

**Introduction to *Philosophy of Psychedelics* (Oxford University Press, 2021)**  
**Chris Letheby**

Please quote from published version.

## **1 Introduction**

*The unacceptability of psychedelic therapy... stems in part at least from this fundamental empirical fact: Through the psychedelic experience persons tend to accept beliefs which are at variance with the usual conception of the “scientific world view.”*

Willis Harman, ‘The Issue of the Consciousness Expanding Drugs’.

Something very strange is happening in psychiatry. Clinicians and researchers are becoming increasingly interested in a “new” experimental treatment, with some suggesting it might herald a “new paradigm” in the treatment of disorders such as anxiety, depression, and addiction (Nichols et al. 2017, Schenberg 2018). If promising initial results are borne out, this new treatment may significantly outperform existing pharmacological and psychological therapies in terms of rapidity and efficacy. And yet this treatment has a number of properties that, from the standpoint of mainstream philosophy and science, can only be described as extremely weird.

First, although it is a drug treatment, there is no daily dosing regimen of the kind familiar from existing antidepressant and antipsychotic medications. Patients receive one, or very few, doses of the drug—albeit with considerable interpersonal support and preparation—and then, in successful cases, their symptoms dramatically decrease for many months; no re-dosing required. Second, although it is *in some sense* a drug treatment, it’s far from clear that the drug itself is the direct cause of clinical improvement. Rather, it seems that the drug is merely a catalyst for a brief but intense conscious experience, and it is this experience that causes therapeutic effects. Third, and most weirdly, the kind of experience that seems to lead to the best outcomes is one that patients describe as “mystical” or “spiritual”. Many patients given this treatment report transcendent experiences of oneness, cosmic consciousness, or “ego dissolution”, and evidence suggests that these patients show the greatest clinical improvement.

I am referring, of course, to *psychedelic therapy*: the treatment of psychiatric disorder by the supervised administration of “hallucinogenic” serotonin-2a agonist substances such as LSD, psilocybin (the active ingredient in “magic” mushrooms), and DMT (a key psychoactive constituent of the Amazonian beverage *ayahuasca*.) Given their widespread reputation as psychotomimetic (psychosis-mimicking) or psychotogenic (psychosis-generating) chemicals, the very idea of using psychedelics to *treat* mental illness may seem absurd (Osmond 1957, Carhart-Harris et al. 2016a). However, these drugs were studied and prescribed as experimental psychiatric treatments in the 1950s and 60s, prior to the public controversy surrounding their use by the hippie counterculture. Promising results were reported from this line of research, which was curtailed prematurely due to the crackdown on psychedelics and the subsequent War on Drugs. For decades, virtually no human psychedelic research was conducted.

Since the 1990s, this research has been slowly but steadily resuming, and the results to date are intriguing. Although it is still relatively early days, there is now sufficient evidence to take seriously the claims of earlier researchers. Supervised psychedelic sessions may after all have an

acceptable safety profile (dos Santos et al. 2018a) coupled with significant and lasting psychological benefits for psychiatric patients (dos Santos and Hallak 2020) and even for healthy volunteers (Gandy 2019). Meanwhile, neuroimaging studies of the psychedelic state are providing tantalizing clues about the biological bases of consciousness and self-awareness (dos Santos et al. 2016b). If these preliminary findings are replicated, then psychedelics may well find a place in 21<sup>st</sup> century psychiatry (Sessa 2005, 2012, 2018). But many questions remain. The transformative mechanisms of psychedelic experience are incompletely understood, and this strikingly novel type of therapeutic intervention raises many fascinating and puzzling issues, both scientific and philosophical.

This book is organised around one specific philosophical question that relates to the role of the mystical experience in psychedelic therapy. Patients and subjects who show the greatest psychological benefit from psychedelic experiences tend to be those who report a mystical experience, as defined by widely used psychometric questionnaires, and often the degree of mystical experience predicts the degree of benefit. A mystical experience is sometimes described as an overwhelmingly powerful apparent encounter with “ultimate reality”. Space, time, and the sense of individual selfhood fade away, to be replaced by a sense of union with “another Reality that puts this one in the shade” (Smith 2000, p. 133). In some cases, this Reality is experienced as a divine or “cosmic” consciousness that underlies and unifies the entire manifest universe, evoking philosophical doctrines such as idealism and pantheism (Shanon 2002, p. 162). The apparent centrality of the mystical experience has led researcher Charles Grob to describe psychedelic therapy as an “existential medicine” (Grob 2007, p. 213).

But this picture of psychedelic therapy raises an obvious worry: what if the divine universal consciousness is not real? Many philosophers and scientists today subscribe to a broadly naturalist, materialist, or physicalist worldview, according to which mind and consciousness are not fundamental in the universe, but are relatively recent products of complex evolution. From this perspective, the mystical apprehensions of psychedelic subjects look like “metaphysical hallucinations” (Flanagan and Graham 2017)—subjectively compelling, but ultimately misleading, by-products of aberrant brain activity, on a par with psychedelic subjects’ visions of walls “breathing” or kaleidoscopic fractals (cf. Roche 2010). And there seems to be something seriously questionable about a treatment or enhancement modality that works by inducing metaphysical hallucinations. The journalist Michael Pollan, discussing the use of psychedelics to treat psychological distress in the terminally ill, puts the point forcefully: “Is psychedelic therapy simply foisting a comforting delusion on the sick and dying?” (Pollan 2015). The worry is that psychedelics bring about their salutary psychosocial effects by a deceptive, epistemically (and therefore ethically) bad mechanism.

Call this the *Comforting Delusion Objection* to psychedelic therapy. There are three popular responses. The first is that there’s no problem because the mystical experience is veridical. Far from inducing metaphysical hallucinations, psychedelics afford subjects a direct and transformative apprehension of ultimate reality (Smith 2000, Richards 2015). The second response is that the epistemic status of psychedelic experiences is relatively unimportant; what is more important is that they help people live better lives (Flanagan and Graham 2017.) The third response is that the epistemic status of psychedelic experiences is very important, and also poor; we should be wary of permitting or prescribing psychedelic therapy, because it does in fact work by the objectionable induction of comforting delusions (Lavazza 2017).

In this book I present a fourth, relatively unexplored response to the Comforting Delusion Objection. My central thesis is that the objection fails. My strategy is to start by assuming that (a) the naturalist worldview is true, so no cosmic consciousness exists, and (b) the epistemic status of psychedelic therapy is very important. This is the worst-case scenario for someone seeking to answer this objection. Clearly if a cosmic consciousness or divine Reality exists, the objection fails (response number one above); and clearly if the epistemic status of psychedelic therapy is relatively unimportant, the objection fails (response number two). Here I aim to show that the objection fails *even if* neither of these conditions holds—and this will show that the objection fails, whatever is the case.

My grounds for concluding that the Comforting Delusion Objection fails, even given assumptions (a) and (b) above, are twofold. First, the epistemic risks of psychedelic therapy, from a naturalistic standpoint, are less than one might suppose. Despite first appearances, psychedelic therapy does not work by instilling comforting metaphysical beliefs in a divine Reality. Such beliefs are sometimes acquired, but they are not necessary for therapeutic benefits, nor do they always accompany them. The construct of a “mystical-type experience”, which mediates clinical outcomes, sometimes reflects experiences as of a cosmic consciousness or divine Reality. However, it sometimes reflects more naturalistic experiences of “ego dissolution”, connectedness, emotional catharsis, and psychological insight. Psychedelics promote well-being by disrupting the sense of self, allowing patients to access new ways of seeing themselves and their lives. Such existentially significant changes to self-awareness do not depend on changes to metaphysical beliefs about the ultimate nature of reality.

Second, psychedelic therapy is *epistemically innocent*, in the technical sense defined by Lisa Bortolotti (2015). Bortolotti defines epistemically innocent imperfect cognitions as those which have real epistemic flaws, but also offer significant epistemic benefits that are unavailable by any other means. This is the status of psychedelic therapy, given naturalism. Psychedelic experiences can lead to knowledge acquisition, both directly and indirectly—*even if* naturalism is true and the cosmic consciousness experience is a metaphysical hallucination. In the course of defending these claims, I will deal with issues concerning the nature of conscious experience, the relations between psychology and neuroscience, and the theory and philosophy of self-awareness. The second main thesis of the book is that a sustained and detailed interaction between philosophy and psychedelic science can be mutually beneficial: I aim to show that psychedelic science reveals phenomena of serious philosophical interest, while philosophy offers valuable tools for clarifying and interpreting results from psychedelic science.

I begin in chapter two by reviewing recent scientific evidence concerning the safety, and therapeutic and transformative efficacy, of carefully controlled psychedelic administration. The recent wave of clinical studies suggests that psychedelics can indeed be given safely and responsibly, without serious adverse effects, to carefully selected and prepared subjects in carefully controlled conditions. These studies also provide evidence for durable psychological benefits following one (or very few) psychedelic sessions conducted in this fashion. But they also provide evidence that these benefits are mediated, at least in part, by mystical experiences, giving rise to the Comforting Delusion Objection. I outline the objection and my proposed response to it in more detail. I suggest that what is needed is a *natural philosophy* of psychedelics: a synthetic, big-picture inquiry integrating multidisciplinary evidence to address philosophical issues in a manner continuous with science and consistent with naturalism.

In the spirit of such a natural philosophy, I begin with detailed attention to the phenomena under investigation. As such, chapter three presents an overview of the phenomenology of psychedelic therapy, including patients' own impressions of the therapeutic mechanisms. Psychedelic experiences are notoriously variable and often held to be ineffable; nonetheless, significant progress has been made on identifying typical themes from quantitative and qualitative reports. Subjects often describe an expansion of consciousness, a heightening of emotional experience, strange visions and insights, and a blurring of boundaries between self and world. Many subjects interpret these experiences in non-naturalistic terms, but many do not, instead emphasising experiences of psychological insight, emotional catharsis, acceptance, and *connectedness* to various aspects of self and world (Carhart-Harris et al. 2018a, Breeksema et al. 2020). This provides our first major clue that psychedelic therapy is not a simple matter of inducing existentially comforting metaphysical beliefs.

Although they provide suggestive evidence, we cannot assume that subjects' impressions of the therapeutic mechanisms are correct. To determine how psychedelic therapy works, we need to examine a broader range of evidence. In chapter four I survey three different theories of how psychedelics cause lasting psychological benefits. The first ascribes these benefits to psychedelics' direct effects on the molecular-level mechanisms of neuroplasticity. On this view, psychedelic therapy is an experience-independent pharmacotherapy, and the remarkable phenomenology is a mere cluster of therapeutically epiphenomenal "psychotomimetic" side-effects. The second and third theories differ in detail, but both ascribe lasting benefits to non-naturalistic metaphysical ideations—the transcendent vision of a "Joyous Cosmology" (Watts 1962) supposedly encountered in the mystical experience.

I argue that all three theories are inadequate. The neuroplasticity theory struggles to account for the robust correlation between measures of mystical-type experience and lasting benefits; this correlation suggests that the psychedelic experience itself is causally relevant to those benefits. However, on closer examination, we find that some subjects satisfy psychometric criteria for a "mystical-type experience" without having a transcendent vision of cosmic consciousness. These criteria are broad enough to capture more naturalistic experiences of "ego dissolution" and connectedness as well. Not all mystical-type experiences *in the psychometric sense* are non-naturalistic metaphysical hallucinations. The conclusion is that the key causal factor in psychedelic therapy is genuinely psychological—an aspect of the experience itself—but is not essentially tied to non-naturalistic metaphysical ideations.

In chapter five, I argue that the relevant psychological factor is changes to the sense of self. A few studies pinpoint experiences of psychological insight as strongly linked to lasting benefits, and the relevant insights are often autobiographical in character. Moreover, several studies have shown that psychedelics can induce lasting increases in "mindfulness-related capacities" for taking an open, non-reactive stance toward one's inner experience—a stance which intrinsically involves changes to the sense of self. These studies constitute an important experimental vindication of the old idea that there are deep commonalities between psychedelic and meditative states (e.g. Huxley 1954, Watts 1962, Leary et al. 1964). Finally, neuroimaging research has consistently implicated certain large-scale brain networks, the Default Mode and Salience networks, in psychedelics' lasting benefits, and both networks are linked to self-representation by considerable independent evidence.

At this point, a question arises: how exactly do neurobiological changes, such as modulation of the Default Mode and Salience networks, relate to psychological changes, such as

autobiographical insights and increased mindfulness-related capacities? I propose that *neurocognitive theory*, which attributes computational or information-processing functions to neural structures, provide a vital explanatory bridge between biological and psychological accounts of the psychedelic state (cf. Gerrans 2014). If we can specify the cognitive functions performed by the neural systems that psychedelics target, then this will allow us to explain why modulating those networks should lead to transformative experiences of ego dissolution, connectedness, catharsis, and insight.

In chapters six and seven I outline such a theory: the *predictive self-binding* account of psychedelic therapy (Letheby and Gerrans 2017). According to this account, one function of the networks targeted by psychedelics is to maintain a hierarchical predictive model of the self. This predictive self-model acts as a “centre of representational gravity”: by parsing information into self-relevant and self-irrelevant, into “me” and “not-me”, it functions as an organising principle that governs and constrains cognitive processing. In pathological conditions, detrimental forms of self-modelling often become rigidly entrenched. By “unbinding” the self-model, psychedelics facilitate experiences of ego dissolution and psychological insight in which pathological self-models can be revised. On this view, psychedelic therapy has a two-factor structure: it involves (a) the induction of neural and psychological plasticity at multiple levels, and (b) the discovery and consolidation of new forms of self-modelling.

I conclude chapter seven with some brief remarks on philosophical questions about self and self-consciousness. I have argued elsewhere that psychedelic evidence supports two controversial philosophical claims: that the self does not exist (Letheby and Gerrans 2017), and that there can be conscious experiences lacking all forms of self-consciousness (Letheby 2020). However, both arguments face serious objections. For present purposes, I content myself with two weaker, but still significant claims: (i) there can be conscious experiences lacking anything like the *ordinary* sense of self, and (ii) the *kind* of self that we automatically take ourselves to be does not exist. Theoretically and existentially, this is plenty to be getting on with.

The upshot of chapters four to seven is that the epistemic risks of psychedelic therapy, given naturalism, are surprisingly small. In chapter eight I argue that its epistemic benefits, given naturalism, are surprisingly large. The concept of “epistemic innocence” (Bortolotti 2015, 2020) encapsulates the overall epistemic status of psychedelic therapy on naturalism: this intervention carries non-trivial epistemic risks, insofar as some subjects do acquire strong beliefs in a cosmic consciousness or spirit world, but these risks are offset by the fact that it also offers significant, often unique, epistemic *benefits*. I survey the major proposals about psychedelic-induced knowledge gain that are consistent with naturalism, and argue that psychedelics offer what philosophers call *knowledge by acquaintance* with various often unrevealed aspects of the human mind, including its potential for diverse and beneficial modes of attention and cognition. This has connections with philosophical discussions of Frank Jackson’s (1982, 1986) famous thought experiment about Mary the super-neuroscientist. It also justifies the increasing use of the appellation “psychedelic”, meaning “mind-manifesting” or “mind-revealing” (Osmond 1957), in preference to the many available alternatives.

At later times, subjects can re-evolve these beneficial modes of attention and cognition, at least to some extent; therefore psychedelics also make available ability knowledge, or *knowledge-how*. The question of factual or propositional knowledge is more vexed. It is highly likely that psychedelics facilitate genuine psychodynamic insights into previously unconscious or unattended mental states, but the possibility of “placebo insights” (Jopling 2001)—spurious

apparent insights with real therapeutic benefits—must be kept in mind. The only viable solution is for psychedelic-induced apparent epiphanies to be subjected to sober scrutiny during the post-session integration period. I argue that psychedelic experiences also facilitate the acquisition of *new knowledge of old facts*, allowing subjects to experience existing beliefs in more vivid and motivating ways. Finally, I argue that psychedelic experiences have indirect epistemic benefits consequent on their lasting psychological benefits (Letheby 2016).

In chapter nine I turn to the philosophical project of “naturalising spirituality”. I argue that psychedelic research vindicates the claim that there are transformative experiences and practices that (a) can legitimately be called “spiritual” and (b) are compatible with adherence to a naturalistic worldview. The existential transformation afforded by some psychedelic experiences provides a paradigm for naturalistic spirituality: the temporary suspension of our default, self-referential mode of cognition, making available broader perspectives, experiences of connectedness, and feelings of wonder and awe. This has connections to Iris Murdoch’s (1973) notion of “unselfing” (cf. Kähönen 2020). There is considerable convergence between typical features of psychedelic-induced spiritual experiences and themes common to multiple philosophical accounts of naturalistic spirituality (Stone 2012), which provides further support for the general approach.

Finally, in chapter ten, I summarise the discussion and make some suggestions for future research. My account makes several testable predictions; for instance, psychedelic-induced changes to the activity of specific brain networks should correlate with distinct types of ego dissolution experiences, and psychedelic-induced changes to metaphysical beliefs (e.g. about the mind-body relation) should account for relatively little variance in clinical outcomes. In closing, I reflect on the broader significance of my arguments. Psychedelic science is a fast-moving field, and this book will be out of date in some respects before it hits the shelves—which is a good thing. In the coming years we are set to learn more than ever before about these controversial substances, their risks, benefits, and potential applications, and the mechanisms underlying their remarkable effects. The philosophical discussion of psychedelics, in particular, is in its infancy. But I think it is not premature to bet that, when psychedelics cause lasting therapeutic benefits, these benefits are not brought about mainly by the induction of comforting delusions. Psychedelics can, as many have insisted, facilitate genuine insights and spiritual experiences—and this is a claim that even a philosophical naturalist should endorse.