Empirical Challenges to Conventional Mind-Brain Theory

Edward F. Kelly
Emily Williams Kelly
Department of Psychiatry and Neurobehavioral Sciences
University of Virginia
210 Tenth St. NE
Charlottesville, VA USA 22902
ek8b@virginia.edu
ewc2r@virginia.edu
434-924-2281
434-924-1710 (fax)

Synopsis: *Despite many significant accomplishments, mainstream scientific psychology has not provided a satisfactory theory of mind, or solved the mind-body problem, and physicalist accounts of the mind are approaching their limits without fully accounting for its properties. The computational theory of mind has collapsed, forcing physicalism to retreat into what necessarily constitutes its final frontier, the unique biology of the brain, but this biological naturalism seems destined to fare little better. Some critical properties of human mental life can already be recognized as irreconcilable in principle with physical operations of the brain, and others appear likely to prove so as well.*

I. Introduction
II. Extreme Psychophysiological Influence
III. Extremes of Informational Capacity and Precision
IV. Memory
V. Psychological Automatisms and Secondary Centers of Consciousness
VI. Psi Phenomena
VII. Genius-Level Creativity
VIII. Mystical Experience
IX. The Unity of Conscious Experience
X. The Heart of the Mind
XI. Conclusion: Toward an Expanded Scientific Psychology

I. INTRODUCTION

Nearly all contemporary psychologists, neuroscientists, and philosophers subscribe – explicitly or implicitly – to some version of physicalism [see **Mind-Body Problem**]. Physicalist conceptions of human mind and personality, contrary to traditional and everyday notions, run along roughly the following lines: We human beings are nothing but extremely complicated biological machines. Everything we are and do is in principle causally explainable from the bottom up in terms of our biology, chemistry, and physics – ultimately, that is, in terms of local contact interactions among bits of matter moving in accordance with mechanical laws under the influence of fields of force. Some of what we know, and the substrate of our general capacities to

learn more, are built-in genetically as complex resultants of biological evolution. Everything else comes to us directly or indirectly by way of our sensory systems, through energetic exchanges with the environment of types already largely understood. Mind and consciousness are generated by – or in some mysterious way identical with – neurophysiological events and processes in the brain. Mental causation, free will, and the "self" do not really exist; they are mere illusions, ineffectual by-products of the grinding of our neural machinery. And since mind and personality are entirely products of our bodily machinery, they are necessarily extinguished, totally and finally, by the demise and dissolution of that body.

Views of this sort unquestionably hold sway over the vast majority of contemporary scientists, and they have also percolated widely through the public at large. We believe, however, that they are at best seriously incomplete, and at certain critical points demonstrably false, empirically. In this article, we will briefly catalogue a variety of interrelated empirical phenomena that appear difficult or impossible to explain in conventional physicalist terms. We emphasize from the outset that these phenomena must be considered collectively, not piecemeal; they not only challenge the conventional physicalist picture individually, but converge in pointing to the need for a radically novel way of understanding the intimate relationship of mind and brain. We also emphasize that we are presenting here only a skeletal outline of the kinds of phenomena to which we wish to direct readers' attention. Much fuller treatments of relevant empirical evidence and the issues raised can be found through works cited in the Bibliography.

II. EXTREME PSYCHOPHYSIOLOGICAL INFLUENCE

We begin with a variety of phenomena especially suggestive of the involvement of direct mental agency in the production of physiological effects not fully explainable in terms of physiological mechanisms alone. The following examples will serve to capture their flavor.

Placebo effects and related kinds of psychosomatic phenomena have long been informally recognized and are now widely accepted, but they were accepted by modern biomedical science only grudgingly, as new mechanisms of brain-body interaction came to light that seemed potentially capable of explaining them. In particular, psychoneuroimmunology has demonstrated the existence, previously unknown, of interactions between the central nervous system and the immune system. Nevertheless, the adequacy of such explanations to account even for placebo effects remains in question, and there are many types of kindred phenomena that pose progressively greater challenges to explanation in such terms.

For example, Sigmund Freud and F. W. H. Myers were impressed by hysterical "glove anesthesias," in which a patient loses sensation from the skin of a hand in the absence of organic lesion, in a way that typically corresponds only to the patient's idea, in complete disregard of the underlying anatomical organization. Related phenomena have often been reported in the context of hypnosis [see **Hypnosis**]. For example, highly suggestible persons who can vividly imagine undergoing an injurious circumstance such as receiving a burn to the skin sometimes suffer effects closely analogous to those that the physical injury itself would produce, such as a blister. More rarely, the correspondence between the hypnotic blister and its imagined source extends even to minute details of geometric shape, details too specific to account for in terms of known mechanisms of brain/body interaction. A closely related and well-documented phenomenon is that of "stigmata," in which fervently devout or pious believers in Christ develop wounds

analogous to those inflicted during the crucifixion. The injuries are again localized and specific in form, vary in locus and character in accordance with their subjects' differing conceptions of Christ's own injuries, and appear and disappear, often suddenly and regularly, also in accordance with the subject's expectations. Similarly dramatic phenomena have occasionally been documented in psychiatric patients in connection with their recall of prior physical trauma.

The conventional hope, of course, is that even the most extreme of the phenomena just mentioned might ultimately be explained in terms of brain processes. Continuing allegiance to this hope, despite the indicated explanatory difficulties, is undoubtedly encouraged by the fact that the phenomena described so far all involve effects of a person's mental states on that person's own body. Still more drastic explanatory challenges are posed, however, by cases in which one person's mental state seems to have directly influenced another person's body. Such phenomena include "maternal impressions" (birthmarks or birth defects on a newborn that correspond to an unusual and intense experience of the mother during the pregnancy), distant healing (including experimental studies of effects of prayer on healing), experimental studies of distant mental influence on living systems, and cases in which a child who claims to have memories of the life of a deceased person also displays extremely unusual birthmarks or birth defects corresponding closely with marks (usually fatal wounds) on the body of that person. In addition, there has been a considerable accumulation of spontaneous cases and experimental evidence demonstrating the reality of psychokinesis (PK), which by definition involves direct mental influence on the physical environment (see Section VI below).

III. EXTREMES OF INFORMATIONAL CAPACITY AND PRECISION

A number of well-documented psychological phenomena involve levels of detail and precision difficult to account for in terms of a brain operating in statistical fashion with neural components of low intrinsic precision and reliability. Here are some examples.

The first involves a case of "automatic writing" observed by William James. The subject wrote with his extended right arm on large sheets of paper, his face meanwhile buried in the crook of his left elbow. For him to see what he was doing was "a physical impossibility." Nevertheless, James continues: "Two or three times in my presence on one evening, after covering a sheet with writing (the pencil never being raised, so that the words ran into each other), he returned to the top of the sheet and proceeded downwards, dotting each *i* and crossing each *t* with absolute precision and great rapidity."

This episode illustrates two features that have often appeared together in the large but neglected scientific literature dealing with automatic writing: The subject is in an altered state of consciousness, and the motor performance, itself remarkable, is apparently guided by an extremely detailed memory record, an essentially photographic representation of the uncompleted page.

The latter property relates to the phenomenon of eidetic imagery, our second example, the most dramatic demonstration of which has been provided by Charles Stromeyer using Julesz stereograms. These are essentially pairs of computer-generated pictures, each of which by itself looks like a matrix of randomly placed dots, but constructed in such a way that when viewed simultaneously (by presentation to the two eyes separately) a visual form emerges in depth. Stromeyer presented pictures of this type to the eyes of his single subject, a gifted female

eidetiker, at different times, ultimately as much as 3 days apart. Under these conditions, the subject could only extract the hidden form if she could fuse current input to one eye with an extremely detailed memory-image of previous input to the other eye. Remarkably, she was able to succeed under a wide variety of increasingly demanding conditions. The original stereograms, for example, were 100 x 100 arrays, but she ultimately succeeded under double-blind conditions with arrays as large as 1000 x 1000, or a million "bits," viewed up to 4 hours apart.

These results were understandably shocking to many psychologists, who sought to escape their force by pointing to the dependence on a single subject and the absence of replications. At least one successful replication has subsequently occurred, however. Moreover, the literature already contains many additional examples of prodigious memory. Stromeyer mentions Luria's famous mnemonist and the case of the "Shass Pollaks," who memorized all 12 volumes of the Babylonian Talmud, and Oliver Sacks has reported a similar case of a person who among other things knew by heart all 9 volumes and 6000 pages of Grove's *Dictionary of Music and Musicians*. Other examples could easily be cited. Prodigious memory of this sort is a real psychological phenomenon.

Third in this group is the family of "calculating prodigies." Of special interest is the "savant syndrome," often associated with autistic disorders, in which islands of spectacular ability appear in the midst of generalized mental disability [see **Savant Syndrome**]. The abilities are of many types, but almost invariably involve prodigious memory. The depth of the problems they pose for brain theory is exemplified by the case of "The Twins," also described by Sacks. These profoundly impaired individuals, unable to perform even simple additions and subtractions with any accuracy, nonetheless proved able to generate and test prime numbers in their heads. Sacks was able to verify the primacy up to10 digits, but only by means of published tables, while the twins themselves went on exchanging numbers of steadily greater length, eventually reaching 20 digits. Sacks makes the intriguing suggestion that they may not literally be calculating these enormous numbers, but discovering them by navigating through some vast inner iconic landscape in which the relevant numerical relations are somehow represented pictorially. The twins themselves of course cannot say how they do it.

Phenomena of these sorts look hard to explain in terms of brain processes. The most serious attempt to do so known to us is in fact devoid of specific neural mechanisms. Its central argument is rather that early-stage brain processes like those subserving visual perception, for example, must also be savant-like in terms of their speed, precision, and informational capacity; what is unusual about savants, therefore, may consist merely in their access to these mechanisms. This explanation of course presupposes a positive answer to the fundamental question at issue, whether the brain alone can accomplish any of these things including perceptual synthesis itself (see Section IX below).

As proved long ago by mathematician John von Neumann, the only practical way to get increased arithmetical precision out of individually unreliable neurons is to use more of them. This biocomputational perspective clearly implies that calculating prodigies must use large portions of their brains in very abnormal ways to achieve the observed effects. The cognitive deficits that often accompany savant-type skills could conceivably reflect such substitutions, but we must remember that comparable skills sometimes also occur in geniuses such as the mathematicians Gauss and Ampére.

IV. MEMORY

The previous section focused on phenomena such as high-precision calculations and prodigious memory that appear incompatible with the physical properties of the brain considered as a kind of computing device. Problems also arise, however, in regard to memory in its more familiar and everyday forms. Here we briefly sketch some relevant issues.

Memory is central to all human cognitive and perceptual functions, yet we remain largely ignorant of where and in what forms our past experience is stored and by what means it is brought to bear upon the present. Generations of psychologists and neurobiologists have taken it as axiomatic that all stored memories must exist in the form of "traces," physical changes produced in the brain by experience, but there has been little real progress toward scientific consensus on the details of these mechanisms despite many decades of intensive research [see **Memory**].

Significant progress has recently been made, to be sure, in regard to "learning" and "memory" in simple creatures such as the sea-slug (Aplysia), and more generally in regard to what might be called "habit memory," the automatic adjustments of organisms to their physical environments. But these discoveries fall far short of providing satisfactory explanations of the most central and important characteristics of the human memory system, including in particular our supplies of general knowledge (semantic memory) and our ability to recall voluntarily and explicitly our own past experience (autobiographical or episodic memory). Furthermore, recent functional neuroimaging studies, although generating vast amounts of data, have yielded little if any progress toward a comprehensive and coherent account of memory based on trace theory.

Meanwhile, deep conceptual problems have been identified in trace theory itself. For example, autobiographical memory clearly involves something more than mere revival of traces of experiences past, something that allows us to interpret what is experienced now as a representation of our own past rather than a contemporary perception, dream, or hallucination. Traces as such, that is, only provide memory-aids rather than memories per se, and it has proven extremely difficult to specify in conventional physicalist terms what that extra something is, without falling into regressive forms of explanation that presuppose and hence cannot explain the phenomenon of memory itself. Similarly, the content of a concept or semantic memory typically transcends any finite set of experienced circumstances that can plausibly be imagined as having deposited corresponding "traces" in a form capable of explaining its effective deployment in an unlimited variety of novel contexts.

These conceptual problems regarding trace theories of memory have deep connections with issues discussed in Section X below, and similar issues arise in relation to allied components of current cognitive theory such as "information" and "representation." See also Section VI for an additional empirical problem.

V. PSYCHOLOGICAL AUTOMATISMS AND SECONDARY CENTERS OF CONSCIOUSNESS

Phenomena catalogued under this heading involve what looks like multiple concurrent engagement, in potentially incompatible ways, of major cognitive skills (linguistic skills, for example) and the corresponding brain systems.

Current cognitive neuroscience pictures the mind or "cognitive system" as a hierarchically ordered network of subprocessors or "modules," each specialized for some particular task and corresponding (it is hoped) to some particular brain region or regions. Leaving aside major issues regarding the details of its specification, this picture seems broadly consistent with the overall manner in which our minds seem normally to operate. Our basic way of consciously doing things, that is, is essentially one at a time in serial fashion. Although psychologists recognize that with suitable training people can do more things in parallel than they customarily suppose, this generalization applies mainly to relatively divergent things, and conspicuously fails as the simultaneous tasks become more complex and more similar.

Nevertheless, a large body of credible evidence, some dating back to the late 19th century, demonstrates that additional "cognitive systems," dissociated psychological entities indistinguishable from full-fledged conscious minds or personalities as we normally understand these terms, can sometimes occupy the same organism simultaneously, carrying on their varied existences as it were in parallel, and largely outside the awareness of the primary, everyday consciousness. In essence, the structure that cognitive psychology conventionally pictures as unitary, as instantiated within and identified with a particular organization of brain systems, can be functionally divided – divided, moreover, not "side-to-side", leading to isolation of the normal cognitive capacities from each other, but "top-to-bottom", leading to the appearance and concurrent – not alternating – operation of what seem to be two or more complete cognitive systems each of which includes all of the relevant capacities. Emergent "multiple" or "alter" personalities also can differ widely, not only in demeanor, interests, and knowledge but even in regard to non-voluntary physiological characteristics such as visual defects and susceptibilities to allergies. Even worse, it sometimes happens that one of these personalities appears to have direct access to the conscious mental activity of one or more others, but not vice-versa.

Two brief examples, drawn from an enormous literature, may help convey a more concrete sense of the character of these phenomena.

The first comes from a report by Oxford philosopher F. C. S. Schiller on automatic writing produced by his brother. As is characteristic of this genre of automatisms, the writer was typically unaware of the content of his writing, which went on continuously while he was fully and consciously engaged in some other activity such as reading a book or telling a story. Of particular relevance here, however, were occasions on which he wrote simultaneously with both hands and on completely different subjects, one or the other of these streams of writing also sometimes taking mirror-image form.

The second example is the case of Anna Winsor, described by William James in his report on automatic writing. The case was protracted and bizarre, but only superficially resembles the neurological "alien hand" (Dr. Strangelove) syndrome. Its central feature is that the patient, Anna, at a certain point lost voluntary control of her right arm, which was taken over by a distinctive secondary personality. This personality, whom Anna herself named "Old Stump," was benign, often protecting Anna from her pronounced tendencies toward self-injury. As in the case of Schiller's brother, Stump typically wrote or drew while Anna was occupied with other matters. But Stump also continued writing and drawing even when Anna was asleep, and sometimes in total darkness. This secondary personality also remained calm and rational during periods when Anna was feverish and delusional, and it manifested knowledge and skills which Anna herself did not possess.

VI. PSI PHENOMENA

Here we refer to experimental and field observations systematically adduced in the course of over a century of effort by workers in "psychical research" and its modern descendent, "parapyschology." The phenomena in question involve, by definition, correlations occurring across physical barriers that should be sufficient, on presently accepted physicalist principles, to prevent their formation. This occurs, for example, when person A spontaneously experiences an apparition of his friend B, as B unknown to A lies dying from a fatal accident. Over a thousand detailed cases of this sort – carefully documented experiences that are not dismissible *en masse* as mere "anecdotes" – have been published in the peer-reviewed literature. It also occurs when an experimental subject consistently succeeds in identifying randomly selected forced-choice targets displayed in a remote location. It is not difficult to set up controlled experiments of this sort and to evaluate their outcomes using rigorous statistical procedures. A large amount of careful experimental work has been carried out along these lines, with results more than sufficient, in our opinion, to demonstrate beyond reasonable doubt to open-minded persons that the sheer existence of the basic input/output phenomena – "extrasensory perception" (ESP) and "psychokinesis" (PK) in the popular vocabulary, or in more theory-neutral terminology, "psi" – is a fact of nature with which we must somehow come to scientific terms.

Psi phenomena in general are important because they provide examples of human behavioral capacities that appear impossible to account for in terms of presently recognized computational, biological, or classical-physics principles. Even more important for our purposes, however, is a further body of evidence suggestive of post-mortem survival, the persistence of elements of mind and personality following bodily death. It is simply not true, as most scientists presume, that we possess no such evidence. We in fact possess a lot of such evidence, much of it of very high quality, deriving for example from studies of veridical apparitions, trance mediumship, and "cases of the reincarnation type", in which young children spontaneously report verifiable events from the lives of distant and ordinary persons now deceased. Ironically, the primary threat to a survivalist interpretation of this accumulated evidence arises not from considerations of evidential quality, but from the difficulty of excluding alternative explanations based upon psi interactions involving only living persons.

Quite apart from any personal or theological interests readers may bring to this subject, it should be evident that post-mortem survival, if it occurs, demonstrates dramatically the limitations of present-day reductive physicalism. If it is the case, for example, as much evidence indicates, that autobiographical, semantic, and procedural (skill) memories can survive bodily death, then memory in living persons must presumably exist at least in part outside the brain and body as conventionally understood.

Either horn of this interpretive dilemma – postmortem survival or psi among the living – is lethal to current physicalist orthodoxy, which undoubtedly explains the widespread scientific resistance to both. But as we are arguing here, and have argued in much more detail elsewhere, these phenomena cannot be isolated and quarantined, because similarly difficult explanatory challenges are posed by many other well-evidenced psychological phenomena. Evidence for the occurrence of psi phenomena in general and post-mortem survival in particular must, we believe, play an important role in the formulation of an empirically adequate mind/brain theory, and our

efforts here will be amply rewarded if they lead scientifically-minded readers to examine these subjects more seriously than they otherwise might.

VII. GENIUS-LEVEL CREATIVITY

Any scientific theory of personality and cognition truly worthy of the name surely must help us to understand this humanly vital topic, but by this standard we have so far made distressingly little progress [see **Genius, Eminence, and Giftedness**; **Creative and Imaginative Thinking**; and **Creativity**]. The reason, in our opinion, is that for the most part we have tried to understand the exceptional – real genius, in its fullest expressions – as an amplification of the commonplace – "creativity," as found in convenience samples of undergraduates and the like.

All of the challenging phenomena catalogued in this article – including extreme psychophysiological influence, psychological automatisms and secondary centers of consciousness, flashes of inspiration involving unusual forms of thinking and symbolism, prodigious memory, spontaneous psi phenomena, and altered states of consciousness verging on the mystical realm – are inescapably bound up with genius in its fullest expressions, but these connections go virtually unmentioned in contemporary mainstream discussions. A particularly dramatic case which exemplifies our central point is that of the Indian mathematical genius Ramanujan, rated by his distinguished discoverer Hardy as standing alone at 100 atop a scale of mathematical ability on which most of us lie at or near zero, while the magnificent David Hilbert rated 80 and Hardy himself a mere 25. Replete with examples of prodigious memory, psychological automatisms, mathematical discoveries presented in the form of dreams, and profound and beautiful intuitions of hidden but ultimately verifiable properties of the physical world, this astonishing case fairly beggars the theoretical apparatus currently available to cognitive science and hence could well serve as a kind of reality check and navigational aid for further investigations of genius.

VIII. MYSTICAL EXPERIENCE

Experiences of this type lie at the core of the world's major religious traditions and have continued to occur throughout history and across cultures [see **Psychology and Religion**]. Their existence as a distinctive and important class of psychological phenomena can scarcely be denied, yet they have largely been ignored by mainstream psychology and neuroscience, and generations of clinical psychologists, psychiatrists, and neuroscientists have tended with few exceptions to devalue and pathologize them, treating them as products of malfunctioning brains. Even when acknowledging that such experiences are typically life-transforming and self-validating for those who have them, the historically standard epistemological approaches in psychology and philosophy – beginning with William James in his *Varieties of Religious Experience* – treat them as purely subjective events having authority only for those who experience them, and thus deny their objective significance and the testability of the associated truth-claims. However, a large though scattered literature testifies to the common occurrence in connection with such experiences, or in individuals who have them, of genius-level creativity, spontaneous psi-type events, and many other unusual but verifiable empirical phenomena of the sorts described in this article. Mystical-type states of consciousness are also at least partially

reproducible by pharmacological (psychedelic) means [see **Drugs, the Brain, and Behavior**], and they can be induced by protracted self-discipline involving transformative practices such as the various forms of meditation [see **Meditation**]. An objective and informed appraisal of mystical experience thus finds within it much additional support for an enlarged conception of human personality, and many new opportunities for empirical research.

IX. THE UNITY OF CONSCIOUS EXPERIENCE

Under this heading we will briefly address two interrelated problems. The first and narrower is the so-called "binding" problem, which emerged as a consequence of the success of contemporary neuroscientists in analyzing sensory mechanisms, particularly in the visual system [see **Visual Perception**]. It turns out that different properties of a visual object such as its form, color, and motion in depth are handled individually by largely separate regions or mechanisms within the brain. But once the stimulus has been thus dismembered, so to speak, how does it get back together again as a unit of visual experience?

Only one thing is certain: The unification of experience is not achieved anatomically. There are no privileged places or structures in the brain where everything comes together, either for the visual system itself or for the sensory systems altogether. Some early theorists such as James and McDougall argued that the evident disparity between the multiplicity of physiological processes in the brain and the felt unity of conscious experience could only be resolved in materialist terms by anatomical convergence, and since there is no such convergence, materialism must be false. This argument, although ingenious, relied upon the faulty premise that the only possible physical means of unification must be anatomical in nature. All current neurophysiological proposals for solving the binding problem are instead functional in nature; the essential concept common to all of them is that oscillatory electrical activity in widely distributed neural populations can be rapidly and reversibly synchronized, particularly in the "gamma" band of EEG frequencies (roughly 30-70 Hz), thereby providing a possible mechanistic solution to the binding problem [see **Electroencephalography**].

A great deal of sophisticated experimental and theoretical work over the past 20 years has demonstrated that such mechanisms do in fact exist in the nervous system, and that they are active in conjunction with normal perceptual synthesis. Indeed, contemporary physicalism has crystallized neurophysiologically in the form of a family of "global workspace" theories, all of which make the central claim that conscious experience occurs specifically – and only – in conjunction with large-scale patterns of gamma-band oscillatory activity linking widely separated regions of the brain [see **Consciousness**].

The neurophysiological global workspace, however, cannot be the whole story, because a large body of recent evidence demonstrates that elaborate, vivid, and life-transforming conscious experience sometimes occurs under extreme physiological conditions, such as deep general anesthesia and cardiac arrest, that categorically preclude workspace operation. In short, it appears to us that the early theorists were right after all, albeit for the wrong reason. In effect, we believe, recent progress in theoretical neuroscience, coupled with advances in our capacity to retrieve patients from the borderland of death, has provided new means for the falsification of physicalist theories of mind-brain relations [see **Near-Death Experiences**].

Availability of this emerging evidence emboldens us to make some further and more speculative remarks regarding the larger problem of perceptual synthesis, and the direction in which things seem to us to be moving.

It is an historical fact that mainstream psychology has always tended on the whole to try to solve its problems in minimalist fashion and with as little reference as possible to what all of us experience every day as central features of our conscious mental life. The early workers in "mechanical translation," for example, imagined that they could do a decent job simply by constructing a large dictionary that would enable substitution of words in one language for words in the other. This approach failed miserably, and we were slowly driven, failed step by failed step, to the recognition that truly adequate translation presupposes understanding, or in short a full appreciation of the capacities underlying the human use of language.

A similar evolution is underway in regard to perceptual theory [see **Perceptual Systems (Overview)**]. Most of the work to date has taken a strongly "bottom-up" approach, which views perceptual synthesis as a kind of exhaustive calculation from the totality of input currently present at our sensory surfaces. Machine vision and robotics, for example, necessarily took this approach, and even in neuroscience it seemed to make sense to start with the most accessible parts of the perceptual systems – the end organs and their peripheral connections – and work our way inward. The great sensory systems themselves – vision, audition, somatosensation, and so on – were also presumed to operate more or less independently, and were in fact typically studied in isolation.

A separate tradition dating back at least to Kant and the early Gestalt theorists, and carried forward into the modern era by psychologists such as Ulric Neisser and Jerome Bruner, has been sensitive to the presence of "top-down" influences, both within and between sensory modalities. Although a few perceptual subsystems (such as those that produce incorrigible visual illusions) may be truly autonomous or "cognitively impenetrable," these seem to be isolated and special cases. A very different overall picture of perceptual synthesis is currently emerging in which top-down influences predominate. On this view perceptual synthesis is achieved not from the input, but with its aid. This is necessarily the case for example in regard to ambiguous figures such as the Necker cube, where the stimulus information itself is insufficient to determine a uniquely correct interpretation. More generally, we routinely ignore information that is present in the input and supply information that is not, speed-reading providing a characteristic example. Something within us, a sort of world-generating or virtual-reality system, is continuously updating and projecting an overall model of the perceptual environment and our position within it, guided by limited samplings of the available sensory information.

As in the case of understanding spoken or written language, an enormous amount of general knowledge is constantly mobilized in service of this projective activity, which freely utilizes whatever information it finds relevant. Top-down and cross-modal sensory interactions have recently been recognized as the rule rather than the exception in perception, and neuroscientist Rodolfo Llinás and his co-workers have advanced the view, which we believe is profoundly correct, that dreaming, far from being an odd and incidental part of our mental life, represents the fundamental form of this world-creating activity. Ordinary perceptual synthesis, on this inverted view of things, amounts to oneiric (dreamlike) activity constrained by sensory input. Psychoanalyst Ernest Hartmann has proposed similar ideas in regard to hallucinatory activity more generally, with dreaming included. On his view such activity is again a ubiquitous

and fundamental feature of our mental life, and the critical question is not "why do we sometimes hallucinate?" but rather "what keeps us from hallucinating most of the time?" The answer, he suggests, lies in inhibitory influences exerted by the brain activity that accompanies ongoing perceptual and cognitive functions of the ordinary waking sorts.

So far so good, but where exactly is the "top," the ultimate source of this top-down world-creating activity? The mainstream neuroscientists who have already recognized its existence invariably presume that it arises entirely within the brain itself, but evidence such as that of near-death experiences occurring under extreme physiological conditions, and the more direct evidence of post-mortem survival, suggests that it may originate outside the brain as conventionally understood.

X. THE HEART OF THE MIND

In this section we will comment briefly on a hornet's nest of issues lying at the core of mental life as all of us routinely experience it, every day. These issues have been the focus of extensive recent debates, especially in the philosophical literature, precisely because of their resistance to understanding in conventional physicalist terms. The issues are deep, individually complex, and densely interconnected, and what we can say here will necessarily amount to little more than a summary of our own opinions. Our central point is that the prevailing *a priori* commitment to physicalism has rendered us systematically incapable of dealing adequately with the mind's most central and characteristic properties. We should rethink that commitment.

Consider first the issue of semantic content, the "meaning" of words and other forms of representation. Throughout our history, we have tried unsuccessfully to deal with this by "naturalizing" it, reducing it to something else that seems potentially more tractable. An old favorite among psychologists was that representations work by resembling what they represent, by virtue of some sort of built-in similarity or structural isomorphism, but any hope along these lines was long ago exploded by philosophical arguments. The central move subsequently made by classical cognitive psychology is essentially the semantic counterpart of the prevailing "functionalist" doctrine in philosophy of mind: Meanings are not to be conceived as intrinsic to words or concepts, but rather as deriving from and defined by the functional role those words or concepts play in the overall linguistic system. Currently there is great interest in "externalist" causal accounts of this functionalist type; in connectionism, dynamic systems theory, and neuroscience, for example, the "meaning" of a given response, such as the settling of a network into one of its "attractors" or the firing off a volley of spikes by a neuron in visual cortex, is typically identified with whatever it is in the organism's environment that produces that response. But this simply cannot be right: How can such an account deal with abstract things, for example, or non-existent things? Responses do not qualify *ipso facto* as representations, nor signs as symbols. Something essential is being left out. That something, as John Searle has so effectively argued, is precisely what matters, the semantic or mental content.

Closely related to this is the more general and abstract philosophical problem of intentionality, the ability of any and all representational forms to be "about" things, events, and states of affairs in the world. Mainstream psychologists and philosophers have struggled to find ways of making intentionality intrinsic to the representations themselves, but again it just does not and cannot work, because something essential is left out. That something is the user of the

representations. Intentionality is inherently a three-way relation involving users, symbols, and things symbolized, and the user cannot be eliminated. As Searle puts it in various places, the intentionality of language is secondary and derives from the intrinsic intentionality of the mind. Searle thus agrees in part with 19th-century philosopher Franz Brentano, for whom intentionality was the primary distinguishing mark of the mental. At the same time, however, Searle ignores the other and more fundamental part of Brentano's thesis, which is that intentionality cannot be obtained from any kind of purely physical system, including brains.

Talk of "users" and the like raises for many contemporary psychologists and philosophers the terrifying specter of the self as a homunculus, a little being within who embodies all the capacities we sought to explain in the first place. Such a result would clearly be disastrous, because that being would evidently need a similar though smaller being within itself, and so on without end. Cognitive modelers seeking to provide strictly physicalist accounts of mental functions must therefore do so without invoking a homunculus, but in attempting this they routinely fail. Often the homuncular aspect is hidden, slipped into a model by its designers or builders and covertly enlisting the semantic and intentional capacities of its users or observers. Much contemporary work on computational modeling of memory, metaphor, and semantics harbors subtle problems of this sort. Sometimes, however, the homunculus is more brazenly evident. One example is David Marr's account of vision, which applies computations to the two-dimensional array of retinal input in order to generate a "description" of the three-dimensional world that provided that input, but then needs someone to interpret the description. Another is Stephen Kosslyn's model of visual imagery, which essentially puts up an image on a sort of internal TV screen, but then needs somebody else to view the image.

Cognitive models cannot function without a homunculus, we believe, precisely because they lack what we have – minds, with their capacities for semantics, intentionality, and all the rest built in. No homunculus problem, however, is posed by the structure of our conscious experience itself. The efforts of Daniel Dennett and other physicalists to claim that there is such a problem, and use that to ridicule any residue of dualism, rely upon the deeply flawed metaphor of the "Cartesian theater," a place where mental contents get displayed and we pop in separately to view them. Descartes himself, James, and Searle, among others, all have this right; conscious experience comes to us whole and undivided, with the qualitative feels, phenomenological content, unity, and subjective point of view all built-in, intrinsic features. We and our experience cannot be separated in this way.

Finally, we wish simply to record our own deepest intuition as to where these issues lead. All of the great unsolved mysteries of the mind – semantics, intentionality, volition, the self, and consciousness – seem to us inextricably interconnected, with consciousness somehow at the root of all.

The consciousness we have in mind, however, is emphatically not that of people such as David Chalmers, irreducible but ineffectual, consisting merely of phenomenological properties or "qualia" arbitrarily tacked on to some sort of computational intelligence that supposedly does all the cognitive work. Ordinary perception, memory, and action are saturated with conceptual understanding, and conceptual understanding is saturated with phenomenological content. Volition too has an intentionality aspect, for as Nietzsche somewhere remarked, one cannot just will, one must will something. And as William James so forcibly argued at the dawn of our science, all of this perceptual, cognitive, and volitional activity somehow emanates from a

mysterious and elusive "spiritual self," which can often be sensed at the innermost subjective pole of our ongoing conscious experience.

We find it astonishing, and predict that it will be found so as well by our intellectual descendants, that so much of 20th-century psychology and philosophy sought – consciously! – to slight or ignore these first-person realities of the mind, and sometimes even to deny their existence. There is perhaps no better example of the power of pre-existing theoretical commitments to blind able persons to countervailing facts. The gloomy and counterintuitive modern conclusions summarized in Section I about mind, consciousness, free will, and the self really do follow – inexorably – from the physicalism that prevails today. But as we will next briefly explain, that kind of physicalism is itself incompatible with our deepest physical science.

XI. CONCLUSION: TOWARD AN EXPANDED SCIENTIFIC PSYCHOLOGY

It cannot be emphasized too strongly that these unresolved explanatory problems concerning consciousness, the heart of the mind, and the other empirical phenomena surveyed in this article all have a common source in the narrow physicalist consensus which undergirds practically everything now going on in mainstream psychology, neuroscience, and philosophy of mind. But that consensus rests ultimately upon a classical-physics-based conception of nature, deriving from people such as Descartes, Galileo, Newton, Laplace, and Kelvin, that began its career by deliberately banishing conscious human minds from its purview! Given that historical background, it should occasion little surprise that William James – like Newton and Leibniz before him, and like increasing numbers of philosophers and scientists today – clearly recognized the inherent impossibility of explaining consciousness and allied phenomena within that Procrustean framework. James himself cautioned that the physical-science concepts underlying classical physicalism were "provisional and revisable things," but he had no good alternatives in sight. As he correctly anticipated, however, that conception of nature was soon radically undermined by a tectonic shift in the foundations of physics itself, associated especially with the rise of quantum mechanics.

The founders of quantum mechanics discovered to their horror that the fundamental ideas of classical physics were not just limited but wrong, leading repeatedly to predictions falsified by experiment. The theory they were driven to in response, quantum theory, is a more fundamental and better physical theory that explains everything explainable in classical terms and a host of additional things as well, often to extraordinary levels of accuracy. No outcome predicted by it has ever been experimentally falsified. Furthermore, in at least some of its various interpretations quantum mechanics appears able to accommodate phenomena of the sorts surveyed here. Mathematical physicist Henry Stapp in particular has shown that a strictly orthodox interpretation derived from the mathematical formalization achieved by von Neumann leads naturally to a non-Cartesian form of dualism in which the human mind with its powers of attention and decision-making plays a necessary and fundamental role in completing the quantum dynamics. As a corollary, the classical doctrine of "causal closure of the physical", which underlies most contemporary physicalist denials of free will, is specifically rejected [see **Free Will**]. And although details remain to be supplied, many of the challenging behavioral phenomena cited above, from stigmata and hypnotic blisters to psi phenomena and even post-mortem survival, seem potentially understandable within this broader framework.

The empirical challenges briefly surveyed here should be sufficient in themselves, we believe, to compel and to some extent foreshadow a radical reworking of central parts of our science of the mind. But it is also important to recognize that a scientific psychology enlarged in these ways will likely prove more compatible than present-day physicalist psychology both with everyday human experience and with our most fundamental physical science.

BIBLIOGRAPHY

Chalmers, D. (1996). The conscious mind: In search of a fundamental theory. New York: Oxford University Press.

Gauld, A. (1982). Mediumship and survival: A century of investigations. London: Heinemann.

Hartmann, E. (1991). Boundaries in the mind: A new psychology of personality. New York: Basic Books.

James, W. (1890). Principles of psychology. (2 vols.). New York: Henry Holt.

James, W. (1986). Essays in psychical research. (F. H. Burkhardt, Ed.). Cambridge, MA: Harvard University Press.

Kanigel, R. (1991). The man who knew infinity: A life of the genius, Ramanujan. New York: Charles Scribner's Sons.

Kelly, E. F., Kelly, E. W., Crabtree, A., et al. (2007). Irreducible mind: Toward a psychology for the 21st century. Lanham, MD: Rowman and Littlefield.

Llinás, R., & Ribary, U. (1994). Perception as an oneiric-like state modulated by the senses. In C. Koch & J. L. Davis (Eds.), Large-scale neuronal theories of the brain (pp. 111-124). Cambridge, MA: MIT Press.

Marshall, P. (2005). Mystical encounters with the natural world: Experiences and explanations. Oxford: Oxford University Press.

Radin, D. (2006). Entangled minds: Extrasensory experiences in a quantum reality. New York: Simon and Schuster.

Sacks, O. (1987). The man who mistook his wife for a hat. New York: Simon and Schuster.

Searle, J. R. (1992). The rediscovery of the mind. Cambridge, MA: MIT Press.

Stapp, H. P. (2007). Mindful universe: Quantum mechanics and the participating observer. Berlin: Springer.

Stevenson, I. (1997). Reincarnation and biology: A contribution to the etiology of birthmarks and

birth defects. (2 vols.). Westport, CT: Praeger.

Strawson, G., et al. (2006). Consciousness and its place in nature: Does physicalism entail panpsychism? (A. Freeman, Ed.). Exeter: Imprint-Academic.

Stromeyer, C. F. III and Psotka, J. (1970). The detailed structure of eidetic images. **Nature** 225, 346-349.

Treffert, D. (1989). Extraordinary people: Understanding “idiot savants”. New York: Harper and Row.