

Evaluating Reliance on Narratives in the Scientific Study of Religious Experiences

Wesley J. Wildman, Ph.D. ¹*

Patrick McNamara, Ph.D. ²

* Corresponding author.

¹ Boston University School of Theology, 745 Commonwealth Avenue, Boston, MA 02215

Office: 617-353-6788; wwildman@bu.edu

² Boston University School of Medicine and Boston VA Healthcare System Department of Neurology, 72 East Concord St., Robinson Bldg., 5th Floor, Rm. 528 Boston, MA 02118

Running head: Evaluating Narratives of Religious Experiences

Abstract

The scientific study of religious and spiritual experiences (RSEs) has relied heavily on narratives. What are the benefits and liabilities of this ongoing reliance? We employed a survey instrument for generating a multi-dimensional quantitative profile of a state of consciousness: the Phenomenology of Consciousness Inventory (PCI). In addition to yielding important information about the characteristic features of RSEs, the results permitted comparison with independent ratings of narratives. The comparison revealed that significant information about the phenomenological qualities of recalled experiences drops out when moving to narrative interpretation. We conclude that scientists studying the phenomenological features of RSEs should approach narratives with care, relying on them where they prove effective at communicating experiential content, and supplementing them with other approaches where narratives are less effective modes of communication.

Key Words: religious experiences, spirituality, narrative, reliability, phenomenology, consciousness, methodology, Phenomenology of Consciousness Inventory, PCI, Linguistic Inquiry and Word Count, LIWC 2001

Introduction

Billions of people across all nations describe themselves as religious. Many more readily describe themselves as spiritual even when they don't readily acknowledge any religious affiliation and are skeptical about religious claims. Most of these people appear to have deeply meaningful experiences directly related to their religious or spiritual beliefs and practices. The characteristic appeal of these experiences is that they help to knit the varied parts of life together in an emotionally and intellectually satisfying way. Even when they are terrifying such experiences typically reinforce a religious or spiritual outlook on life. A person narrates such an experience using the plausibility structures and conceptual frameworks furnished by their religious or spiritual beliefs and practices, even when the experience in some ways defies comprehension within extant patterns of understanding. Such experiences can be bracketed for the purposes of studying some social and intellectual features of religious belief and behavior, but RSEs are a central aspect of human religiosity and spirituality.

The value of studying RSEs is not difficult to see. For the large majority of people worldwide, religion influences access to resources, education, and health; it affects who you will marry, who your friends will be, and what your long-term life plans are. It identifies who your enemies are and whether and why you must fight them. Because of the importance of religion to most people's lives, and the centrality of RSEs to religion, RSEs have received a great deal of scholarly and scientific attention. Psychologists of religion have focused on their personal and social effects, and philosophers of religion have analyzed the reliability of RSEs as sources of confirmation for religious belief. To understand RSEs is to gain insight into the almost universal appeal of religious beliefs and behaviors, and to shed light on the enormous existential and social power of religious movements.

Narrative Approaches in the Study of Religious Experiences

RSEs are almost universally studied using narratives supplied by people who accept invitations to share them. Many RSEs are recorded for posterity in the writings of mystics or other religious authors. The largest collections of RSEs, such as those of Edwin Starbuck, William James, and the Sir Alister Hardy Institute, were built through collecting accounts submitted in response to newspaper advertisements, word-of-mouth invitations, personal encounters, and (more recently) web-based outreach. Often these invitations include prompts to ensure that elements thought important by researchers but apt to be neglected were in fact included as much as possible.

It was not easy in the late nineteenth century getting the scientific study of RSEs started. Christopher White's fascinating account explains why (see White, 2008, on which the following material depends). People were reluctant

to narrate their experiences. Early researchers, such as Starbuck, experienced recrimination and outrage from those who felt that his requests for detailed narratives of RSEs improperly reduced wondrous living experiences to lifeless text. For example, while a student at Harvard Divinity School and conducting research under the supervision of William James, Starbuck circulated a request for narratives of conversion and sanctification experiences. He asked specifically for details about personal motivations for the experience, social influences on it, contextual features of it, belief in its supernatural origins, strange alterations of experience during it, and aftereffects of it including relapse. James warned Starbuck that this would not sit easily with conservative New Englanders but permitted the study to go ahead. Leading liberal reformer Thomas Wentworth Higginson demanded to know whether James had truly signed off on the project, which he thought amounted to “moral and spiritual vivisection.” And Starbuck’s early results were greeted by many with dismay and anger in the Divinity School context. Yet Starbuck’s effort on this and future studies eventually paid off. James relied heavily on his student’s work in his *Varieties of Religious Experience* (James, 1902). Eventually psychologists of religion became more comfortable with asking for and supplying detailed narratives of RSEs without fear that the meaning and value of such experiences were somehow being eliminated in the process.

After this rocky start, researchers compiled vast collections of RSE narratives. In most cases, written accounts were solicited by asking people to identify experiences that they themselves (rather than the researchers) considered religious or spiritual. Most of the insights we have into the nature of RSEs derive from these collections. James based his analysis in *Varieties* on extreme and colorful experiences, drawing on Starbuck’s and his own collections of narratives. Documenting the extremes sheds light on the range of experiences possible in human life—possible for some people, at least—and helps to explain the certainty with which people undergoing such exceptional experiences hold beliefs deriving from what they take to be their cognitive content. Hardy took a broader view than James and based his analysis on a much wider range of RSEs, including everyday experiences that people were prepared to call religious or spiritual (see Hardy, 1979). He analyzed thousands of narratives for their distinctive features, producing a rich descriptive classification that continues to guide the field. The analysis of mystical experiences, which (despite James’ and Hardy’s wider interests) was the leading edge of the philosophical wing of the study of RSEs during the twentieth century, depended almost solely on narratives found in the writings of mystics long dead. Recently interviews with living mystics and functional imaging brain scans of meditating monks

in mystical states have entered the fray but almost every aspect of the study of RSEs depends on narratives and almost every insight we have into RSEs derives from analysis of narratives.

This cursory survey underlines the actual importance of narratives in the study of RSEs. It also makes clear why this is fundamentally a good and useful state of affairs. Indeed, a researcher might prefer narratives over any other conceivable approach on the grounds that, without narrative texture, the actual meaning of an experience in a person's life and social context is lost and subsequent interpretation is severely hobbled (see Yamane 2000, who reminds sociologists of the dangers of isolating the study of RSEs from their concrete meaning in personal lives, as manifested in narratives).

Evaluating Strengths and Weaknesses of Narrative Approaches

Despite the undeniable practical importance of narrative approaches to studying RSEs, and perhaps because of it, researchers are wise to ask about their potential weaknesses. The evaluation of narrative methods has been extremely difficult for the lack of anything with which to compare them. Complementary methods of studying RSEs have become increasingly well established, however, so comparative evaluations should become steadily more tractable, in principle. Consider four major alternatives to narrative methods in the study of RSEs, in order of increasing usefulness.

First, functional imaging studies help researchers generate neural models of RSEs—an extraordinarily difficult process that has barely begun. But they offer no insight into the strengths and weaknesses of RSE narratives beyond the painfully obvious twin points that brain scans know nothing of semantic content and narratives don't include accurate information about the neural realization of RSEs.

Second, linguistic analysis seeks patterns in the linguistically measurable features of narratives expressing meanings that may or may not be obvious in the narrative structure itself. For example, a match between the presence of vocabulary bearing on meaning in a narrative and a subject's affirmation of the meaningfulness of the experience being narrated would rightly increase confidence in the ability of narratives to convey the phenomenal features of the experience. Similarly, if linguistic analysis suggests rich social embedding of the experience that was not present in the way the narrative is structured, we would be justified in wondering whether the narrative was doing a good job of communicating the experience in respect of its social texture.

A series of studies using linguistic analysis techniques has demonstrated that measurable features of the way people use words when they write and speak do in fact correlate with mental and physical health, with mood and

intention, and with state of mind when writing or speaking (Gottschalk & Glaser, 1969; Rosenberg & Tucker, 1978; Stiles, 1992; Pennebaker, 1997; Pennebaker, Mayne, & Francis, 1997; Campbell & Pennebaker, 2002; Lepore & Smyth, 2002). For this reason, linguistic analysis is used routinely in security applications. Such techniques are underdeveloped in relation to RSEs, by comparison with health and security contexts. In particular, the norms required to make sense of word count statistics from a linguistic analysis do not yet exist for RSEs. As a result, such instruments have to be deployed comparatively in this domain—that is, they should be applied to narratives of both RSEs and control experiences so that salient differences can be identified.

Third, a range of “live” techniques from experimental psychology are proving increasingly useful for studying RSEs, from inducing such experiences in constructed environments to priming experiments. Such techniques can be made fruitful for evaluating RSE narratives because live experiences can also be narrated, and analysis of narratives then compared with analysis of results from suitably designed experiments. As a speculative example, consider that the effect of priming on RSEs induced using visual imagery may or may not be detectable in expert interpretation or linguistic analysis of the corresponding RSE narratives; that would tell us something about the capacity of narratives to convey priming effects.

Fourth, “retrospective” survey instruments are flexible methods for generating meaningful evaluations of RSE narratives. Using questionnaires about recalled RSEs that are also narrated, researchers can determine what kinds of information tend to be offered spontaneously in RSE narratives and what kinds tend not to appear without special prompting. Like “live” techniques, survey instruments are only possible for living subjects. Unlike many “live” techniques, survey instruments are retrospective in character, asking participants to answer questions about experiences in their recent or far past, and thus effectively soliciting accounts of the meaning-for-them of the RSEs in question. The match between narratives and survey instruments in this respect is helpful but it is important to keep in mind that neither narratives nor surveys access RSEs in a pure way, unaffected by the way they are remembered as meaningful (on methodological challenges facing the scientific study of RSEs, including the problem of access to first-person states of consciousness, see Lutz & Thompson, 2003; Varela, 1996; Wildman & McNamara, 2008).

In principle, a variety of survey instruments might furnish a basis for evaluating the strengths and weakness of RSE narratives. The particular angle of each instrument would shed a distinctive light on the narratives. For example, suppose Hood and Rosegrant’s Religious Experience Episodes Measure (REEM; see Hill & Hood, 1999:

220-224) was deployed in conjunction with a request for a RSE narrative. The shortened and revised version of the REEM (Rosegrant, 1976) asks participants to rate on a 9-point Likert scale the similarity of the RSE in question to each of series of ten narratives drawn from James's *Varieties*. When a narrative is read in light of the REEM results, it is quite conceivable that a participant could rate as very similar to their own experience one that to a third-party reader bears little relationship to it. An accumulation of such occurrences would help to isolate the features of RSEs that tend to be important to those who experience them yet suppressed in narratives. Such results would be impressionistic but informative.

With this kind of comparative approach in mind, but aiming for a more precise and detailed comparison, we employed a survey instrument that specializes in helping participants identify the phenomenological characteristics of their states of consciousness. Because this instrument produces a multidimensional quantitative profile of the phenomenological features of a recalled state of consciousness, we reasoned that it would give us a rich basis for assessing the quality of RSE narratives.

Phenomenology of Religious Experience

Our deployment of an instrument focusing on the phenomenological features of a recalled experience (rather than behavioral consequences, cognitive structures, or causal triggers, for example) is not unproblematic, owing to the intrinsic difficulties associated with studying the phenomenology of experiences accessed through recall procedures. These difficulties can be appreciated from two angles.

On the one hand, literature in the phenomenology of religion has invested considerable energy in discerning the distinctive features of religious consciousness. This has been done using introspection, in some cases, but always with heavy reliance on narratives of RSEs. Unfortunately, the diversity of phenomenological analyses is extremely pronounced and even the most well-known proposals are difficult to reconcile with one another. This has allowed debates to carry on for decades with no clear resolution. For example, full recognition of the cultural, historical, and personal context for mystical RSEs (Bagger, 1999; Katz, 1978; Proudfoot, 1985) rightly corrected the decontextualized and apologetically driven approaches that previously dominated the study of mysticism (classically, Otto, 1917/1923; Schleiermacher, 1799/1893; Stace, 1960; Zaehner, 1957). But contextualism also produced a kind of particularism that threatened to exclude recognition of any patterns in or types of RSE (see the analysis in Burhenn, 1995). The cognitive and emotional similarities of human beings across cultures may be limited but they are significant enough to make such an extreme particularism as implausible as decontextualized abstract

theological analyses of RSEs. The intractability of these disagreements about how RSEs register context suggests that phenomenologically focused research continues to be deeply problematic. Similarly intractable debates persist over whether RSEs are substantially similar or different across cultures, eras, and individuals; whether they occur to religious and non-religious people in similar ways; and whether they have recurring essential features.

On the other hand, the states of consciousness associated with RSEs have simply not received the attention in consciousness studies or in cognitive neuroscience commensurate with the prominence of religion in human life. The recent flourishing of works on the cognitive science of religion promises a change in this situation (Boyer, 1999, 2000; Andresen, 2001; Atran, 2002; essays in McNamara, 2006; McNamara, 2009). Even here, however, the emphasis is often not on the phenomenal qualities of states of religious consciousness, presumably in part because of the difficulties involved.

At root, these two difficulties—apparently intractable disagreement among phenomenological analyses and neglect of phenomenology within the cognitive science of religion—are both due to the fact that the study of RSEs must rely at least initially on subjective reports as primary data sources. It is difficult to penetrate beyond subjective reports of experiences to the phenomenological characteristics of the experiences themselves (for an insightful analysis of the problem of subjective reporting in consciousness studies as a whole, see Revonsuo, 2006). For instance, it is difficult to know whether impressionistic similarities among subjective reports reliably indicate the same type of experience, even when those impressionistic similarities are identified by experts.

These challenges indicate that narrative methods for investigating the phenomenology of RSEs, despite being vital to the scientific study of religion, are not yielding the fruits researchers might expect. Nevertheless, the tradition of investigators employing narrative approaches to access the phenomenological properties of RSEs is well established (Alschuler, 1993; Currie et al., 1982; Donovan, 1979; George, 1995; Hallman, 1963; Hardy, 1979; James, 1902; Kason, 1997; Kokoszka, 1992; Moehle, 1983; Murphy, 1992; Oxman et al., 1998; Stace, 1960; Tart, 1975; Wildman & Brothers, 1999; Wildman, 2002). The reliance of such studies on expert analyses of narrative reports of RSEs, in conjunction with the difficulties just described, naturally raises a crucial question. Is salient information about the phenomenological qualities of states of religious consciousness reliably preserved in verbal reports of RSEs?

This question has led us to begin a research program that combines both third-party analysis of RSE narratives and participant rankings of phenomenological features of RSEs (Wildman and McNamara, 2008). The particular

instrument we used in the current study to identify key phenomenological properties of states of consciousness is the Phenomenology of Consciousness Inventory (PCI, described below). The PCI has not been applied extensively to the study of RSEs, though it has been used in the study of religious meditation and a variety of altered states of consciousness that typically possess spiritual meanings for those who have them. The PCI has been used with considerable success in its original domain of application, which was to investigate hypnotic susceptibility. There it produced predictions of the hypnotizability of experimental subjects that comport well with existing measures such as the Harvard Group Scale of Hypnotic Susceptibility and the Stanford Hypnotic Susceptibility Scales; additionally, the PCI gives a better understanding of the characteristics of conscious states in hypnotically susceptible people.

In adopting this instrument we have had to adapt slightly the method of its use to accommodate the fact that we ask our participants to recall experiences a longer time ago than in the original studies. In short, our use of this instrument is quite ambitious and many issues surrounding its application to RSEs remain unresolved. For the limited purposes of this study, however, the way we deploy it is sound, and the suggestiveness of our results for further study is clear.

The meaning of RSEs and ordinary experiences

Our study asked participants to recall, narrate, and score three types of experience, as described below: happy, ordinary, and religious. There are many definitional questions surrounding these terms, particularly “religious” and “ordinary,” and thus some introductory remarks are in order.

An example from among the narratives of happy experiences we collected is: “I went on a trip to Bermuda with my wife and my kids. I wasn’t very enthused that the father-in-law and the mother-in-law were coming. They come like the package deal. But I accepted that anyway, so they came. And we had a great time. It was nice, enjoyable, relaxing. Spent time with the kids, the family, my wife. Just doing things, you know, ’cause I’m always working. Also it was my daughter’s birthday.”

An example of a narrated ordinary experience is: “One of the traditions in my household is that I make popcorn on Sunday night. It’s a real ordinary event. If popcorn isn’t made by 7pm someone is asking ‘where’s the popcorn?’ So on a typical Sunday night I will go into the kitchen between 6 and 7 in the evening and get out the big silver bowl and an air popper we have had for who knows how many years. I get the popcorn out of the refrigerator. We usually get about two tablespoons of butter and I put it in the microwave first. Then in the air popper I would pour in exactly a quarter cup of popcorn and it usually takes no more than 3 or 4 minutes before it is popped. And when that is done

I dribble the butter into the pan as it is coming out of the air popper so there is an equal amount of butter throughout. And after it is all buttered I always take half a dozen popped corns, throw them into the butter dish, and sap up the last pieces of the butter and eat them. Rewards for being the maker. Then I dish out the rest to all interested parties. And then I take the remains and go upstairs and watch TV and eat popcorn.”

The RSEs participants offered were diverse. Some were recognizably religious, such as the following. “I’m standing in front of a huge door, an open door, which is the entrance to the church I belong to. As I enter I know that I have to take my hat off and make a cross and walk in. I was anxious to walk in because I hadn’t been to church for two years. And I experienced this feeling I had every time I used to walk into church. I knew that it was going to be strong. As soon as I walked in, the smell of the candles—the special candles that we light in the church for good luck, for health, for just acknowledging all the dead people, just thinking of everyone that had touched you or been in your life and had gone—all the smell kind of hit me as I was walking into church. I knew I had to get my own candles to light. Passing by some other religious people standing there in the quiet, I walked all the way down to where the icons are and I just stood there staring at the pictures and all the paintings of all the religious scenes and Jesus. I just stood there and I lit my candle. I felt like I was collecting or ingesting this good energy that was coming to me from the entire place. I had been to church before but I had never felt like this. So that was something that I got addicted to and I had to do it almost like a little ritual every day.”

Other RSEs occurred in non-religious settings, such as the following. “I went running on the edge of the mountains in Utah, near the University of Utah, where it comes right up to the mountains. And it was nice outside, it was green and the sky was blue. And I was exercising. And I felt like, ‘Yeah, there’s something more.’ It could’ve been the euphoria of exercise.”

Following the standard methods of collecting RSEs, we allowed participants to define what the terms “religious” and “spiritual” mean to them rather than stipulating the meanings of these terms. The problems surrounding the formal definition of such terms—problems infamous in the academic study of religion—make this a wise procedure. The result was narratives of apparently quite diverse RSEs, as illustrated in the two sample quotations above. This diversity reflects (1) individual differences in participants’ usage of the key descriptor terms “ordinary”, “happy”, and “religious”; (2) the clarity of the experimental protocol that evoked the narratives; and (3) the features of the underlying experiences themselves. We sought to get behind the limitations imposed by (1) and (2) in order to clarify the relationship between the phenomenological features of an experience and the

features of the narrative produced when a participant recalls that experience. There are several reasons to think that this should be possible, at least to a degree sufficient for the purposes of this study.

First, and most obviously, the key words “happy”, “ordinary”, and “religious” are not hopelessly variable in semantic scope, at least as far as our study population is concerned. As the large collections of RSEs show, prompting for RSEs using people’s own definition of the key terms produces intelligible, analyzable narratives that seem naturally to belong in a meaningful class with fairly clear-cut subclasses and distinctive patterns of phenomenological, cognitive, emotional, behavioral, and social content. The same is true of happy experiences: the affect marker seems more or less universal. Admittedly ordinary experiences are more problematic. Any protocol that asks for more than one type of experience risks priming participants to furnish an experience that is “ordinary” simply in the sense of not being like the other experiences they have shared.

Second, however, we can employ an experimental protocol that minimizes the problem with vagaries of meaning, especially for “ordinary” experiences. Our protocol employed several checks to minimize this problem. The recall prompt for ordinary experiences offered a vague definition of “ordinary” as follows: “I want you to tell me about a personal experience that you remember as being ordinary—neither particularly religious nor particularly intense.” We introduced a challenge procedure when the ordinary experience appeared not to be of the “neither particularly religious nor particularly intense” variety. We asked for a subjective rating of intensity to confirm that the experiences were in fact not intense. We randomized the order in which experiences were obtained, allowing us to evaluate whether that order made a substantive difference in the nature of the experiences offered (it did not). And we used the PCI to obtain analyses of the class of ordinary experiences (and the other classes also) that are far more fine-grained than is possible with impressionistic assessment of narratives. These efforts did not yield a perfect protocol; indeed, we believe that eliminating unwanted demand characteristics is virtually impossible in any contrastive recall procedure of the type required for this study.

Third, with happy experiences and ordinary experiences providing two different types of baseline experience, we uses the PCI to investigate whether the class of RSE experiences contrasts meaningfully with those two baselines. That is, we examined the phenomenological content of the three classes of experiences to determine whether their significant differences among them occurred in PCI dimensions that make theoretical sense. For example, we could evaluate whether affect dimensions were highly positive for happy experiences but not for others, and whether alterations of experience were low for ordinary experiences but higher for RSEs.

Finally, we deployed linguistic analysis as a further check on the stability of the major groupings of happy, ordinary, and religious experiences.

Using this approach, we think that we may have minimized inevitable problems and maximized the value of the results obtained.

Methods

We asked a convenience sample of healthy volunteers to recall and narrate a religious experience, a happy experience, and an ordinary experience. Participants rated the degree to which a number of phenomenological features were present in each of the three experiences. We later asked two raters, one an expert and one a novice, both blinded to the purposes of the study, to rate participant narratives using the same instrument used by the participants themselves. We also performed linguistic analysis on the narratives. We expected significant differences in phenomenological ratings on religious versus happy versus ordinary experiences, but made no predictions about the direction of difference. Centrally for the purposes of this paper, we expected significant differences between participant ratings of the phenomenological properties of recalled RSEs and third-party ratings of RSE narratives. Analyzing participant ratings allowed us to determine that the RSEs our participants reported do possess distinctive phenomenological features. Analyzing these results against third-party ratings of RSE narratives showed that there are significant and intelligible differences between the phenomenological profile of RSEs obtained from participants and the phenomenological profile of the same experiences produced by rating of narratives.

Participants

Thirty-nine healthy participants (15 males, 24 females) between the ages of 18 and 79 (mean age 34) were recruited from the Boston area. This study was part of a larger study on religious social cognition, which was approved by the VA New England Healthcare System, Jamaica Plain Campus, IRB. All volunteers were informed of the nature of the study, consented to participate, and completed a demographics questionnaire. Then each participant completed a number of tests, described below.

Measures

In addition to the PCI (below), we used two instruments to ensure that our participant sample was not extremely atypical regarding mood or religiosity.

Depression, Anxiety, and Stress Scales (DASS)

The DASS, developed by Lovibond and Lovibond (1995), is a 21-item questionnaire that assesses depression, anxiety, and stress on three separate subscales. There are seven questions for each subscale. Participants were asked to indicate on a scale of 0 (did not apply to me at all) to 3 (applied most of the time) how much each question applied to them over the past week. It follows that scores for each scale range from a minimum of 0 (low depression, low anxiety, low stress) to a maximum of 21 (high depression, high anxiety, high stress). Crawford and Henry (2003) and Antony et al. (1998) have reported excellent reliability, validity, and other psychometric properties for the three subscales of the DASS.

Our population yielded means (standard deviations) of 3.26 (3.75) for the DASS Depression subscale, 2.67 (2.67) for the Anxiety subscale, and 6.21 (4.76) for the Stress subscale, and 12.13 (9.844) for the DASS Total (minimum 0, maximum 63). The population was close to normally distributed around these means. None of our participants was unusually depressed, anxious, or stressed by DASS norms.

Brief Multidimensional Measure of Religiousness/Spirituality (BMMRS)

The BMMRS was developed by a panel of experts on religion and health with the assistance of the Fetzer Institute (1999) and the National Institutes of Health and Aging. The BMMRS contains 38 statements in Likert scale formats, along with several yes/no, selection, and narrative responses, covering 11 domains of religious behavior, belief, and experience. Examples of questions and statements on the BMMRS include “To what extent do you consider yourself a religious person?” or “I feel God’s presence.”

For this study, we read to participants the following preface prior to reading the directions that are on the BMMRS and administering the instrument itself:

This is a questionnaire that assesses your religious beliefs and spirituality. This instrument uses the language of God, religious congregation, and religious experience, but we understand that this does not apply naturally to everyone. If you have trouble with the concept “God,” please translate that concept into “Ultimate Reality” or “Ultimate Meaning.” If you have trouble with the idea of a religious congregation, then please translate that into the group of people who are most spiritually important to you. If you have trouble with the idea of religious experience, then interpret it as a spiritual experience.

The BMMRS Overall Self Rating dimension yielded a mean of 4.92 and a standard deviation of 1.48 for our 39 participants (the minimum is 2 and the maximum is 8), with the population close to normally distributed around the

mean. This indicates that the participants in our convenience sample were not especially religious or spiritual, nor were they especially hostile to religion or spirituality.

The Phenomenology of Consciousness Inventory (PCI; Pekala, 1991)

In order to obtain participant and third-party ratings on an extensive range of phenomenological properties, we used the Phenomenology of Consciousness Inventory (PCI) developed by Pekala and associates (Pekala, 1991). The PCI is a self-report 53 item questionnaire on phenomenological aspects of a selected state of consciousness. The subject is asked to complete the inventory while recalling a previous state of consciousness—thus it is one of a small class of *retrospective phenomenological analysis* tools. This method of recall is similar to that required to respond to a request to produce a narrative of any past experience. Thus, our prompts are akin to those used to produce the large databases of RSE narratives, discussed above; this similarity suits our purposes.

The PCI yields a quantitative profile of the contents and quality of personal consciousness along 26 measures, grouped into 12 major dimensions (positive affect, negative affect, altered experience, imagery, attention, self-awareness, altered state of awareness, internal dialogue, rationality, volitional control, memory, and arousal). Pekala and colleagues validated the PCI for the state of “sitting quietly with eyes open for four minutes” on a population of 112 subjects, and it was further tested for the state of “sitting quietly with eyes closed” with a sample size of 217. Pekala provides detailed validity and reliability data on the instrument (Pekala, 1991, chapters 5-7). The PCI has demonstrated excellent internal consistency, yielding coefficient alphas between .70 and .90 (Pekala, Steinberg, & Kumar, 1986).

The PCI has been repeatedly tested and its domain of validity extended in numerous studies including hypnotic susceptibility, out-of-body experiences, eating disorders, shamanic states, drumming, creativity, phobias, depression, anti-social feelings, kleptomania, epilepsy, and meditation (see Pekala & Kumar, 1984; Pekala et al., 1986; Pekala & Kumar, 1987; Pekala & Kumar, 1988; Kumar & Pekala, 1988; Kumar & Pekala, 1989; Pekala & Kumar, 1989; Pekala & Nagler, 1989; Pekala et al., 1989; Maitz & Pekala, 1991; Forbes & Pekala, 1993; Spinhoven et al., 1993; Szabó, 1993; Hand et al., 1995; Pekala, 1995a; Pekala, 1995b; van der Linden, 1995; Forbes & Pekala, 1996; Kumar et al., 1996a; Kumar et al., 1996b; Maurer et al., 1997; Pekala & Forbes, 1997; Venkatesh, 1997; Angelini et al., 1999; McCloskey et al., 1999; Huang et al., 2000; Varga et al., 2001; Pekala, 2002; Grant, 2004; Hutchinson-Phillips et al., 2005; Manmiller et al., 2005; Pekala et al., 2006; Johanson et al., 2008; Rock et al., 2008).

The 26 dimensions of the PCI, with their meanings and the abbreviations used for them throughout this paper, are shown in Table 1. The 12 major dimensions are in italics.

Pekala originally used the instrument to assess the phenomenological features of subjective states of consciousness very recently recalled. He had large groups of subjects sit quietly for a few minutes and then report on the qualities of that experience using the PCI. We modified this procedure in the way described below. Most importantly, we worked with participants individually because we also collected narratives, we had them complete the PCI after recalling each of three different experiences (randomizing order), and we asked them to recall experiences not from mere minutes earlier but possibly from weeks or months earlier. Similar adjustments to Pekala's original procedure appear in previous studies using the PCI.

The PCI's Internal Reliability Check

The PCI includes five Reliability Item Pairs (RIP) used to assess the consistency of a participant's answers. Each pair consists of two virtually identical questions with reverse polarity, widely separated on the questionnaire. Pekala's guideline in prior studies was that, if the average RIP score exceeded 2.0—that is, if the difference between the answers to the paired items, which range from 0 to 6 on a 7-point Likert scale, exceeded 2.0—then the participant might well be judged unreliable in filling out the PCI questionnaire.

Our participants produced an average RIP score of 0.738 for happy experiences, 0.686 for ordinary experiences, and 0.938 for religious experiences, with a total overall RIP average of 0.787. This indicates high average reliability. None of the participants produced an average RIP score for all three experiences that exceeded the cutoff of 2.0. On this basis, we excluded no participants from the study.

Procedure for Recall of Experiences

All participants were asked to recollect three separate experiences: a religious experience (in the sense that the participant considers it religious or spiritual, following the Hardy Institute's approach to soliciting accounts of such experiences), an ordinary experience (neither particularly religious nor particularly intense), and a happy experience. The order of experiences was randomized. Most participants were able to recall an event from within a year of the testing date. After the event was narrated, an approximate date of the event was requested. They were then asked a series of questions about the event such as how subjectively intense the event was, when the experience occurred, who was involved, how often they had thought about it, how often they had shared it with others, etc. The questions were designed in part to help subjects recall more details about the event, and in part to gather data relevant to

estimating the problem of recall over periods up to several months. After these probe questions were answered subjects were asked to re-narrate the event.

After narrating an experience, dating it, answering questions about it, and then re-narrating it, participants completed the PCI. This recall procedure was repeated for each of the three experiences. The recall procedures—first narrative, answers to questions, second narrative—were tape recorded and then later transcribed, producing narratives and answers to questions that were the basis for later analysis. Also, for each experience, latency to retrieval of the experience was measured from after the participant had heard instructions until the participant began sharing the memory.

Analysis of Narratives

Using the transcribed narratives, two third-party raters not informed about the purpose of the study—an advanced graduate student in religious studies with extensive training in the academic study of RSEs, and an undergraduate student with no training in the academic study of religion—rated the 117 narratives that our 39 participants produced using the PCI instrument. For each narrative, this required the independent rater to read the first version of the narrative, read the answers to probe questions about the experience, read the second version of the narrative, and then assign a score to each of the 26 PCI dimensions for the experience underlying the narrative material they read. Raters had no access to demographic, DASS, or BMMRS information.

Our decision to allow raters access to two versions of each narrative, as well as probe questions about the underlying experience, is contestable. By furnishing information beyond a single narrative we potentially bias our results toward the conclusion that narratives are more effective means of communicating RSEs than they actually are in practice. We elected to do this because our participants were not particularly skilled narrators, whereas those volunteering RSE narratives for large collections, such as that of the Hardy Institute, did so in part presumably because they felt confident about the quality of the narration that they produced. Access to two versions of the narratives and to probe questions probably offers about the same amount of information as a polished, written narrative. Moreover, supplying the extra information to third-party raters before they rate the PCI categories makes the comparison with participant PCI scores more meaningful. Most importantly for the purposes of this study, supplying the extra information will tend to increase the impression of reliability of narratives, if anything, thereby making mismatches between participant and third-party PCI ratings that much more informative about the strengths and weaknesses of narrative communication.

We performed linguistic analysis on the narratives using the Linguistic Inquiry and Word Count (LIWC 2001) program designed by James W. Pennebaker, Roger J. Booth, and Martha E. Francis. The LIWC measures:

- standard linguistic features (e.g. words per sentence, question marks, long words, pronouns),
- psychological processes (affective or emotional processes, cognitive processes, sensory and perceptual processes, social processes),
- relativity (time, space, motion), and
- personal concerns (occupation, leisure activity, money and financial issues, metaphysical issues, physical states and functions).

Statistical Analysis and Multiple Comparisons

Our aim was to determine whether third-party interpreters of RSE narratives could produce PCI ratings matching those supplied by participants who were recalling the original experiences. For this purpose, we calculated Pearson correlations, which identify the closeness to a linear fit between two sets of ratings. A significant correlation would indicate a *relative match* between participant and third-party ratings. That is, even if the means of the ratings differed—say because of a muting effect when moving from recalled experiences to interpreted narratives—Pearson correlations could still detect whether third-party raters were able to make the same kinds of discriminations in *relative* strength of phenomenological features by interpreting narratives that participants produce by recalling experiences. This yields the best approach to detecting narrative adequacy when it is present.

The question of multiple comparisons is controversial in relation to any multidimensional instrument. Allowance for multiple comparisons is clearly necessary when comparing groups across widely diverse characteristics (i.e. statistically fully independent features of the groups). Equally clearly, allowance for multiple comparisons is misleading when comparing groups across strongly similar characteristics (i.e. statistically dependent features of the groups). Consensus on these extreme cases is strong in statistical theory but there are many mixed cases that are less clear-cut. The PCI presents such a mixed case. On the one hand, the 26 PCI dimensions depend on 21 independent measures (recall that 5 major dimensions are composites of sub-dimensions), which suggests that allowance should be made for 21 independent multiple comparisons. On the other hand, as with any multidimensional instrument, all of the PCI dimensions measure features of a single, rather narrow phenomenon. That is, the PCI measures dimensions of a single state of consciousness and does not ask about behaviors, cognitive content, religious beliefs, contextual triggers, or any of a host of other questions that one might well ask about RSEs.

This suggests that no allowance should be made for multiple comparisons. Pekala's original studies and most since have taken the latter line, that the multidimensional PCI does not require allowance for multiple comparisons—more precisely, that allowing for multiple comparisons is misleading.

How does this issue affect the current study? On the one hand, if we treat the PCI dimensions as fully statistically independent measures of diverse phenomena and allow for multiple comparisons, then we would tend to find fewer significant correlations between participant and third-party PCI ratings, making RSE narratives seem less reliable. On the other hand, if we were to follow Pekala and not make allowance for multiple comparisons, then we would tend to find more significant correlations between participant and third-party PCI ratings, making RSE narratives seem more reliable.

It is likely that a multidimensional instrument, even one measuring a single circumscribed phenomenon such as a state of consciousness, requires some allowance for multiple comparisons, particularly when the instrument is well designed and the instruments' various dimensions have been shown to vary significantly independently of one another (which is the case for the PCI). It is also likely that treating all dimensions of a multidimensional instrument as statistically fully independent and thus allowing for multiple comparisons will produce distorted results. A compromise can be struck using empirical methods on massive data sets to determine confidence levels that reduce the probability of Type I errors to a specified level, but that possibility does not exist for this study. Thus the controversy over multidimensional instruments will continue.

Fortunately, much of this important debate proved to be irrelevant in the current study. This is because comparison of participant and third-party ratings on most PCI dimensions falls clearly on one side or the other of the line of significance, whether or not that line is drawn making allowance for multiple comparisons. That is, whether significance is set to the $\alpha=0.05$ level (not allowing for multiple comparisons), at the $\alpha=0.0038$ level (allowing for 21 multiple comparisons using the standard formula $\alpha=1-(1-0.05)^{21}$), at the 0.0024 level (allowing for 21 multiple comparisons using $\alpha=0.05/21$), at the $\alpha=0.0054$ level (allowing for 12 multiple comparisons of the major dimensions using the standard formula $\alpha=1-(1-0.05)^{12}$), or at the 0.0042 level (allowing for 12 multiple comparisons using $\alpha=0.05/12$), the correlations we obtained are consistently either significant or not significant for almost all PCI dimensions. This being the case, we decided to proceed by making allowance for 21 multiple comparisons and assessing significance at the $\alpha=0.0038$ level.

Results: Using the PCI to compare introspective recall and narrative analysis

The muting effect of narratives

Our aim was to determine how much information about the phenomenological features of an experience—information available to participants through introspective recall, as reflected in their PCI ratings—was adequately conveyed in the narratives and interview material available to third-party raters. We calculated means (and standard deviations) across the 117 narratives for our 39 participants, the expert rater, and the novice rater. The results are presented in Table 2.

A suggestive way to present these data is a radar graph. The mean for each PCI dimension is plotted on a radius stretching from the center outward to 6 at the circumference. The plot in Figure 1 displays the result for all 117 experiences (happy, ordinary, and RSE), and the plot in Figure 2 does the same for the 39 RSEs. Note that the radius has been lengthened to unpack results close to zero at the center of the plot.

The most obvious feature of these results is that the participant scores tend to be higher than the scores of both raters for most PCI dimensions. This is the case both for all 117 narratives and for the 39 RSE narratives considered separately. This impression is confirmed by the fact that high correlations between participant and rater PCI scores were obtained on many dimensions where the Z test for equality of means produced no significant results. Taking account of the meaning of high versus low PCI ratings, this pattern suggests that the vividness of the phenomenological features of experiences is more pronounced for those who recollect and narrate the experiences and than for those who approach the same experiences through reading narratives. Unless composed by gifted writers and story tellers, it is evident that *narratives are relatively muted ways of introducing readers and hearers to the vivid phenomenological intricacies of personal experience.*

Breaking down the comparison into PCI dimensions

In order to discriminate between the phenomenological features that narratives communicate well and those that narratives convey poorly, we need to allow for the muting effect just described. As discussed earlier, this is accomplished by calculating Pearson correlations between participant PCI ratings and third-party rater PCI scores, because such correlations detect linear relationships regardless of mean differences, and thus register *relative accuracy* of narrative interpretation. For the reasons discussed above, *significance of correlations is assessed at the 0.0038 level.* We note below the very few PCI dimensions for which correlations were significant at the 0.05 level (not allowing for multiple comparisons) but not at the 0.0038 level (allowing for 21 multiple comparisons). The

most notable conclusions to be drawn from these correlation results are as follows. Note that a summary of these results appears in the discussion section, below.

First, certain phenomenological features of experiences communicate quite clearly through narratives, in the sense that their relative strength is registered by both expert and novice interpreters in much the same way. Note that this result is consistent with the “relative muting” result from the previous section. For example, a narrative reader gets the sense that anger is stronger than fear and weaker than sadness in roughly the same degree that the narrator recalls these features of experience, even though all of the impressions are muted for the narrative reader relative to the narrator. The correlation analysis by PCI dimension draws attention to the phenomenological features that have this characteristic. Significance values are reported for all 117 narratives (happy, ordinary, and religious), distinguishing between expert and novice rater comparisons with participant ratings where necessary. We add comments about ways in which results for the 39 RSEs differ from results for the all 117 narratives.

- Affect dimensions, both negative affect ($r=0.63$ for the participant-expert correlation and $r=0.59$ for the participant-novice correlation)—including anger ($p<0.0001$ for both correlations), fear ($p<0.0001$), sadness ($p<0.0001$)—and positive affect ($r=0.72$ and $r=0.63$)—including joy ($p<0.0001$), love ($p<0.0001$), sexual excitement ($p<0.0001$). Even inexperienced narrators seem able to communicate emotion quite well, and both expert and novice interpreter accurately registered the relative strength of affect dimensions.

Interestingly, love was more difficult to convey in RSE narratives ($r=0.37$ for the participant-expert correlation and $r=0.42$ for the participant-novice correlation) than in other types of narratives ($r=0.48$ and $r=0.49$ for all 117 narratives combined), suggesting that particularly complex dimensions of intense meaning attach to religious experiences of love.

- Alterations of awareness ($r=0.48$ for the participant-expert correlation and $r=0.39$ for the participant-novice correlation; $p<0.0001$)—again, inexperienced narrators seem able to convey when their state of consciousness departs from the normal and familiar and heads toward the abnormal and unfamiliar. Our expert interpreter ($r=0.48$) did slightly better than the novice interpreter ($r=0.42$) at picking up on the strength of altered awareness in RSEs.

Second, another set of phenomenological features of experiences proved easier to convey by means of narratives to our expert interpreter than to our novice interpreter. This result is merely suggestive, because thorough investigation would require a population of expert interpreters and another group of novice interpreters as controls.

But the finding is strong enough to invite closer study, and sufficiently important to make researchers cautious about the qualifications of the people they charge with the task of interpreting narratives. It is likely that extensive exposure to RSE narratives and deep familiarity with the way a wide variety of people experience the world, both religiously and more generally, increases awareness and empathy. This in turn helps expert interpreters make better judgments about what kinds of phenomenological characteristics are likely to be present in many kinds of experiences. Thus, narratives of experiences are more resonant for expert interpreters than for novice interpreters, at least in the respects listed below. Presumably, great writers and story tellers make it easier for both experts and novices to understand an experience. For ordinary narrators, by contrast, the expertise and experience of the interpreter matters a great deal. The phenomenological features that have this characteristic correspond to the following PCI dimensions. Again, significance is assessed at the 0.0038 level (“ns” means not significant, “ms” means marginally significant), significance values are reported for all 117 narratives (happy, ordinary, and religious), and we distinguish between expert and novice rater correlations with participant ratings where necessary. We report correlations to conveniently indicate effect size. We add comments about ways in which results for the 39 RSEs differ from results for the all 117 narratives.

- Alterations of experience—changes in experience overall ($r=0.45$ with $p<0.0001$ for the expert-participant correlation versus $r=0.10$ [ns] for the novice-participant correlation), changes in meaningfulness ($r=0.59$ with $p<0.0001$ versus $r=0.25$ with $p=0.0070$ [ms]), and changes in sense of time ($r=0.31$ with $p=0.0006$ versus $r=0.039$ [ns]); *marginally* alterations of body sense ($r=0.22$ with $p=0.019$ versus $r=-0.08$ [ns]); but *not* alterations of perception ($r=0.11$ versus $r=0.03$). Alterations of experience appear to be more difficult for ordinary narrators to explain, giving expert interpreters an advantage in discerning the qualities of the underlying experience. When moving from all 117 experiences to just the 39 RSEs, only the alteration-of-meaning sub-dimension produced a significant correlation, and here again the expert interpreter enjoyed a large advantage over the novice ($r=0.58$ with $p<0.0001$ versus $r=0.13$ [ns]).
- Physical arousal, in the medical rather than sexual sense—marginal for all narratives ($r=0.25$ with $p=0.0056$ for experts versus $r=0.060$ [ns] for novices) and significant for the 39 RSEs ($r=0.47$ with $p=0.0025$ versus $r=0.055$ [ns]).
- Attention in general—marginal for all narratives ($r=0.26$ with $p=0.0045$ [ms] versus $r=0.051$ [ns]) and significant for the 39 RSEs ($r=0.49$ with $p=0.0014$ versus $r=0.22$ [ns])—and the inward directedness of

- Amount of imagery ($r=0.36$ with $p<0.0001$ versus $r=0.11$ [ns] for all experiences; $r=0.51$ with $p<0.0008$ versus $r=0.25$ [ns] for RSEs alone). This did *not* apply to vividness of imagery or to imagery overall, though our expert interpreter achieved a marginally significant result for vividness of imagery in RSEs ($r=0.36$ with $p=0.023$ [ms]).

Third, the remaining dimensions of the PCI correspond to phenomenological features of experience that appear to be difficult to convey in narrative form, or unlikely to come to mind when narrating an experience. The result is that both expert and novice interpreters do a poor job of appreciating the relative importance of these dimensions in the narrator's experience. This holds for all experiences generally and also for RSEs specifically; indeed, in the case of RSEs interpretation seems to be more challenging in regard to most PCI dimensions. These are the danger dimensions for researchers in the scientific study of religion, where narrative approaches seem clearly inadequate.

These results suggest that narrative analysis of RSEs and indeed other complex experiences, *when aiming to identify phenomenological features* of associated states of consciousness, should be approached with caution. Narrative analysis performed by expert reviewers *for this purpose* is likely to be reliable in some respects and misleading in other respects. Thus, the ideal approach to the scientific study of phenomenological features of RSEs will combine a validated tool for identifying phenomenological features of experience (such as the PCI) with expert interpretation of narrative reports of these experiences.

Results: Using the PCI and LIWC to evaluate narrative content

In addition to seeking correlations between participant and third-party rater PCI scores, we attempted to assess the adequacy of narratives at the level of narrative content. The most obvious content question is whether the basic meaning of the three types of narratives—namely, that they convey happy, ordinary, or religious experiences—corresponds to characteristics in the underlying experiences themselves. This question is so elementary that it may seem surprising. But checking that narrative content preserves the basic features of the experiences that participants nominated as happy, ordinary, and religious is an important step in assuring ourselves of the reliability of narratives.

The discussion of methodological complexities above suggests that simply reading the narratives and deciding whether they appropriately report on happy, ordinary, or religious experiences will not do. Instead, we used the PCI to detect significant differences in the phenomenological profile of happy, ordinary, and religious experiences. We

also used the LIWC to analyze the measurable features of narratives in order to determine whether they displayed the sorts of features that the PCI analysis of the underlying recalled experiences suggest that they should. Most LIWC results do not line up conveniently with PCI dimensions but a few do, and in those few cases we have a basic check on the reliability of the entire process, from the protocol that produces narratives of recalled experiences to the adequacy of the narratives themselves.

Happy versus ordinary versus religious experiences

We analyzed differences among the three experience types by performing a one-way analysis of variance (ANOVA) on the 26 PCI dimensions for ordinary versus happy versus religious experiences. Results are displayed in Table 3. All p values are Bonferroni adjusted for multiple group comparisons (i.e. ordinary versus happy versus religious). Significance is assessed at the 0.05 level. The main results are as follows.

- Altered Experience was rated as significantly higher for religious experiences than both happy ($p=.012$) and ordinary ($p<.0001$) experiences; in particular, Altered Meaning was rated as significantly higher for religious experiences than both happy ($p<.0001$) and ordinary ($p<.0005$) experiences. Altered Perception was significantly higher for religious experiences than ordinary ($p=.002$). Altered Awareness was significantly higher for religious experiences than both happy ($p=.004$) and ordinary ($p<.0005$).
- Attention was rated as significantly more inward and/or absorbed for religious experiences than both happy ($p=.009$) and ordinary ($p=.046$) experiences; in particular, attention was significantly more inwardly directed for religious experiences than both happy ($p=.007$) and ordinary ($p=.017$) experiences.
- Imagery content was rated as significantly higher in religious experiences than in both happy ($p<.0005$) and ordinary ($p=.005$) experiences; in particular, the amount of imagery is significantly higher for religious experiences than both happy ($p<.0001$) and ordinary ($p<.0001$).
- Internal Dialogue was rated as significantly higher for religious experiences than happy ($p=.016$) experiences.
- Negative Affect was rated as significantly higher for religious experiences than happy ($p<.0001$) and ordinary ($p=.010$) experiences; this is especially due to Anger being lower for happy experiences than both ordinary ($p=.005$) and religious ($p=.001$). Interestingly, Fear was rated as being significantly higher for religious experience than both happy ($p=.003$) and ordinary ($p=.003$) experiences, and Sadness was rated as being significantly higher for religious experiences than both happy ($p<.0001$) and ordinary ($p=.008$).

- Positive Affect was rated as significantly higher in happy experiences than in ordinary ($p < .0001$) and religious experiences ($p < .0001$); this is especially due to Joy for happy experiences being significantly higher than ordinary ($p < .0001$) and religious ($p < .0001$) experiences.
- Volitional Control was rated as significantly lower for religious experiences than for both happy ($p < .0005$) and ordinary ($p < .0001$).

The graph in Figure 3 displays the PCI dimensions for which religious experiences displayed significant differences (at the 0.05 level) from both happy and ordinary experiences. The error bars indicate 95% confidence levels.

The radar graphs in Figure 4 and Figure 5 display the results, with the scale ranging from 0 to 5 (the theoretical maximum is 6 but there were no average scores higher than 5). PCI dimensions with significant pair-wise mean differences (at the 0.05 level) are marked on the radar graph using the same convention as in Table 3: a–significant difference for H vs O, b–significant difference for H vs R, c–significant difference for O vs R. The first graph displays PCI scores for only the 12 major dimensions for the sake of clarity; the second shows all 26 dimensions.

In passing, we make an important remark about the stability of these findings. It is conceivable that the differences among happy, ordinary, and religious experiences could be due in significant part to differential difficulty in recalling these three types of experiences. Another factor that could conceivably account for the differences in PCI scores among happy, ordinary, and religious experiences, perhaps by differentially affecting ease of recall, is the subjective intensity of the experience. Recall that we timed latency to recall and asked participants to rate the subjective intensity of each experience. Using this information, we ran multivariate analyses of covariance (MANCOVA) for happy versus ordinary versus religious experiences, controlling for both latency and intensity, for just one or the other, and for positive affect. The result of these analyses is that the formal phenomenological features distinguishing RSEs from happy and ordinary experiences remain substantially the same when we control for latency to retrieval, subjective intensity, both latency to retrieval and subjective intensity, and positive affect. These phenomenological signatures are thus robust and offer numerous potential insights into RSEs.

PCI ratings and affect

Our rater analysis showed that narrators communicate all aspects of both positive affect and negative affect to both expert and novice interpreters quite easily and accurately. That being the case, can we use the rater analysis to confirm the reliability of narratives for communicating the features of affective complexity just described? We ran a

one-way ANOVA for happy versus ordinary versus religious experiences, comparing the two major PCI affect dimensions for participants and both raters, using Bonferroni correction for multiple comparisons (across the happy, ordinary, and religious groups). The results are presented in Table 4.

The striking conclusion to be drawn from these data is that both expert and novice raters derived from narratives the same picture of affective complexity in RSEs that the PCI analysis of recalled experiences yielded—and they did so with high resolution given the F values involved. Specifically, for participants recalling experiences and for expert and novice raters interpreting narratives of those experiences, we see that:

- RSEs are affectively more positive than negative overall,
- RSEs contain more positive affect than ordinary experiences and less positive affect than happy experiences; and
- RSEs contain more negative affect than both ordinary experiences and happy experiences.

In the case of affect dimensions, then, we are entitled to have high confidence in the power of narratives—even RSE narratives produced by ordinary people without specific prompting for elaboration—to convey the affective qualities of the underlying experiences.

PCI phenomenological signatures and LIWC analysis of narratives

Our evaluation of narrative reliability so far has depended on blinded raters scoring PCI categories. We turned to the LIWC (2001 version) for an alternative approach. As noted earlier, the aim is to locate measurable linguistic features of narratives that can be meaningfully compared to what narrators report about their recalled experiences, all without having to rely on third-party interpreters. One aspect of the LIWC supports precisely this type of connection between narratives and underlying experiences: negative and positive affect. The LIWC 2001 counts words related to positive emotions and negative emotions using its in-built dictionary. We ran a one-way ANOVA on LIWC counts of positive and negative emotion words for happy versus ordinary versus religious narratives, using Bonferroni correction for multiple comparisons (across the three groups). The results are presented in Table 5.

Unfortunately, these results are not especially useful. LIWC positive emotion words are significantly more frequent for happy than for ordinary ($p=0.011$) or religious ($p<0.0005$) experiences, as expected, but the distinction between ordinary and religious experiences is not significant so we can infer nothing safely about that comparison. The means for LIWC negative emotion words fit the pattern we saw above, with religious experiences showing more negative emotion words than either happy or ordinary experiences, but the result is not significant overall and

there are no significant pair-wise comparisons, either. The most we can conclude from the LIWC results, therefore, is that they are suggestive. With better dictionaries and categorizations, it may become possible to conduct more effective narrative analysis without third-party interpreters.

Discussion

The major goal of this study was to evaluate the adequacy of RSE narratives for conveying the phenomenological features of the RSEs underlying those narratives. This question is vital because the scientific study of RSEs has relied heavily on RSE narratives.

Reliability of narratives using the PCI and blinded raters

We used a variety of methods in order to establish meaningful ways to test narrative reliability. The most important and useful tool was the PCI, which is a retrospective phenomenological analysis instrument applied to recalled experiences. We acknowledged complexities surrounding the recall protocol used in this study, arguing that these complexities are inevitable, and described the procedures we introduced to mitigate any problems that may arise. We also used expert and novice raters to score narratives on the 26 dimensions of the PCI, and we submitted narratives to the LIWC 2001 linguistic analysis software. This combination of approaches furnished a sound basis for evaluating the reliability of narratives of recalled experiences in general, and of RSE narratives in particular.

We conclude that narratives of RSEs produced by ordinary people (as against gifted writers and storytellers), and without specific prompting to elaborate beyond basic probe questions as described above, are reliable means of accessing phenomenological features of the underlying RSEs in some respects, reasonably reliable in certain other respects when read by expert interpreters but not by novice interpreters, and simply unreliable in yet other respects. The breakdown in terms of the 26 major and minor PCI dimensions is as follows.

- PCI dimensions for which narratives communicate phenomenological features of experiences clearly to both expert and novice interpreters: *negative affect* (including anger, fear, sadness sub-dimensions), *positive affect* (including joy, love, sexual excitement sub-dimensions), and *altered awareness*.
- PCI dimensions for which narratives communicate phenomenological features of experiences clearly to expert interpreters but *not* to novice interpreters: *alterations of experience* (meaningfulness, time sense, and marginally also body sense sub-dimensions), *physical arousal*, *attention* (and inward directedness sub-dimension), and *amount of imagery*.

- PCI dimensions for which narratives communicate phenomenological features of experiences clearly *neither* to expert interpreters *nor* to novice interpreters: *perception* sub-dimension of alterations of experience (marginally also body sense sub-dimension), *absorption* sub-dimension of attention, *imagery* (especially vividness of imagery sub-dimensions), *internal dialogue*, *memory*, *self-awareness*, *rationality*, and *volitional control*.

This conclusion offers valuable guidance to researchers who depend on narratives to access information about the phenomenological characteristics of RSEs. It should be supplemented with another important conclusion, namely, that the vividness of experiences tends to be muted across the board when they are narrated, relative to the way they are recalled.

The phenomenological profile of RSEs and extant knowledge about RSEs

Our study yielded a phenomenological profile of RSEs. Specifically, the PCI's phenomenological signature for RSEs differs significantly from the profiles for baseline normal and happy experiences on the following PCI major dimensions: altered experience, altered awareness, attention, imagery, internal dialogue, negative affect, positive affect, and volitional control. This result allows us to evaluate the connection between the phenomenological content of recalled experiences (which is what the PCI profiles register) and the semantic content of narratives based on those experiences (which is expressed in the most basic way in the three-way classification of narratives as happy versus ordinary versus religious). The key to evaluation is whether the phenomenological signatures for the three classes of experiences differ *in theoretically intelligible ways* relative to the existing phenomenological literature. We consider the intelligibility of these dimensions of difference in order.

First, alterations of experience—particularly changes in meaningfulness, perception, and time—and altered awareness—feeling that one's conscious awareness has slipped into a strange or unfamiliar state—are classic markers of RSEs. Hardy (1979) documents numerous ways in which such changes in experience and awareness occur within RSEs and our own collection of narratives is replete with examples. As we saw in our blinded rater analysis, however, our narratives most clearly indicate altered awareness in RSEs. Expert interpreters do better than novices in discerning the presence within RSEs of alterations in meaning and time flow, and it is genuinely difficult discerning when RSEs involve alterations in perception and body sense, at least when judgment is based on narratives produced by ordinary people rather than by talented writers gifted in introspection or prompted for

elaboration. When expert interpreters are in use, however, the PCI yields a phenomenological signature for recalled RSEs that closely matches the features discernible in RSE narratives.

Second, the PCI signatures also suggest that RSEs dominate attention, particularly in the sense of catapulting it outwards, beyond the self, to some transcendent Other. Again, this is a classic feature of many types of RSEs, richly documented in the phenomenological literature, and it is a feature that experts (though not novices) readily detect in RSE narratives. Here again we have a close match between the phenomenological signature for recalled RSEs and long-documented features of RSE narratives.

Third, the most spectacular contrast between RSEs and happy or ordinary experiences is in the PCI sub-dimension measuring amount of imagery. The phenomenological literature likewise suggests that RSEs are often loaded with imagery, and expert interpreters appear to detect this feature of RSEs from narratives quite easily. Thus, we have another match between PCI signature and existing knowledge of RSEs based on analysis of narratives.

Fourth, the PCI signature suggests that internal dialogue markedly increases and volitional control markedly drops relative to happy and ordinary experiences. The internal dialogue result is slightly surprising, as the phenomenological literature does not stress this feature of RSEs except in a few types of RSEs such as supernatural encounter experiences and internalized moral debates. By contrast, the drop in volitional control is a well-known feature of RSEs: the experiences are often so powerful that those undergoing them feel less able than usual to make choices about how to regard what is happening to them or about how to respond. Our blinded rater analysis showed that our narratives were difficult to interpret accurately in regard to both internal dialogue and volitional control. We can conclude, however, that at least the volitional control aspect is intelligible against the wider phenomenological literature on RSEs.

Fifth and finally, the PCI signature offers a rich portrayal of affect within RSEs. Religious experiences appear to be affectively complex. They are more positive than negative overall, with positive affective markedly less than for happy experiences and more than for ordinary experiences. But negative affect is decidedly stronger in RSEs than in both happy and ordinary experiences, meaning that RSEs are both affectively positive and negative in a rather complex way. While the phenomenological tradition in the study of RSEs does occasionally present accounts of almost purely affectively positive and almost purely affectively negative experiences, most RSEs appear to be affectively complex in just the way. A person might simultaneously feel fear and love in the presence of a powerful

spiritual being, or a conversion experience will trigger both sadness and joy, intensely intermingled. Thus, the PCI signature makes good sense against the background of existing knowledge of RSEs.

We conclude that the PCI-derived phenomenological signatures for RSEs versus happy and ordinary experiences make good sense against the long-run of narrative-guided interpretation of the phenomenological features of RSEs. This result should increase our confidence in the reliability of narratives, in respect of their power to convey the basic phenomenological features that distinguish RSEs from other types of experience.

Reliability of narratives on affect dimensions

Because narratives proved particularly reliable at communicating affect, the PCI's affect dimensions permitted a more precise kind of reliability testing. Our investigation of affect dimensions of RSEs showed that participants as well as both expert and novice raters rank affect in the same way. Specifically, (i) RSEs are more positive than negative in overall affect, (ii) positive affect in RSEs lies somewhere between happy and ordinary experiences, and (iii) negative affect in RSEs is significantly greater in RSEs than in either happy or ordinary experiences. The agreement between participant PCI scoring of recalled experiences and blinded rater scoring of narrated reports of the same experiences suggests excellent reliability of narratives in regard to communicating affect.

We attempted to confirm this analysis of affect in RSEs using the LIWC 2001 to perform linguistic analysis of narratives on positive and negative emotion words. Though there was rough agreement at the level of means, the LIWC did not produce results of sufficiently high significance to ramify our conclusions. Linguistic analysis requires greater sophistication in dictionaries in order to achieve the sensitivity necessary to play a role in evaluating narrative reliability.

The value of the PCI

The PCI is a valuable tool for generating quantitative profiles of states of consciousness. This approach is superior to blinded rater analysis of narratives alone—*for the particular purpose of identifying distinctive phenomenological features of states of consciousness*. For other purposes, narrative analysis may be quite serviceable; indeed, narrative analysis remains indispensable for identifying the specific cognitive content of states of consciousness. We believe the PCI instrument can be tailored to the specific application of studying RSEs but even in its original form it proved to be a sensitive and robust tool.

The PCI also permits finer discriminations among types of RSEs, such as the major types discriminated within the Hardy classification (Hardy 1979), though we do not report on that here. Because the Hardy subtypes distinguish

neurologically quite different types of experiences (e.g. dominantly cognitive-emotional versus dominantly perceptual), the PCI has the potential to constrain cognitive models of the production of RSEs.

Conclusion

The longstanding tradition within the scientific study of religion has been to detect the phenomenological features of RSEs by means of analyzing narratives of RSEs solicited from those who have such experiences and are willing to recall and narrate them. Our study shows that this traditional reliance on narratives is sound in some respects but questionable in other respects. Though narratives remain indispensable for interpreting RSEs, future research into the phenomenology of RSEs should proceed with clear awareness that narratives can be misleading as well as illuminating. Researchers should ensure that their reliance on narratives for detecting phenomenological features of RSEs should emphasize the respects in which narratives tend to be reliable and supplement interpretation with other resources (such as PCI results) for the respects in which narratives tend to be unreliable.

Acknowledgements

We are grateful to Ms. Erica Harris, MPH and Mr. Paul Butler who tested subjects, to Mr. Ephraim Josephs and Mr. Derek Michaud who rated narratives, and to Dr. Deborah Kelemen and Dr. Catherine Caldwell-Harris for discussions and advice. We acknowledge funding support from the Center for Theology and the Natural Sciences and from Boston University School of Theology.

References

- Alschuler, A. S. (1993). Inner teachers and transcendent education. In K. R. Rao (ed.), *Cultivating consciousness: Enhancing human potential, wellness, and healing* (pp. 181-93). Westport, CT: Praeger Press.
- Andresen, Jensine (ed.). (2001). *Religion in mind: Cognitive perspectives on religious belief, ritual, and experience* (Cambridge: Cambridge University Press).
- Angelini, F. J., Kumar, V. R., & Chandler, L. (1999). The Harvard Group of Scale Hypnotic Susceptibility and related instruments: Individual and group administrations. *International Journal of Clinical and Experimental Hypnosis*, 47(3), 236-50.
- Antony M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales (DASS) in clinical groups and a community sample. *Psychological Assessment*, 10, 176-81.
- Atran, S. (2002). *In Gods we trust: The evolutionary landscape of religion*. Oxford: Oxford University Press.

- Bagger, M. C. (1999). *Religious experience, justification, and history*. Cambridge: Cambridge University Press.
- Boyer, P. (1999). Human cognition and evolution. In H. Moore (ed.), *Anthropological theory today* (206-33). Cambridge: Polity Press.
- Boyer, P. (2000). Evolutionary psychology and cultural transmission. *American Behavioral Scientist*, 43(6), 987-1000.
- Burhenn, H. (1995). Philosophy and religious experience. In R. W. Hood (ed.), *Handbook of religious experience* (pp. 144-59). Birmingham, AL: Religious Education Press.
- Campbell, R. S., Pennebaker, J. W. (2003). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science*, 14, 60-65.
- Crawford, J. R., & Henry, J. D. (2003). The Depression Anxiety Stress Scales (DASS): Normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology*, 42, 111-31.
- Currie, R., Klug, L. F., & McCombs, C. R. (1982). Intimacy and saliency: Dimensions for ordering religious experiences. *Review of Religious Research*, 24(1), 19-32.
- Donovan, P. (1979). *Interpreting religious experience*. New York: Seabury Press.
- Fetzer Institute, (1999). *Multidimensional measurement of religiousness/spirituality for use in health research*. Bethesda, MD: Fetzer Institute, National Institute on Aging Working Group.
- Forbes, E. J., & Pekala, R. J. (1993). Predicting hypnotic susceptibility via a phenomenological approach. *Psychological Reports*, 73, 1251-56.
- Forbes, E. J., & Pekala, R. J. (1996) Types of hypnotically (un)susceptible individuals as a function of phenomenological experience: A partial replication. *Australian Journal of Clinical and Experimental Hypnosis*, 2, 92-109.
- George, L. (1995). *Alternative realities: The paranormal, the mystic, and the transcendent in human experience*. New York: Facts on File.
- Gottschalk, L. A., Gleser, G. C. (1969). *The measurement of psychological states through the content analysis of verbal behavior*. Berkeley: University of California Press.
- Grant, J. E. (2004). Dissociative symptoms in kleptomania. *Psychological Reports*, 94(1), 77-82.
- Hallman, R. J. (1963). The necessary and sufficient conditions of creativity. *Journal of Humanistic Psychology*, 3(1), 14-27.

- Hand, J., Pekala, Ronald J., & Kumar, V. K. (1995). Prediction of Harvard and Stanford Scale scores with a phenomenological instrument. *Australian Journal of Clinical and Experimental Hypnosis*, 23(2), 124-34.
- Hardy, A. C. (1979). *The spiritual nature of man: A study of contemporary religious experience*. Oxford and New York: Oxford University Press.
- Hill, P. C., & Hood, Ralph W., Jr., (eds.). (1999). *Measures of religiosity*. Birmingham, AL: Religious Education Press.
- Huang, M. P., Himle, J., & Alessi, N. E. (2000). Vivid visualization in the experience of phobia in virtual environments: Preliminary results. *CyberPsychology and Behavior*, 3(3), 315-20.
- Hutchinson-Phillips, S., Jamieson, G. A., & Gow, K. (2005). Differing roles of imagination and hypnosis in self regulation of eating behavior. *Contemporary Hypnosis*, 22(4), 171-83.
- James, W. (1902). *The varieties of religious experience: A study in human nature*. New York: Longmans, Green, & Company.
- Johanson, M., Valli, K., Revonsuo, A., Chaplin, J. E., & Wedlund, J. E. (2008). Alterations in the contents of consciousness in partial epileptic seizures. *Epilepsy and Behavior*, 13(2), 366-71.
- Kason, Y. (1997). *A farther shore: How near-death and other extraordinary experiences can change ordinary lives*, 2nd ed. New York: HarperCollins.
- Katz, S. (1978). *Mysticism and philosophical analysis*. New York: Oxford University Press.
- Kokoszka, A. (1992). Occurrence of altered states of consciousness among students: Profoundly and superficially altered states in wakefulness. *Imagination, Cognition, and Personality*, 12, 231-47.
- Kumar, V. K., & Pekala, R. J. (1988). Hypnotizability, absorption, and individual differences in phenomenological experience. *International Journal of Clinical and Experimental Hypnosis*, 36(2), 80-8.
- Kumar, V. K., & Pekala, R. J. (1989). Variations in phenomenological experience as a function of hypnosis and hypnotic susceptibility: A replication. *British Journal of Experimental and Clinical Hypnosis*, 6, 17-22.
- Kumar, V. K., Marcano, G., & Pekala, R. J. (1996a). Behavioral and subjective scoring of the Harvard Group Scale of Hypnotic Susceptibility: Further data and an extension. *American Journal of Clinical Hypnosis*, 38(3), 191-9.
- Kumar, V. K., Pekala, R. J., & Cummings, J. (1996b). Trait factors, state effects, and hypnotizability. *International Journal of Clinical and Experimental Hypnosis*, 44(3), 232-49.

- Lepore, S. J., Smyth, J. M. (2002). *The writing cure: How expressive writing promotes health and emotional well-being*. Washington, DC: American Psychological Association.
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales*. Sydney, Australia: The Psychology Foundation of Australia.
- Maitz, E. A., & Pekala, R. J. (1991). Phenomenological quantification of an out-of-the-body experience associated with a near-death event. *Omega*, 22, 199-214.
- Manmiller, J. L., Kumar, V. K., & Pekala, R. J. (2005). Hypnotizability, creative capacity, creativity styles, absorption, and phenomenological experience during hypnosis. *Creativity Research Journal*, 17(1), 9-24.
- Maurer, R. L., Sr., Kumar, V. K., Woodside, L., & Pekala, R. J. (1997). Phenomenological experience in response to monotonous drumming and hypnotizability. *American Journal of Clinical Hypnosis*, 40(2), 130-45.
- McCloskey, M. S., Kumar, V. K., & Pekala, R. J. (1999). State and trait depression, physical and social anhedonia, hypnotizability and subjective experiences during hypnosis. *American Journal of clinical Hypnosis*, 41(3), 231-52.
- McNamara, P. (2009). *The neuroscience of religious experience*. Cambridge and New York: Cambridge University Press.
- McNamara, P., (ed.). (2006). *Where God and science meet: How brain and evolutionary studies alter our understanding of religion, 3 vols. Vol. 1: Evolution, genes, and the religious brain; Vol. 2: The neurology of religious experience; Vol. 3: The psychology of religious experience*. Westport, CT: Praeger Press.
- Moehle, D. (1983). Cognitive dimensions of religious experiences. *Journal of Experimental Social Psychology*, 19(2), 122-45.
- Murphy, M. (1992). *The future of the body: Explorations into the further evolution of human nature*. New York: Putnam.
- Otto, R. (1923). *The Idea of the Holy: An Inquiry into the non-rational factor in the idea of the divine and its relation to the rational*. London and New York: Oxford University Press; original work published 1917.
- Oxman T. E., Rosenberg, S. D., Schnurr, P. P., Tucker, G. J., & Gala, G. (1998). The language of altered states. *Journal of Nervous & Mental Disease*, 176, 401-8.
- Pekala, R. J. (1991). *Quantifying consciousness: An empirical approach*. New York and London: Plenum Press.
- Pekala, R. J. (1995a). A short unobtrusive hypnotic induction for assessing hypnotizability level: 1. Development

- and research. *American Journal of Clinical Hypnosis*, 37, 271-83.
- Pekala, R. J. (1995b). A short unobtrusive hypnotic induction for assessing hypnotizability II: Clinical case reports. *American Journal of Clinical Hypnosis*, 37, 284-93.
- Pekala, R. J. (2002). Operationalizing "trance" II. Clinical application using a psychophenomenological approach. *American Journal of Clinical Hypnosis*, 44, 241-55.
- Pekala, R. J., & Forbes, E. J. (1997). Types of hypnotically (un)susceptible individuals as a function of phenomenological experience: Towards a typology of hypnotic types. *American Journal of Clinical Hypnosis*, 39(3), 212-24.
- Pekala, R. J., & Kumar, V. K. (1984). Predicting hypnotic susceptibility by a self-report phenomenological state instrument. *American Journal of Clinical Hypnosis*, 27, 114-21.
- Pekala, R. J., & Kumar, V. K. (1987). Predicting hypnotic susceptibility by a self-report phenomenological state instrument. *American Journal of Clinical Hypnosis*, 30, 57-65.
- Pekala, R. J., & Kumar, V. K. (1988). Phenomenological variations in attention across low, medium, and high hypnotically susceptible individuals. *Imagination, Cognition, and Personality*, 7(4), 303-314.
- Pekala, R. J., & Kumar, V. K. (1989). Phenomenological patterns of consciousness during hypnosis: Relevance to cognition and individual differences. *Australian Journal of Clinical and Experimental Hypnosis*, 17, 1-20.
- Pekala, R. J., & Nagler, R. (1989). The assessment of hypnoidal states: Rationale and clinical application. *American Journal of Clinical Hypnosis*, 31, 231-36.
- Pekala, R. J., Forbes, E. J., & Contrisciani, P. A. (1989). Assessing the phenomenological effects of several stress management strategies. *Imagination, Cognition, and Personality*, 8, 265-81.
- Pekala, R. J., Kumar, V. K., Maurer, R., Elliott-Carter, N. C., & Moon, E. (2006). "How deeply hypnotized did I get?" Predicting self-reported hypnotic depth from a phenomenological assessment instrument. *International Journal of Clinical and Experimental Hypnosis*, 54(3), 316-39.
- Pekala, R. J., Steinberg, J., & Kumar, V. K. (1986). Measurement of phenomenological experience: Phenomenology of consciousness inventory. *Perceptual and Motor Skills*, 63, 983-89.
- Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, 8, 162-166.
- Pennebaker, J. W., Mayne, T., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of*

Personality and Social Psychology, 72, 863-71.

Proudfoot, W. (1985). *Religious experience*. Berkeley, CA: University of California Press.

Revonsuo, A. (2006). *Inner presence: Consciousness as a biological phenomenon*. Cambridge, MA: The MIT Press.

Rock, A. J., Wilson, J. M., Johnston, L. J., & Levesque, J. V. (2008). Ego boundaries, shamanic-like techniques, and subjective experience: An experimental study. *Anthropology of Consciousness*, 19(1), 60-83.

Rosegrant, J. (1976). The impact of set and setting on religious experience in nature. *Journal for the Scientific Study of Religion*, 15, 301-10.

Rosenberg, S. D., & Tucker, G. J. (1978). Verbal behavior and schizophrenia: The semantic dimension. *Archives of General Psychiatry*, 36, 1331-7.

Schleiermacher, F. D. E. (1893). *On religion: Speeches to its cultured despisers*. London: K. Paul, Trench, Trübner; original work published 1799.

Spinhoven, P., Labbe, M. R., & Rombouts, R. (1993). Feasibility of computerized psychological testing with psychiatric outpatients. *Journal of Clinical Psychology*, 49, 440-7.

Stace, W. T. (1960). *Mysticism and philosophy*. Philadelphia: J. B. Lippincott.

Stiles, W. B. (1992). *Describing talk: A taxonomy of verbal response modes*. Newbury Park, CA: Sage.

Szabó, C. (1993). The phenomenology of the experiences and the depth of hypnosis: Comparison of direct and indirect induction techniques. *International Journal of Clinical and Experimental Hypnosis*, 41(3), 225-33.

Tart, C. T. (1975). *States of Consciousness*. New York: Dutton.

Vanderlinden, J., Spinhoven, P., Vandereycken, W., & van Dyck, R. (1995). Dissociative and hypnotic experiences in eating disorder patients: An exploratory study. *American Journal of Clinical Hypnosis*, 38(2), 97-108.

Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, 3(4), 330-49.

Lutz, A., & Thompson, E. (2003) Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies*, 10(9-10), special issue on "Trusting the Subject," 31-52.

Varga, K., Józsa, E., Bányai, É. I., & Gösi-Greguss, A. C. (2001). Phenomenological experiences associated with hypnotic suggestibility. *International Journal of Clinical and Experimental Hypnosis*, 49(1), 19-29.

Venkatesh, S., Raju, T. R., Shivani, Y., Tompkins, G., & Meti, B. L. (1997). A study of structure of phenomenology

- of consciousness in meditative and non-meditative states. *Indian Journal of Physiology and Pharmacology*, 41(2), 149-53.
- White, C. (2008). A measured faith: Edwin Starbuck, William James, and the scientific reform of religious experience. *Harvard Theological Review*, 101(3-4), 431-50.
- Wildman, W. J. (2002). Consciousness expanded. In S. M. A. Sinha, & B. V. Sreekantan (eds.), *Science and metaphysics: a discussion on consciousness and genetics* (pp. 125-41). Bangalore: National Institute of Advanced Studies.
- Wildman, W. J., & Brothers, L. A. (1999). A neuropsychological-semiotic model of religious experiences. In R. J. Russell, N. Murphy, T. C. Meyering, & M. A. Arbib (eds.), *Neuroscience and the person: scientific perspectives on divine action* (pp. 347-416). Vatican City State, Vatican Observatory, Berkeley, CA: Center for Theology and the Natural Sciences.
- Wildman, W. J., & McNamara, P. (2008). Challenges facing the neurological study of religious belief, behavior and experience. *Method and theory in the study of religion*, 20(3), 212-42.
- Yamane, D. (2000). Narrative and religious experience. *Sociology of religion*, 61(2), 171-89.
- Zaehner, R. C. (1957). *Mysticism, sacred and profane: An inquiry into some varieties of praeter-natural experience*. Oxford: Clarendon Press.

Tables and figures (see captions below for the order in which they appear in the paper)

Table 1. The 26 dimensions of the Phenomenology of Consciousness Inventory (PCI), grouped into 12 major dimensions (major dimensions in italics)

PCI Dimensions	Meaning of dimension on left (0) ... right (6) poles of scale
<i>Altered Experience (AE)</i>	<i>Average of:</i>
>Body Image (Abody)	Bodily feelings: within skin...expand into world
>Meaning (Amean)	Insight, awe, reverence: none...strong
>Perception (Aperc)	World's appearance: no difference... strong difference
>Time Sense (Atime)	Flow of time: no change...strong change
<i>Altered Awareness (ALA)</i>	<i>State of awareness: normal... strikingly different</i>
<i>Physical Arousal (AR)</i>	<i>Muscular tightness, tension: low...high</i>
<i>Attention (AT)</i>	<i>Average of:</i>
>Inward Direction (Direction)	Attention: directed to external world... to internal experience
>Absorption (Absorption)	Attention: distracted, low concentration...completely absorbed
<i>Imagery (IM)</i>	<i>Average of:</i>
>Amount (Amount)	Amount of visual imagery: low...high
>Vividness (Vivid)	Imagery: vague, dim, diffuse, normal...vivid like real-world objects
<i>Internal Dialogue (IND)</i>	<i>Silent talking, internal dialogue: none...strong</i>
<i>Memory (M)</i>	<i>Memory: blurred, hazy, vacant... sharp, distinct, complete</i>
<i>Negative Affect (NA)</i>	<i>Average of:</i>
>Anger (Anger)	Anger, rage, upset: none...strong
>Fear (Fear)	Fear, terror: none...strong
>Sadness (Sad)	Sadness, unhappiness, dejection: none...strong
<i>Positive Affect (PA)</i>	<i>Average of:</i>
>Joy (Joy)	Ecstasy, extreme happiness, joy: none... strong
>Sexual Excitement (SE)	Sexual feelings: none...strong
>Love (Love)	Feelings of love, living-kindness: none...strong
<i>Rationality (R)</i>	<i>Thinking: unclear, obscure, irrational...clear, distinct, rational</i>
<i>Self-Awareness (SA)</i>	<i>Self-awareness, self-consciousness: low...high</i>
<i>Volitional Control (VC)</i>	<i>Control over thoughts, attention: none...high</i>

Table 2. Comparison of participant PCI ratings based on introspective recall versus two independent rater's narrative-based PCI scores.

PCI Dimensions 117 Narratives	Participants Mean (StdDev)	Rater1 (expert) Mean (StdDev)	Rater2 (novice) Mean (StdDev)
AE	1.859 (1.390)	0.857 (0.908)	1.825 (1.254)
AE_Abody	1.892 (1.729)	0.171 (0.931)	1.376 (1.607)
AE_Amean	2.224 (1.772)	2.470 (1.976)	1.462 (1.822)
AE_Aperc	1.234 (1.411)	0.350 (1.309)	2.094 (1.961)
AE_Atime	1.997 (1.854)	0.436 (1.289)	2.368 (1.937)
ALA	1.892 (1.857)	1.761 (2.242)	1.103 (1.783)
AR	1.786 (1.536)	1.120 (1.543)	2.521 (1.442)
AT	3.417 (1.172)	3.376 (1.275)	3.436 (1.274)
AT_Absorption	4.333 (1.310)	4.051 (1.231)	4.120 (1.353)
AT_Direction	2.806 (1.584)	2.769 (1.699)	2.752 (1.838)
IM	2.015 (1.279)	0.671 (1.198)	2.321 (1.231)
IM_Amount	2.957 (2.941)	0.838 (1.364)	2.043 (1.522)
IM_Vivid	2.829 (1.662)	0.504 (1.222)	2.598 (1.204)
IND	2.521 (1.926)	0.846 (1.535)	1.726 (2.168)
M	4.823 (1.065)	3.906 (1.196)	5.641 (0.895)
NA	1.040 (1.239)	1.043 (1.410)	0.587 (1.038)
NA_Anger	1.047 (1.527)	1.077 (1.662)	0.573 (1.227)
NA_Fear	0.833 (1.413)	0.983 (1.514)	0.470 (1.208)
NA_Sad	1.192 (1.582)	1.068 (1.828)	0.718 (1.490)
PA	2.444 (1.409)	1.755 (1.224)	1.313 (1.287)
PA_Joy	3.256 (1.982)	3.128 (1.989)	2.026 (2.139)
PA_Love	3.321 (1.973)	1.974 (1.927)	1.821 (2.024)
PA_SE	0.765 (1.401)	0.162 (0.742)	0.094 (0.525)
R	4.801 (1.306)	4.556 (0.969)	5.624 (0.728)
SA	4.598 (1.141)	4.103 (1.094)	3.615 (1.644)
VC	4.231 (1.395)	4.188 (1.152)	5.692 (0.748)

Table Notes: Abbreviations of PCI dimensions are defined in Table 1.

Table 3. PCI means (with standard deviations) for happy versus ordinary versus religious experiences.

PCI Dimension	Happy	Ordinary	Religious	Overall Significance
AE	1.76 (1.26)	1.20 (1.01)	2.61 (1.50)	p<.0005 ^{b c *}
AE_Abody	2.08 (1.73)	1.36 (1.45)	2.24 (1.89)	ns (p=.056)
AE_Amean	1.96 (1.36)	1.17 (1.39)	3.54 (1.68)	p<.0005 ^{b c *}
AE_Aperc	1.14 (1.32)	.735 (.986)	1.82 (1.66)	p=.002 ^{c *}
AE_Atime	1.89 (1.75)	1.56 (1.71)	2.54 (2.00)	ns (p=.060)
ALA	1.72 (1.59)	.991 (1.46)	2.97 (1.96)	p<.0005 ^{b c *}
AR	1.51 (1.43)	1.67 (1.52)	2.18 (1.60)	ns (p=.133)
AT	3.11 (1.04)	3.26 (1.19)	3.89 (1.16)	p=.007 ^{b c *}
AT_Absorption	4.03 (1.41)	4.37 (1.42)	4.60 (1.03)	ns (p=.147)
AT_Direction	2.42 (1.53)	2.51 (1.58)	3.49 (1.44)	p=.004 ^{b c *}
IM	2.46 (1.33)	2.23 (1.26)	1.36 (.970)	p<.0005 ^{b c *}
IM_Amount	1.74 (1.49)	1.51 (1.64)	5.62 (3.26)	p<.0005 ^{b c *}
IM_Vivid	3.17 (1.58)	2.91 (1.45)	2.41 (1.88)	ns (p=.142)
IND	2.00 (2.06)	2.36 (1.90)	3.21 (1.64)	p=.017 ^{b *}
M	4.87 (1.12)	4.85 (1.01)	4.75 (1.08)	ns (p=.874)
NA	.461 (.703)	.945 (1.22)	1.71 (1.38)	p<.0005 ^{b c *}
NA_Anger	.282 (.605)	1.32 (1.71)	1.53 (1.70)	p<.0005 ^{a b *}
NA_Fear	.487 (.885)	.487 (.970)	1.53 (1.91)	p=.001 ^{b c *}
NA_Sad	.512 (.949)	1.03 (1.41)	20.4 (1.88)	p<.0005 ^{b c *}
PA	3.49 (1.03)	1.61 (1.18)	2.24 (1.32)	p<.0005 ^{a b *}
PA_Joy	4.81 (1.03)	2.24 (1.61)	2.72 (2.13)	p<.0005 ^{a b *}
PA_Love	4.38 (1.52)	2.06 (1.77)	3.51 (1.90)	p<.0005 ^{a c *}
PA_SE	1.29 (1.76)	.526 (1.01)	.474 (1.20)	p=.014 ^{a b *}
R	5.06 (1.05)	4.88 (1.41)	4.46 (1.39)	ns (p=.115)
SA	4.63 (.901)	4.71 (1.21)	4.45 (1.29)	ns (p=.599)
VC	4.41 (1.06)	4.96 (1.22)	3.32 (1.38)	p<.0005 ^{b c *}

Table Notes: Abbreviations of PCI dimensions are defined in Table 1. Pair-wise significance results are expressed using superscripts, as follows: a–significant difference between happy and ordinary; b–significant difference between happy and religious; c–significant difference between ordinary and religious; ns–not significant at the 0.05 level. Significance is assessed at the 0.05 level.

Table 4. One-way ANOVA showing means (with standard deviations) for Happy versus Ordinary versus Religious experiences on major PCI affect dimensions.

PCI Dimension	Happy	Ordinary	Religious	F	Overall Sig.
Participant_NA	0.461 (0.703)	0.945 (1.217)	1.714 (1.376)	12.063	p<0.0005 ^{b c}
Participant_PA	3.487 (1.028)	1.611 (1.179)	2.235 (1.320)	25.485	p<0.0005 ^{a b}
Expert_NA	0.479 (0.752)	0.726 (1.081)	1.923 (1.765)	14.396	p<0.0005 ^{b c}
Expert_PA	2.718 (0.786)	0.966 (0.945)	1.581 (1.201)	31.301	p<0.0005 ^{a b c}
Novice_NA	0.154 (0.332)	0.470 (0.790)	1.085 (1.344)	10.340	p<0.0005 ^{b c}
Novice_PA	2.120 (1.407)	0.786 (0.978)	1.103 (1.185)	13.086	p<0.0005 ^{a b}

Table Notes: PA=positive affect, NA=negative affect. Table shows participant scores as well as expert and novice blinded rater scores. Overall significance results are Bonferroni corrected for multiple comparisons. Pair-wise significance results are expressed using superscripts, as follows: a–significant difference between happy and ordinary; b–significant difference between happy and religious; c–significant difference between ordinary and religious. Significance is assessed at the 0.05 level.

Table 5. One-way ANOVA showing means (with standard deviations) for Happy versus Ordinary versus Religious experiences on LIWC 2001 emotion word.

LIWC Category	Happy	Ordinary	Religious	F	Overall Sig
Positive Emotion	3.085 (2.070)	1.979 (1.674)	1.618 (1.030)	8.390	p<0.0005 ^{a b}
Negative Emotion	0.649 (0.897)	0.951 (0.982)	1.082 (1.101)	1.933	p=0.149

Table Notes: Counts are words in category as percentage of words in entire narrative. Overall significance results are Bonferroni corrected for multiple comparisons. Pair-wise significance results are expressed using superscripts, as follows: a–significant difference between happy and ordinary; b–significant difference between happy and religious; c–significant difference between ordinary and religious. Significance is assessed at the 0.05 level

Figure 1. Radar plot of PCI ratings of 117 experiences (happy, ordinary, and religious) for participants and two raters. Abbreviations of PCI dimensions are defined in Table 1.

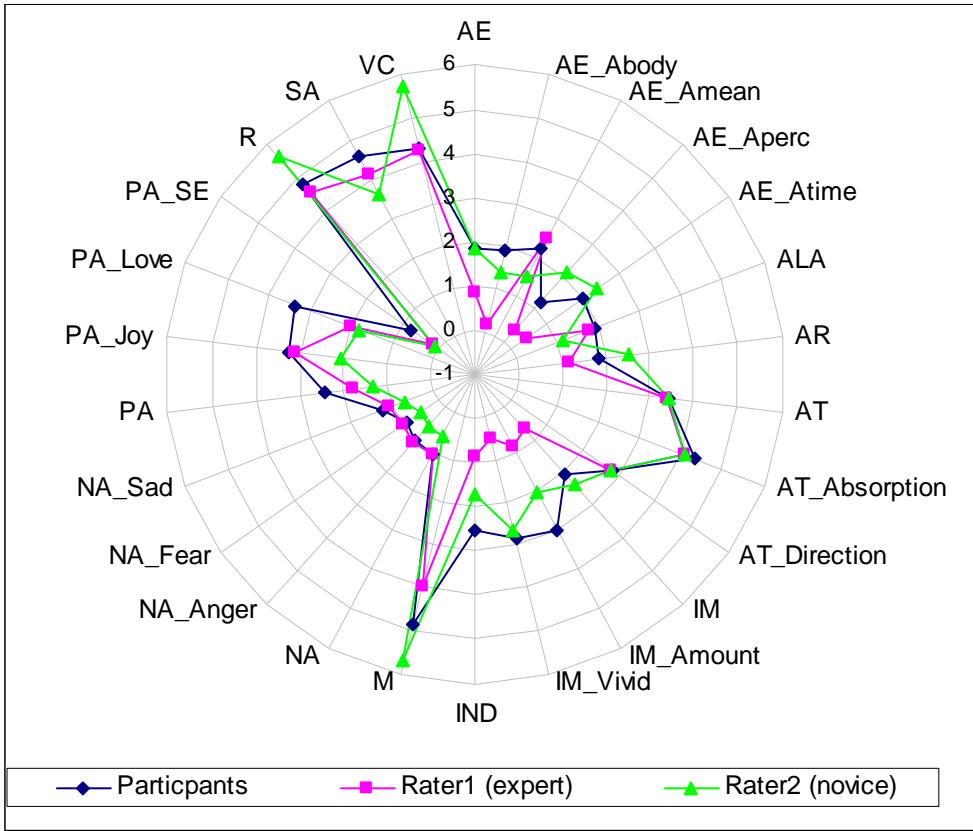


Figure 2. Radar plot of PCI ratings of 39 RSEs for participants and two raters. Abbreviations of PCI dimensions are defined in Table 1.

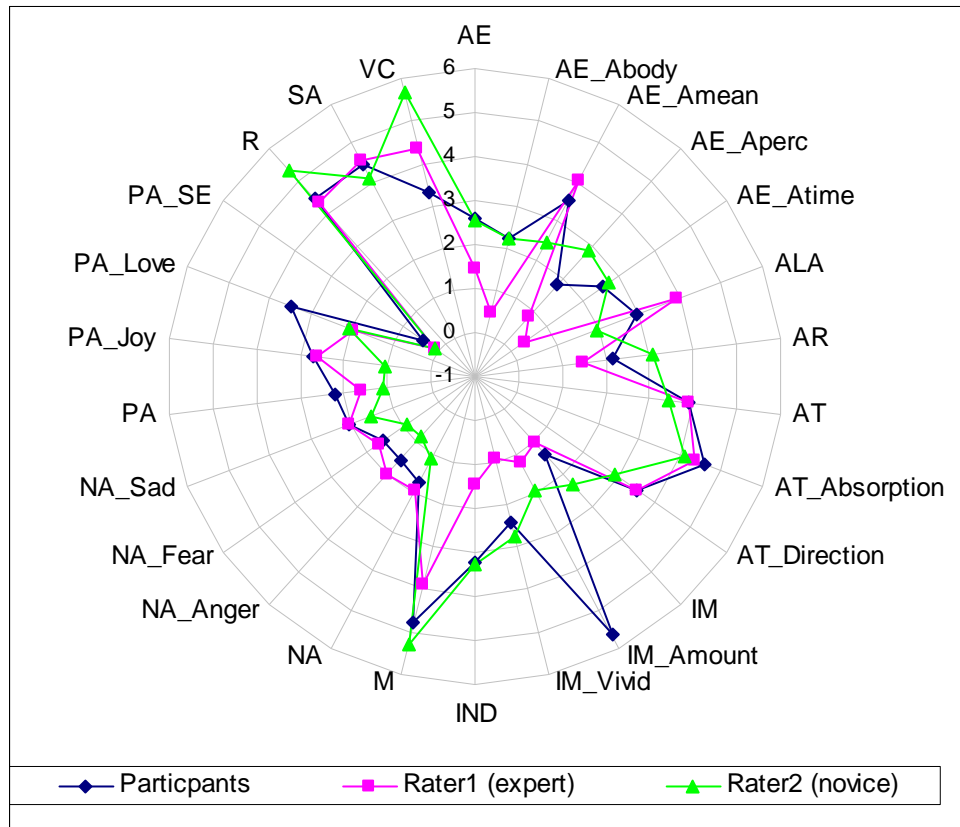


Figure 3. PCI dimensions for which religious experiences displayed significant differences from both happy and ordinary experiences. Error bars indicate 95% confidence levels.

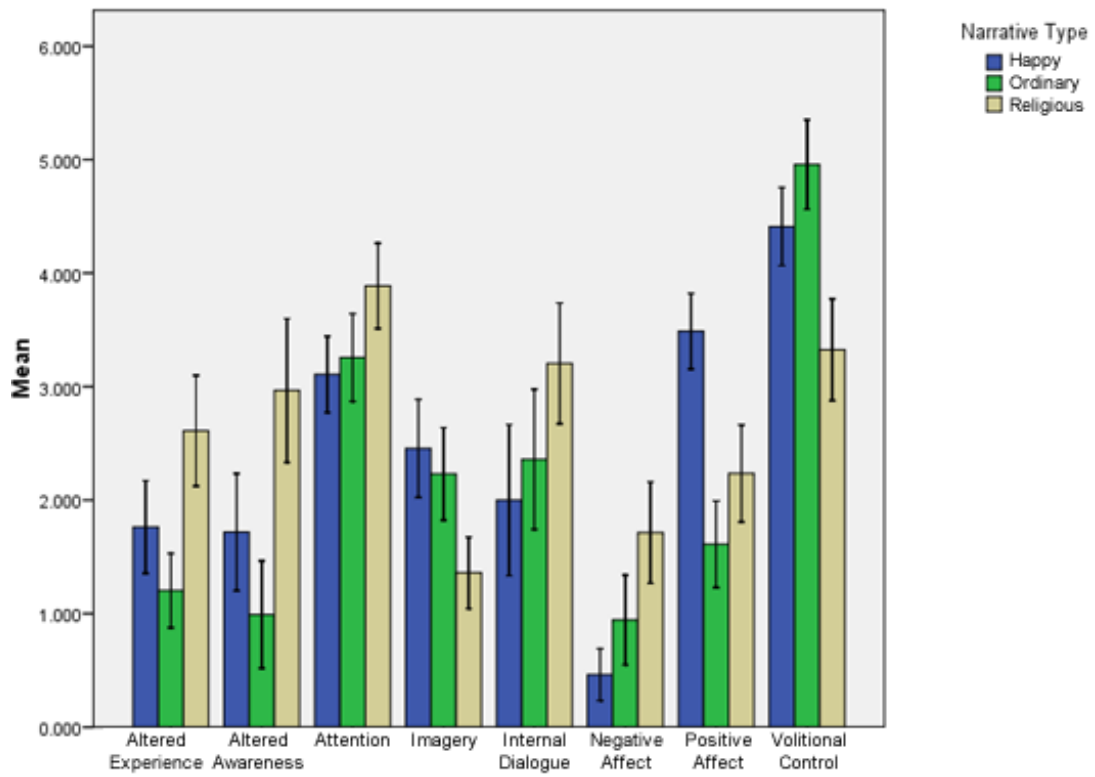


Figure 4. The PCI's 12 major dimensions for happy vs ordinary vs religious experiences.

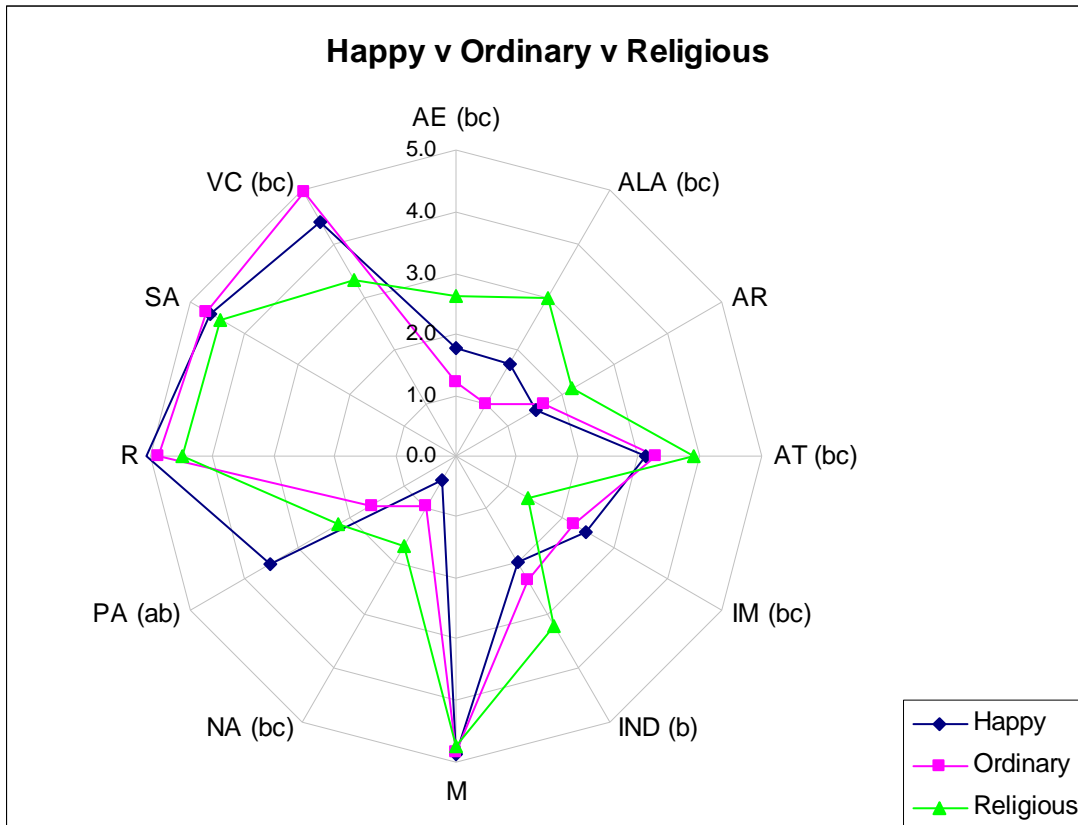


Figure Notes: Abbreviations of PCI dimensions are defined in Table 1. Pair-wise significance results are expressed using superscripts, as follows: a–significant difference between happy and ordinary; b–significant difference between happy and religious; c–significant difference between ordinary and religious.

Figure 5. The PCI's 26 major and minor dimensions for happy vs ordinary vs religious experiences.

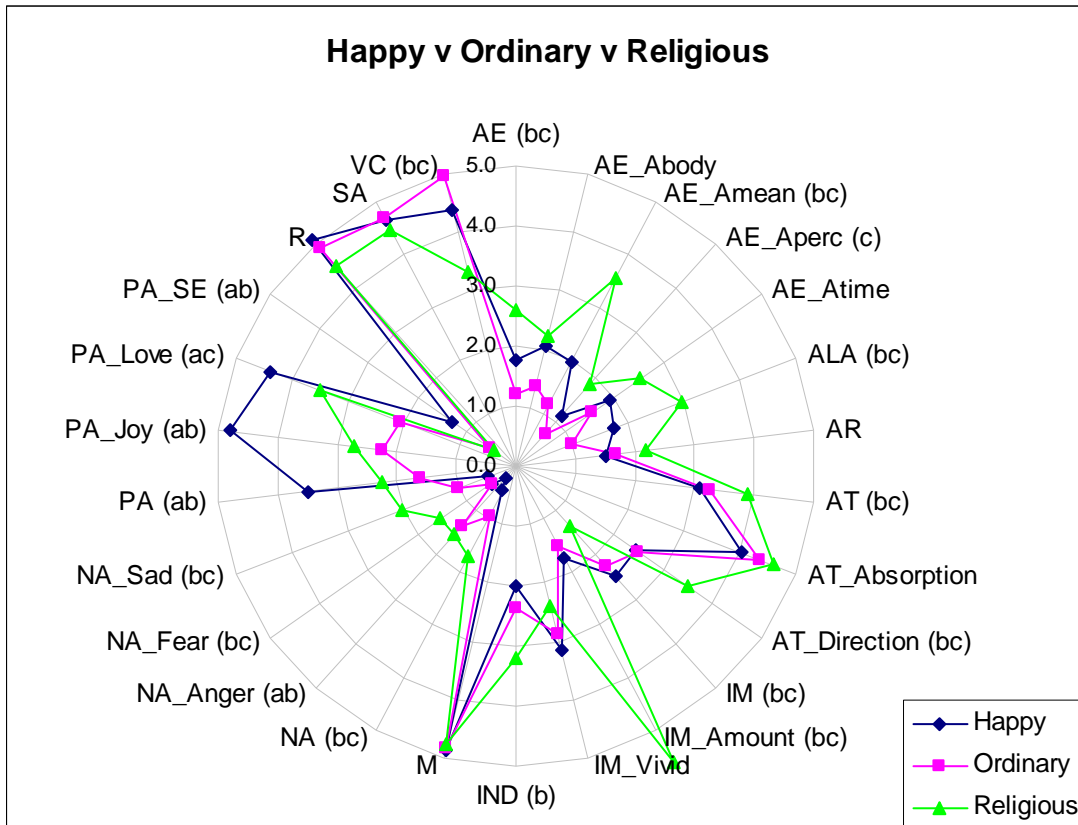


Figure Notes: Abbreviations of PCI dimensions are defined in Table 1. Pair-wise significance results are expressed as follows: a—significant difference between happy and ordinary; b—significant difference between happy and religious; c—significant difference between ordinary and religious.

Table and figure captions in the order in which they appear in the paper

[PAGE 13] Table 1. The 26 dimensions of the Phenomenology of Consciousness Inventory (PCI), grouped into 12 major dimensions (major dimensions in italics)

[PAGE 17] Table 2. Comparison of participant PCI ratings based on introspective recall versus two independent rater's narrative-based PCI scores.

[PAGE 17] Figure 1. Radar plot of PCI ratings of 117 experiences (happy, ordinary, and religious) for participants and two raters. Abbreviations of PCI dimensions are defined in Table 1.

[PAGE 17] Figure 2. Radar plot of PCI ratings of 39 RSEs for participants and two raters. Abbreviations of PCI dimensions are defined in Table 1.

[PAGE 21] Table 3. PCI means (with standard deviations) for happy versus ordinary versus religious experiences.

[PAGE 22] Figure 3. PCI dimensions for which religious experiences displayed significant differences from both happy and ordinary experiences. Error bars indicate 95% confidence levels.

[PAGE 22] Figure 4. The PCI's 12 major dimensions for happy vs ordinary vs religious experiences.

[PAGE 22] Figure 5. The PCI's 26 major and minor dimensions for happy vs ordinary vs religious experiences.

[PAGE 23] Table 4. One-way ANOVA showing means (with standard deviations) for Happy versus Ordinary versus Religious experiences on major PCI affect dimensions.

[PAGE 23] Table 5. One-way ANOVA showing means (with standard deviations) for Happy versus Ordinary versus Religious experiences on LIWC 2001 emotion word.