness that is as alive and vital, valid and significant as the ordinary waking consciousness of the mind. How to understand the interaction of matterless mind and mindless matter raises a problem only as long as the ideas of a purely insentient brain and a purely nonphysical consciousness are maintained (Wright, 1977).

If it is a metaphysical, scientific, and creative error to separate matter from consciousness and consciousness from matter, then an adequate explanatory strategy of how brain transmits mind and why brain functioning is accompanied by conscious experience would not commit this error. It would recognize consciousness both as the agent that initiates and directs the transduction of energy into matter and matter into energy and as a quality intrinsic to matter itself (Clarke, 2003; de Quincey, 2002; Freeman, 2006; Griffin, 1988, 1997; Pfeiffer, Mack, and Devereux, 2007; Skrbina, 2005). As McGinn (1991) put it, “If neurons possessed some elementary form of awareness, then it would be easy enough to see how neurons could generate consciousness” (p. 28). The ramifications that such an alternative explanatory strategy might have for addressing what Chalmers
(1997) has called the “hard problem” of explaining consciousness — the problem of experience and why it should accompany brain functioning at all — requires further analysis and discussion.

References


