Post-acute psychological effects of classical serotonergic psychedelics: a systematic review and meta-analysis

Author for correspondence:
Simon B. Goldberg,
E-mail: sbgoldberg@wisc.edu

Simon B. Goldberg[1] [iD], Benjamin Shechet[1], Christopher R. Nicholas[2], Chi Wing Ng[1], Geetanjali Deole[1], Zhuofan Chen[1] and Charles L. Raison[3,4]

[1]Department of Counseling Psychology, University of Wisconsin – Madison, Madison, WI, USA; [2]Department of Family Medicine and Community Health, University of Wisconsin – Madison, Madison, WI, USA; [3]School of Human Ecology, University of Wisconsin – Madison, Madison, WI, USA and [4]Usona Institute, Fitchberg, WI, USA

Abstract

Background. Scientific interest in the therapeutic effects of classical psychedelics has increased in the past two decades. The psychological effects of these substances outside the period of acute intoxication have not been fully characterized. This study aimed to: (1) quantify the effects of psilocybin, ayahuasca, and lysergic acid diethylamide (LSD) on psychological outcomes in the post-acute period; (2) test moderators of these effects; and (3) evaluate adverse effects and risk of bias.

Methods. We conducted a systematic review and meta-analysis of experimental studies (single-group pre-post or randomized controlled trials) that involved administration of psilocybin, ayahuasca, or LSD to clinical or non-clinical samples and assessed psychological outcomes $\geqslant 24$ h post-administration. Effects were summarized by study design, timepoint, and outcome domain.

Results. A total of 34 studies (24 unique samples, $n = 549$, mean longest follow-up = 55.34 weeks) were included. Classical psychedelics showed significant within-group pre-post and between-group placebo-controlled effects on a range of outcomes including targeted symptoms within psychiatric samples, negative and positive affect-related measures, social outcomes, and existential/spiritual outcomes, with large between-group effect in these domains (Hedges' gs = 0.84 to 1.08). Moderator tests suggest some effects may be larger in clinical samples. Evidence of effects on big five personality traits and mindfulness was weak. There was no evidence of post-acute adverse effects.

Conclusions. High risk of bias in several domains, heterogeneity across studies, and indications of publication bias for some models highlight the need for careful, large-scale, placebo-controlled randomized trials.

Introduction

Humans have intentionally consumed psychoactive substances for thousands of years (Guerra-Doce, 2015). Psychedelic substances, in particular, figure prominently in indigenous medical and religious practices around the world (Samorini, 2019; Schultes, 1969). Scientific interest during the 1950s and 1960s in the therapeutic potential of both plant-based psychedelics (e.g. psilocybin) and synthetic psychedelics [e.g. lysergic acid diethylamide (LSD)] largely ceased following legislative changes during the 1970s and 1980s (Bonson, 2018). Research has resumed in the past two decades. While early work in this contemporary period focused on pharmacokinetics (e.g. Callaway et al., 1999) or the use of psychedelics as a model for psychiatric conditions (e.g. schizophrenia; Vollenweider, Vollenweider-Scherpenhuyzen, Bäbler, Vogel, and Hell, 1998), a growing number of studies are again evaluating the therapeutic potential of psychedelics (Reiff et al., 2020).

Classical psychedelics are a class of psychoactive substances that share both mode of action (agonism of the 5-HT$_{2A}$ receptor; Carhart-Harris, 2019) and psychoactive effects (marked cognitive, affective, and perceptual changes). Members of this class that have received recent scientific attention include psilocybin, ayahuasca, and LSD (dos Santos, Bouso, Alcázar-Córcoles, & Hallak, 2018). Psilocybin (4-phosphoroyloxy-N,N-dimethyltryptamine) is a naturally occurring plant alkaloid used ritualistically for spiritual and healing purposes by indigenous cultures in Mexico and South America (Guzmán, 2008). Ayahuasca is a plant-based serotonergic psychedelic also used ritualistically by indigenous cultures in South America (McKenna, 2004). The psychoactive effects of ayahuasca are due to N,N-dimethyltryptamine coupled with reversible monoamine oxidase inhibitors (Ott, 1999). LSD is a synthetic psychedelic first synthesized in 1943 by Hofmann (1980) that is both a serotonin and dopamine receptor agonist (Giacomelli, Palmery, Romanelli, Cheng, & Silvestrini, 1998; Preller et al., 2017).

Numerous studies in the 1960s investigated the therapeutic effects of LSD for the treatment of addiction (Krebs & Johansen, 2012) and other clinical applications (e.g. end-of-life distress; Ross, 2018). Research halted as LSD became associated with the countercultural revolution of the late 1960s coupled with concerns regarding its safety (Nutt, King, & Nichols, 2013).

Studies have begun reexamining the therapeutic potential of classical psychedelics for clinical conditions including depression (Carhart-Harris et al., 2016a, 2018a; Palhano-Fontes et al., 2019), anxiety (Gasser et al., 2014; Ross et al., 2016), and substance use (Bogenschutz et al., 2015; Johnson, Garcia-Romeu, Cosimano, & Griffiths, 2014). Often psychedelics are paired with behavioral interventions intended to maximize benefits by enhancing the mental 'set' and physical 'setting' (Carhart-Harris et al., 2018b). Other studies have examined effects in non-clinical samples on measures of well-being, personality, and associated constructs (e.g. mindfulness, spirituality; MacLean, Johnson, and Griffiths, 2011; Soler et al., 2018).

Several systematic reviews have examined the safety and efficacy of psychedelics for both clinical and non-clinical populations. These narrative reviews consistently suggest psychedelics can be safely administered (i.e. adverse effects are minimal and transient) and may reduce depression and anxiety symptoms (Muttoni, Ardissino, & John, 2019), provide psychological benefits in the context of life-threatening disease (Reiche et al., 2018), and induce mystical experiences associated with enduring changes in personality and attitudes (Aday, Mitzkovitz, Bloesch, Davoli, & Davis, 2020). Despite several well-conducted systematic reviews, only two quantitative reviews (i.e. meta-analyses) have characterized the efficacy of psychedelics. Krebs and Johansen (2012) meta-analyzed six randomized controlled trials (RCTs) published between 1966 and 1970 testing LSD for alcoholism, finding LSD substantially reduced substance misuse (odds ratio = 1.96). Goldberg, Pace, Nicholas, Raison, and Hutson (2020) found that psilocybin was associated with large reductions in depression and anxiety across four recent studies (Hedges' gs = 0.82 to 1.47).

The available reviews suggest psychedelics may have therapeutic potential. Yet, a clear quantitative depiction of the breadth of this literature is lacking. A comprehensive meta-analysis would be valuable for characterizing the magnitude and variability (i.e. heterogeneity) of the effect of psychedelics across psychological outcomes, including but not limited to psychiatric symptoms. Such a meta-analysis would be particularly valuable for clarifying effects that have been inconsistent in prior studies (e.g. effects on personality; Barrett, Doss, Sepeda, Pekar, & Griffiths, 2020; MacLean et al., 2011). The small sample size in many primary studies (e.g. mean n = 29.25; Goldberg et al., 2020) also recommends the use of meta-analysis which allows aggregation across studies. Lastly, meta-analysis offers the opportunity to examine whether various study-level features (e.g. psychedelic type, behavioral support) moderate effects.

The current study sought to address this gap in the literature by quantitatively synthesizing psychological effects from experimental studies testing psilocybin, ayahuasca, or LSD. We focus on these three substances due to their shared mechanism of action (5-HT$_{2A}$ receptor agonism) and subjective effects. Other psychoactive compounds that produce partially overlapping effects through partially overlapping mechanisms were not considered [e.g. enactogens such as 3,4-methylenedioxymethamphetamine (MDMA); Reiff et al., 2020]. Given our interest in therapeutic applications, we focus on effects outside of the acute period of intoxication. To provide the most comprehensive depiction, we included studies with either

clinical or non-clinical (i.e. healthy) samples. Likewise, we included both between-group (e.g. RCTs) and within-group (e.g. pre-post) designs. Four study-level characteristics (psychedelic type, clinical sample, presence of behavioral support, percentage female) were examined as moderators. We also assess adverse effects and risk of bias within and between studies.

Method

Protocol and registration

We followed the PRISMA guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009). This meta-analysis was pre-registered through the Open Science Framework (https://osf.io/79y5v/). Upon reviewing the available studies, we made several deviations. First, we restricted our focus to post-acute effects given the acute hallucinogenic effects have been well characterized (e.g. Studerus, Kometer, Hasler, and Vollenweider, 2011) and are less relevant for therapeutic purposes. Second, there were insufficient studies to test moderation by specific clinical condition (e.g. depression *v.* anxiety disorders). Instead, we report results restricted to clinical samples and to samples with depression. Third, no waitlist control conditions were available to compare with placebo-controlled studies. Fourth, we aggregated outcomes into conceptually coherent categories based on measures reported across studies. This led to the addition of some categories (e.g. adverse effects) and exclusion of some that were rarely reported (e.g. substance use).

Eligibility criteria

Eligible studies involved the administration of psilocybin, ayahuasca, or LSD within an experimental setting (i.e. not a naturalistic setting). Studies were required to report at least one psychological outcome. We maintained a broad definition of psychological to include psychiatric symptoms as well as non-clinical measures (e.g. well-being, spirituality). However, measures primarily focused on the acute psychedelic experience itself (e.g. altered states of consciousness; Studerus, Gamma, and Vollenweider, 2010) were excluded. Outcomes were assessed outside of the period of acute intoxication, which we operationalized as ⩾24 h post-administration of the psychedelic, consistent with prior studies (e.g. Schmid et al., 2015). Studies with and without behavioral support were eligible. Both single group (e.g. within-group pre-post) or between-group designs (e.g. placebo-controlled RCT) were eligible. Both clinical and non-clinical samples were eligible. No restriction was placed on language or publication status. Studies were excluded if they were missing data necessary for computing effect sizes. Studies that only reported post-treatment data without a baseline measurement or a relevant control group (e.g. persisting effects at post-treatment for a single-group design; Nicholas et al., 2018) were excluded. Principal investigators of completed clinical trials were contacted regarding available results.

Information sources

We searched six databases including PubMed, CINAHL, PsycINFO, Web of Science, Scopus, and Cochrane. We restricted our search to studies from the contemporary period of psychedelic research (1990 or later). This window captured the period when research on classical psychedelics resumed (e.g. Strassman, Qualls, Uhlenhuth, and Kellner, 1994) but excluded

early research (1950s to 1960s) conducted under sufficiently different methodological standards such that safety and efficacy data may not be interpretable (Bonson, 2018). The search was conducted between October 23rd and 31st, 2019. In addition, we hand searched recent systematic reviews (Aday et al., 2020; Bouso, dos Santos, Alcázar-Córcoles, & Hallak, 2018; dos Santos et al., 2018; Jungaberle et al., 2018; Muttoni et al., 2019; Reiche et al., 2018; Reiff et al., 2020; Schenberg, 2018).

Search

We paired search terms associated with the three psychedelics of interest (e.g. 'psilocybin,' 'ayahuasca,' 'LSD,' 'psychedelic*') with terms related to both clinical (e.g. 'mental disorders,' 'depression,' 'anx*') and non-clinical populations (e.g. 'well-being,' 'quality of life,' 'healthy'). The full search terms for all six databases are shown in online Supplementary Table 1.

Study selection

Two authors independently reviewed each title and/or abstract of potential studies for inclusion. Full texts were reviewed for studies that passed initial screening. Disagreements were discussed with the first author until consensus was reached.

Data collection process

Standardized spreadsheets were developed for study- and effect size-level coding. The first and second authors independently extracted data. Inter-rater reliabilities were good to excellent (i.e. Ks and ICCs ⩾ 0.74; Cicchetti, 1994).

Data items

In addition to data necessary for computing effect sizes (e.g. sample sizes, means, standard deviations), we extracted: (1) study design, (2) psychedelic type and dose and control condition, (3) inclusion criteria, (4) adverse events, (5) post-treatment and follow-up timing, (6) behavioral support, (7) sample age and sex composition, (8) country, (9) and retention. We also extracted data necessary for coding risk of bias with the Cochrane tool (Higgins & Green, 2008). Outcomes were grouped into categories that were intended to be both parsimonious and conceptually coherent. This yielded 14 categories: adverse effects (i.e. symptoms potentially associated with negative drug effects such as psychotic symptoms or mania), targeted symptoms of psychiatric disorders (e.g. alcohol use for samples with alcohol use disorder), depression for samples with depression (as this was the most common psychiatric disorder studied), negative affect-related outcomes (e.g. negative mood, anxiety), positive affect-related outcomes (e.g. joy), social outcomes (e.g. altruism), behavior (e.g. observer-rated behavior change), existential and spiritual outcomes (e.g. death transcendence, lifetime mystical experience), mindfulness, and the big five personality traits (i.e. openness, neuroticism, extraversion, agreeableness, conscientiousness).

Risk of bias in individual studies

Risk of bias was evaluated using the Cochrane tool (Higgins & Green, 2008). Bias was assessed across five domains: selection bias (random sequence generation, allocation concealment), performance bias (blinding of participants and personnel), detection bias (blinding of outcome assessors), attrition bias (incomplete outcome data), and reporting bias (selective reporting). For each study, an evaluation of low, high, or unclear risk of bias was made.

Summary measures

Effect sizes in standardized units were calculated using standard meta-analytic methods (Cooper, Hedges, & Valentine, 2009). Specifically, a within-group pre-post and pre-follow-up Cohen's (1988) d was computed for all studies providing eligible data. The pre-post effect used baseline and the first available data collected post-treatment. To provide the most conservative estimate of effects at follow-up, pre-follow-up effects used data from the last available follow-up. For within-group effects, we assumed a correlation of $r_{xx} = 0.50$ between timepoints (Hoyt & Del Re, 2018). For controlled studies, a between-group effect size was also computed. When pre-post data were available for both the treatment and control conditions, within-group effects were computed for each group separately. Then, the between-group effect was computed as the difference between within-group effects (i.e. Becker's del; Becker, 1988 del). This effect size has the advantage of accounting for baseline data. When within-group effects were not available (e.g. outcomes like persisting effects assessed only at post-treatment; Griffiths, Richards, McCann, and Jesse, 2006), a between-group Cohen's d was computed. To provide the most conservative estimate of controlled effects, we used data from the last available follow-up timepoint. For randomized controlled cross-over designs in which both groups ultimately received the active treatment (e.g. Ross et al., 2016), we used data from the last timepoint prior to cross-over. For within-person RCTs that included multiple dosages (e.g. Bershad, Schepers, Bremmer, Lee, and de Wit, 2019), we compared the placebo condition with the highest dose condition.

In order to decrease the influence of selective reporting bias (Higgins & Green, 2008), we attempted to represent all outcome measures that were assessed. Authors were contacted regarding measures described in the Method section but not included in the Results section. When data remained missing at the time of analysis, we represented effects described in the text as non-significant as $d = 0.00$. Authors were also contacted when adverse effects were not mentioned in the published report.

Synthesis of results

Using standard meta-analytic methods (Cooper et al., 2009), effects were aggregated first within measure [e.g. subscales of the Depression Anxiety and Stress Scale (Lovibond & Lovibond, 1995)] and then within study using the 'MAd' package (Del Re & Hoyt, 2014) in R (R Core Team, 2018). As noted previously, separate analyses examined effects for specific outcome domains. Meta-analytic effect sizes with an associated 95% confidence interval (CI) were computed when at least two studies were available for a specific estimate (Valentine, Pigott, & Rothstein, 2010). Summary effects were converted from Cohen's d to Hedges' g in order to account for small sample bias (Cooper et al., 2009). As appropriate, the sign for each effect was reversed so that a positive g always indicated improvement (e.g. decreased depression, increased well-being). Magnitude was interpreted based on Cohen's (1988) guidelines. Separate aggregate effect size estimates were computed for within-group effects at post-treatment and follow-up and for between-group effects at last available post-treatment assessment. Heterogeneity was characterized using I^2

(i.e. proportion of heterogeneity that is between-study heterogeneity) and interpreted based on Higgins, Thompson, Deeks, & Altman's (2003) guidelines. Random effects models with weighting based on the inverse of the variance of each study's effect size was implemented through the 'metafor' package (Viechtbauer, 2010).

Risk of bias across studies

We assessed publication bias using trim-and-fill analyses in the 'metafor' package. When funnel plot asymmetry was detected, an adjusted effect size was computed with studies imputed to account for asymmetry. Due to the small number of studies in some analyses, which limits statistical power, these tests were considered exploratory. In addition, we calculated the fail-safe *N*s to represent the number of non-significant results that would need to exist to nullify an observed effect (Rosenthal, 1979).

Additional analyses

We tested four study-level characteristics as moderators. These included the psychedelic type (coded as 1 = psilocybin, 0 = LSD or ayahuasca), whether the sample was clinical (i.e. required elevated symptoms of a medical/psychiatric diagnosis for inclusion) or non-clinical (i.e. healthy controls), whether behavioral support was provided (e.g. pre-treatment preparation), and percentage female. Psilocybin was compared with LSD or ayahuasca as the majority of studies investigated psilocybin ($k = 14$). Insufficient studies were available to adequately compare psilocybin with LSD ($k = 4$) and ayahuasca ($k = 6$) separately, or LSD and ayahuasca with each other. We also conducted sensitivity analyses with outliers excluded. There are several methods for identifying outliers in meta-analysis (Viechtbauer & Cheung, 2010). We used the 'find.outliers' function provided by Harrer, Cuijpers, Furukawa, and Ebert (2019) which defines an outlier as a study whose CI does not overlap the omnibus effect CI.

Results

Study selection

Our search produced a total of 14 591 citations. After removing 4540 duplicates, 10 051 unique titles and/or abstracts were reviewed. After applying our exclusion criteria (Fig. 1), we retained 34 studies representing 24 unique samples and 549 participants (see online Supplementary Table 2 for a list of the 34 studies). Studies were published between 2006 and 2020.

Study characteristics

Study-level characteristics are reported in Table 1. Half of the studies used single-group pre-post designs (50.0%) with the remainder being within-group RCTs (i.e. participants received all conditions in random order; 16.7%), or between-group RCTs (33.3%). The majority of studies tested psilocybin (58.3%) with 25.0% testing ayahuasca and 16.7% testing LSD. Dosages of each psychedelic and placebo control conditions are listed in online Supplementary Table 3. Post-test assessment occurred on average at 5.54 weeks post-treatment (s.d. = 6.48, range = 0 to 26.00). Most studies (54.2%) included a follow-up assessment. For studies with a follow-up assessment, last follow-up occurred on average 53.34 weeks (s.d. = 64.25) post-treatment (range = 3

to 234.90). Retention at post-treatment was 94.5% (s.d. = 10.0) and 85.6% (s.d. = 16.9) at follow-up.

Sample sizes were generally small, on average 22.88 participants (s.d. = 17.42, range = 6 to 85). Mean age was 42.13 years old and the samples were 51.5% female. Among the studies that reported race/ethnicity (37.5% of studies), 74.6% were non-Hispanic white or Caucasian. Studies were conducted in the USA (45.8%), Europe (41.7%), and Brazil (12.5%). Approximately half of the studies (45.8%) included participants with clinical conditions. The most common clinical condition was depression ($k = 4$). Other clinical conditions included cancer/life-threatening diseases with comorbid anxiety and/or depression ($k = 3$), alcohol dependence ($k = 1$), smoking ($k = 1$), and AIDS ($k = 1$).

Risk of bias within studies

Risk of bias varied, often based on whether a single-group design was used (online Supplementary Table 4). Single-group designs lacked randomization and other features (e.g. blinding) that increase confidence that effects are associated with the active treatment. Risk of bias also varied across domains (Fig. 2). Blinding of participants and personnel and blinding of outcome assessment were the domains most at risk for bias. Selective reporting bias was commonly rated as unclear due to difficulty in determining whether the reported outcomes were planned.

Results of individual studies

Effect size-level data are reported by study, domain, timepoint, and design in online Supplementary Table 5. The outcome measures included across studies are listed in online Supplementary Table 6 along with their corresponding domain.

Synthesis of results

Adverse effects

Adverse effects were available for 79.2% of studies (online Supplementary Table 3). Among those reporting adverse effects, none reported serious adverse effects (e.g. death, hospitalization). Commonly reported transient adverse effects included headache, anxiety, nausea, and increased blood pressure.

Several studies (29.2%) also included measures of longer-term adverse effects that could be used to quantify the magnitude of these effects (e.g. psychotic symptoms, mania, persisting negative effects; see online Supplementary Table 6). There was no evidence that psychedelics increased risk for adverse effects. In fact, within-group effects suggested decreased adverse effects at post-treatment and follow-up (gs = 0.40 and 0.50, respectively; Table 2). As noted above, a positive effect size indicates a reduction in adverse effects. Heterogeneity was low for within-group pre-post comparisons but moderate to high for within-group pre-follow-up and between-group comparisons.

Within-group effects

Psychedelics showed statistically significant within-group improvements across several outcome domains at both post-treatment and follow-up (Table 2, Fig. 3). Domains showing beneficial effects included targeted symptoms within psychiatric samples, depression within samples with depression, negative affect, positive affect, social outcomes, and existential/spiritual outcomes. Associated effect sizes ranged from gs = 0.44 (positive affect) to 2.06 (depression) and were fairly similar in magnitude
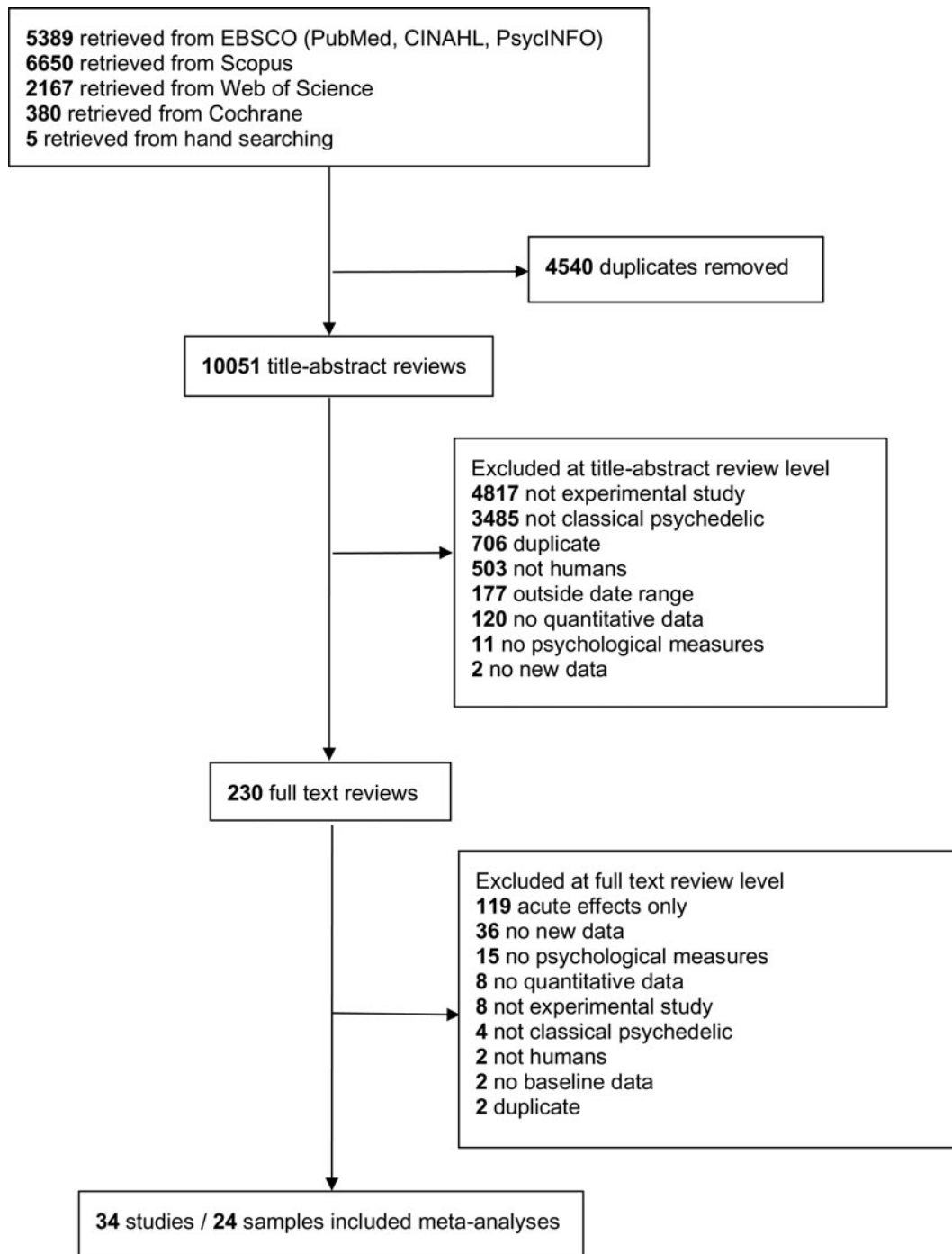
Fig. 1. PRISMA flow diagram.

at post-treatment and follow-up. Psychedelics showed improvements in behavior and mindfulness at post-treatment, although estimates were not available at follow-up. Psychedelics were not associated with changes in big five personality dimensions, with the exception of openness which showed a small increase. Heterogeneity was generally high ($I^2 > 50\%$).

Between-group effects

Moderate to large and statistically significant between-group effects favored psychedelics relative to placebo controls across several outcome domains at longest follow-up. These included

targeted symptoms within psychiatric samples, negative affect, positive affect, social outcomes, behavior, and existential/spiritual outcomes. Effect sizes ranged from gs = 0.84 to 1.16. There was no evidence of between-group effects on personality. Heterogeneity was generally high ($I^2 > 50\%$).

Risk of bias across studies

There was evidence of funnel plot asymmetry (i.e. publication bias) in eight models (Table 2). Statistical significance was not impacted by this adjustment, with one exception [within-group

Table 1. Study characteristics

Study	Design	Sample	Psychedelic	Behav	N_{tx}	N_{cont}	Wk_{post}	Wk_{FU}	Age	% Fem	% White	Ret_{post}	Ret_{FU}	Country
Anderson et al.(2019)	Single group pre-post	AIDS survivors	Psilocybin	Yes	18	NA	NA	NA	NA	NA	NA	100	NA	US
Barrett et al. (2020)	single group pre-post	healthy	Psilocybin	Yes	12	NA	1	4	32.1	58	100	100	NA	US
Bershad et al. (2019)	Within-group RCT	Healthy	LSD	No	20	NA	0	NA	25	60	45	55	NA	US
Bogenschutz et al. (2015)	Single group pre-post	Alcohol dependence	Psilocybin	Yes	10	NA	1	24	40.1	40	30	90	90	US
Carhart-Harris et al. (2012)	Within-group RCT	Healthy	Psilocybin	No	15	NA	2	NA	30.5	13	NA	100	NA	England
Carhart-Harris et al. (2016b)	Within-group RCT	Healthy	LSD	No	20	NA	2	NA	30.9	20	NA	100	NA	England
Carhart-Harris et al. (2018a)	Single group pre-post	Depression	Psilocybin	Yes	20	NA	1	26	44.1	30	75	95	95	England
Gasser et al. (2014)	Between-group RCT	Anxiety + life-threat disease	LSD	Yes	8	4	8	52	51.7	36	NA	92	75	Switzerland
Griffiths et al. (2006)	Between-group RCT	Religious/spiritual healthy	Psilocybin	Yes	15	15	8	56	46	61	NA	100	100	US
Griffiths et al. (2011)	Within-group RCT	Healthy	Psilocybin	Yes	18	NA	4	56	46	56	NA	100	100	US
Griffiths et al. (2016)	Between-group RCT	Cancer + depression/anxiety	Psilocybin	Yes	29	27	5	26	56.3	49	94	91	82	US
Griffiths et al. (2018)	Between-group RCT	Healthy	Psilocybin	Yes	57	28	26	NA	42	60	NA	88	NA	US
Grob et al. (2011)	Between-group RCT	Cancer + anxiety	Psilocybin	Yes	6	6	2	26	47	92	NA	100	67	US
Johnson et al. (2014)	Single group pre-post	Smokers	Psilocybin	Yes	15	NA	15	130.5	51	33	93	100	80	US
Madsen et al. (2020)	Single group pre-post	Healthy	Psilocybin	No	10	NA	12	NA	28.4	40	NA	100	NA	Denmark
Osorio et al. (2015)	Single group pre-post	Depression	Ayahuasca	Yes	6	NA	0	3	44.2	67	NA	100	100	Brazil
Palhano-Fontes et al. (2019)	Between-group RCT	Depression	Ayahuasca	No	17	18	1	NA	42.0	72	59	83	NA	Brazil
Ross et al. (2016)	Between-group RCT	Cancer + anxiety	Psilocybin	Yes	16	15	7	234.9	56.3	62	90	90	45	US
Sampedro et al. (2017)	single group pre-post	healthy	ayahuasca	No	16	NA	8	NA	38.9	38	NA	88	NA	Spain
Sanches et al.(2016)	Single group pre-post	Depression	Ayahuasca	Yes	17	NA	0	3	42.7	82	NA	100	100	Brazil
Schmid & Liechti (2018)	Single group pre-post	Healthy	LSD	No	16	NA	4	52	28.6	50	NA	100	94	Switzerland
Smigielski et al. (2019)	Between-group RCT	Healthy	Psilocybin	Yes	20	20	16	NA	51.7	39	NA	95	NA	Switzerland
Soler et al. (2016)	Single group pre-post	Healthy	Ayahuasca	No	25	NA	0	NA	43.6	56	NA	100	NA	Spain
Soler et al. (2018)	Single group pre-post	General population	Ayahuasca	No	10	NA	4	NA	50	70	NA	100	NA	Spain

Behav, inclusion of behavioral support (e.g. preparation prior to psychedelic administration); N, sample size; tx, treatment; cont, control; Wk_{post}, week of post-treatment assessment; Wk_{FU}, week of follow-up assessment; Fem, female; Ret_{post}, % of sample retained at post-treatment assessment; Ret_{FU}, % of sample retained at follow-up assessment; NA, not available; life-threat disease, life-threatening disease.
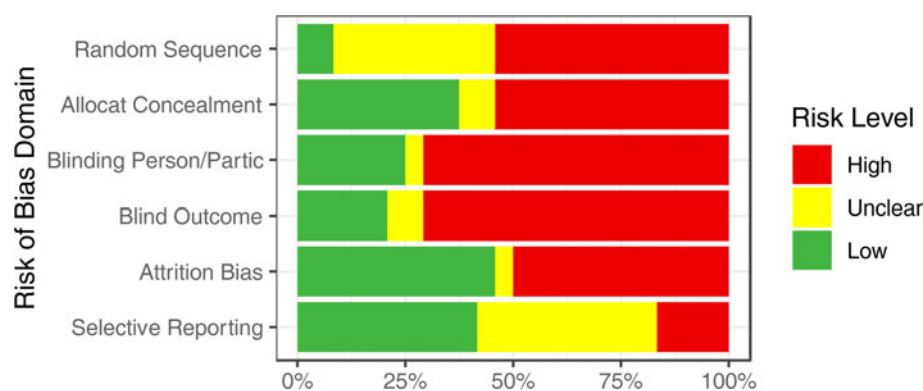
Fig. 2. Cochrane risk of bias assessment. Random sequence, random sequence generation; Allocat Concealment, allocation concealment; Blinding Person/Partic, blinding of personnel and participants; Blind Outcome, blinding of outcome assessment; Attrition bias, incomplete outcome data.

pre-post effect on social outcomes which became non-significant, $g = 0.43$ $(-0.10, 0.97)$]. Fail-safe Ns ranged from 0 to 803. Based on Rosenberg's (2005) guidelines (i.e. fail-safe $N > 5n + 10$, where n = number of published studies), within-group effects on adverse effects, social outcomes, openness, and mindfulness as well as between-group effects on behavior were not robust against publication bias.

Additional analyses

Due to insufficient studies, not all moderators could be tested for all models (see online Supplementary Table 7). Clinical samples were associated with larger improvements for some comparisons in the domains of negative affect, positive affect, adverse effects, existential/spiritual outcomes, and extraversion. Psychedelic type did not moderate effects, with the exception of within-group pre-post effects on mindfulness for which psilocybin produced larger increases. Presence of behavioral support did not moderate effects. Percentage female did not moderate effects, with the exception of within-group pre-follow-up effects on extraversion for which higher percentage female was associated with smaller increases.

Models with outliers removed are reported in online Supplementary Table 8. No significance tests changed as a result of this and effect sizes were similar in magnitude (change in $g \leqslant 0.26$).

Discussion

To our knowledge, this is the first comprehensive meta-analysis of experimental studies testing the post-acute effects of psychedelics.[1] Although based on a relatively small number of studies and participants ($k = 34$ studies and 24 unique samples, $n = 549$), results suggest psychedelics may produce beneficial effects. Most relevant for psychiatric samples, large and statistically significant effects were detected for targeted symptoms ($g = 1.08$) when psychedelics were compared with placebo controls in RCTs. As points of comparison, this effect is on par or larger than that achieved by psychotherapy relative to waitlist (e.g. $d = 0.80$; Wampold and Imel, 2015) and antidepressants relative to placebo (e.g. ds = 0.42 to 0.17; Cipriani et al., 2018). Moreover, this effect appears robust to publication bias and not influenced by outliers. Psychedelics also compared favorably with placebo controls on measures related to negative and positive affect; on

measures of social, behavior, and existential/spiritual outcomes; and on depression in samples with depression (although effect on behavior was not robust to fail-safe N). The superiority over placebo controls supports the possibility of specific effects, however this conclusion is necessarily uncertain given difficultly blinding psychedelics. Within-group effects were similar in magnitude and statistical significance, and support the notion that beneficial effects may persist at follow-up. Although adverse effects were not available for 20.8% studies, effects reported were transient and no serious adverse events occurred. Quantitative assessment of longer-term adverse effects similarly suggests that transient psychological effects do not typically remain elevated during the post-acute period and may even reduce in some instances. Evidence supporting the effects of psychedelics on personality and mindfulness were less compelling and less robust to test of publication bias.

Due to the limited number of studies and variation across studies in design features, we were limited in our ability to test moderators. Nonetheless, it appears that some effects may be larger for clinical samples. Psychedelic type, presence of behavioral support, and percentage female generally did not moderate effects, although confounding with other design characteristics (e.g. amount of behavioral support, clinical sample) makes these null findings tenuous. It does appear that moderate to large reductions in psychiatric symptoms have been achieved in studies testing psilocybin with relatively little behavioral support (e.g. one to three sessions; Carhart-Harris et al. 2018a, b; Grob et al., 2011). Future clinical trials and meta-analyses should clarify the requisite dosage of behavioral support.

Although the most comprehensive quantitative review to date, our study remained limited in sample size and associated statistical power. Indeed, the sample available in the entire literature reviewed ($n = 549$) is considerably smaller than that from large-scale RCTs (e.g. $n = 952$ in Project MATCH; Project Match Research Group, 1998). This highlights the inherent uncertainty in conclusions drawn. An additional complication is the degree to which generalizations can be made from the individuals who chose to participate in the available experimental studies, given psychedelics remain Schedule I substances in most study locations. While selection bias may have produced inflated effect size estimates (e.g. selecting individuals most open to the possibility of change through psychedelic treatments, higher expectancy), some studies included healthy controls with previous use of psychedelics which could have created ceiling effects (i.e. therapeutic effects were achieved at baseline through prior use). A relatively modest amount of racial/ethnic diversity and a lack of reporting on sample race/ethnicity in the available studies is another

---

[1]Since the time of submission, Luoma et al. (2020) published a meta-analysis of nine placebo-controlled RCTs that was restricted to primary outcomes and included trials testing MDMA.

Table 2. Meta-analytic estimates of effects of classical psychedelics across outcome domains

Domain	Comparison	Timepoint	N	K	ES [95% CI]	I^2 [95% CI]	k_{imp}	ES$_{adj}$	FSN
Targeted sx	Within-group	Pre-post	126	10	1.70 [1.16–2.23]	84.19 [65.56–95.68]	0	1.70 [1.16–2.23]	803
Targeted sx	Within-group	Pre-FU	142	9	1.37 [0.95–1.80]	79.52 [53.83–94.48]	0	1.37 [0.95–1.80]	657
Targeted sx	Between-group	Pre-post	67	5	1.08 [0.74–1.43]	0.00 [0.00–93.15]	0	1.08 [0.74–1.43]	69
Depression	Within-group	Pre-post	49	4	2.06 [1.41–2.71]	62.58 [0.00–97.86]	1	1.83 [1.10–2.55]	152
Depression	Within-group	Pre-FU	42	3	1.57 [0.90–2.24]	63.92 [0.00–99.09]	0	1.57 [0.90–2.24]	67
Depression	Between-group	Pre-post	NA	NA	NA	NA	NA	NA	NA
Neg affect	Within-group	Pre-post	182	14	1.05 [0.60–1.49]	89.09 [79.11–96.21]	0	1.05 [0.60–1.49]	818
Neg affect	Within-group	Pre-FU	157	10	0.99 [0.54–1.43]	87.17 [72.29–96.17]	0	0.99 [0.54–1.43]	509
Neg affect	Between-group	Pre-post	82	6	0.87 [0.46–1.28]	43.12 [0.00–91.83]	2	0.67 [0.27–1.08]	66
Pos affect	Within-group	Pre-post	139	7	0.44 [0.15–0.73]	70.83 [24.71–94.67]	0	0.44 [0.15–0.73]	71
Pos affect	Within-group	Pre-FU	115	5	0.47 [0.13–0.82]	73.08 [26.51–96.83]	0	0.47 [0.13–0.82]	47
Pos affect	Between-group	Pre-post	163	8	0.89 [0.52–1.25]	71.45 [30.78–93.59]	1	0.76 [0.33–1.19]	214
Adverse	Within-group	Pre-post	23	2	0.40 [0.08–0.71]	0.00 [0.00–99.89]	NA	NA	2[a]
Adverse	Within-group	Pre-FU	23	2	0.50 [0.05–0.95]	42.78 [0.00–99.94]	NA	NA	5[a]
Adverse	Between-group	Pre-post	96	4	0.46 [−0.42 to 1.33]	92.55 [74.26–99.52]	0	0.46 [−0.42 to 1.33]	9
Social	Within-group	Pre-post	90	4	0.81 [0.36–1.25]	49.16 [0.00–98.69]	2	0.43 [−0.10 to 0.97]	25[a]
Social	Within-group	Pre-FU	41	3	0.53 [0.16–0.90]	24.88 [0.00–97.64]	0	0.53 [0.16–0.90]	10[a]
Social	Between-group	Pre-post	95	4	1.13 [0.76–1.51]	35.87 [0.00–96.61]	0	1.13 [0.76–1.51]	84
Behavior	Within-group	Pre-post	87	3	1.47 [0.90–2.04]	33.57 [0.00–99.89]	NA	NA	28
Behavior	Within-group	Pre-FU	NA	NA	NA	NA	NA	NA	NA
Behavior	Between-group	Pre-post	76	3	1.16 [0.78–1.53]	0.00 [0.00–98.23]	NA	NA	24[a]
Exist/spirit	Within-group	Pre-post	159	8	0.56 [0.35–0.76]	50.73 [0.00–88.07]	2	0.48 [0.28–0.68]	170
Exist/spirit	Within-group	Pre-FU	145	7	0.52 [0.27–0.76]	62.51 [13.36–89.69]	0	0.52 [0.27–0.76]	111
Exist/spirit	Between-group	Pre-post	120	5	0.84 [0.53–1.16]	50.90 [0.00–94.23]	2	0.64 [0.28–1.00]	103
Openness	Within-group	Pre-post	124	5	0.21 [0.04–0.38]	0.00 [0.00–37.54]	0	0.21 [0.04–0.38]	5[a]
Openness	Within-group	Pre-FU	92	5	0.20 [0.00–0.40]	0.00 [0.00–77.94]	0	0.20 [0.00–0.40]	2[a]
Openness	Between-group	Pre-post	85	3	0.07 [−0.22 to 0.36]	0.00 [0.00–78.27]	0	0.07 [−0.22 to 0.36]	0
Neuroticism	Within-group	Pre-post	124	5	0.06 [−0.11 to 0.23]	0.00 [0.00–88.92]	0	0.06 [−0.11 to 0.23]	0
Neuroticism	Within-group	Pre-FU	92	5	0.16 [−0.05 to 0.36]	4.88 [0.00–87.32]	0	0.16 [−0.05 to 0.36]	1
Neuroticism	Between-group	Pre-post	85	3	0.00 [−0.29 to 0.29]	0.00 [0.00–0.00]	0	0.00 [−0.29 to 0.29]	0
Extraversion	Within-group	Pre-post	124	5	0.04 [−0.13 to 0.21]	0.00 [0.00–0.00]	3	0.01 [−0.13 to 0.15]	0
Extraversion	Within-group	Pre-FU	92	5	0.18 [−0.09 to 0.44]	41.19 [0.00–93.22]	0	0.18 [−0.09 to 0.44]	1
Extraversion	Between-group	Pre-Post	85	3	0.00 [−0.29 to 0.29]	0.00 [0.00–0.00]	0	0.00 [−0.29 to 0.29]	0
Agreeable	Within-group	Pre-post	124	5	0.05 [−0.12 to 0.23]	0.00 [0.00–53.53]	0	0.05 [−0.12 to 0.23]	0
Agreeable	Within-group	Pre-FU	92	5	−0.02 [−0.21 to 0.18]	0.00 [0.00–59.10]	0	−0.02 [−0.21 to 0.18]	0
Agreeable	Between-group	Pre-post	85	3	0.09 [−0.20 to 0.38]	0.00 [0.00–87.19]	0	0.09 [−0.20 to 0.38]	0
Conscientious	Within-group	Pre-post	124	5	0.02 [−0.15 to 0.19]	0.00 [0.00–84.35]	1	−0.01 [−0.17 to 0.15]	0
Conscientious	Within-group	Pre-FU	92	5	0.17 [−0.03 to 0.36]	0.00 [0.00–85.92]	0	0.17 [−0.03 to 0.36]	2
Conscientious	Between-group	Pre-post	85	3	0.00 [−0.29 to 0.29]	0.00 [0.00–0.00]	0	0.00 [−0.29 to 0.29]	0
Mindfulness	Within-group	Pre-post	81	5	0.45 [0.14–0.77]	60.84 [0.00–95.86]	0	0.45 [0.14–0.77]	33[a]
Mindfulness	Within-group	Pre-FU	NA	NA	NA	NA	NA	NA	NA
Mindfulness	Between-group	Pre-post	NA	NA	NA	NA	NA	NA	NA

N, sample size; K, number of studies; CI, confidence interval; k_{imp}, number of studies imputed for trim-and-fill adjustment; ES$_{adj}$, trim-and-fill adjusted effect size; FSN, fail-safe N; Targeted sx, targeted symptoms within psychiatric samples; Depression, depression outcomes restricted to samples with depression; Neg, negative; Pos, positive; Exist/spirit, existential/spiritual; ES, effect size in Hedges' g units; FU, follow-up; NA, not available. Estimates based on $k = 1$ not included.
[a]statistically significant result not robust to publication bias based on Rosenberg's (2005) guidelines (i.e. fail-safe $N > 5n + 10$, where n = number of published studies).
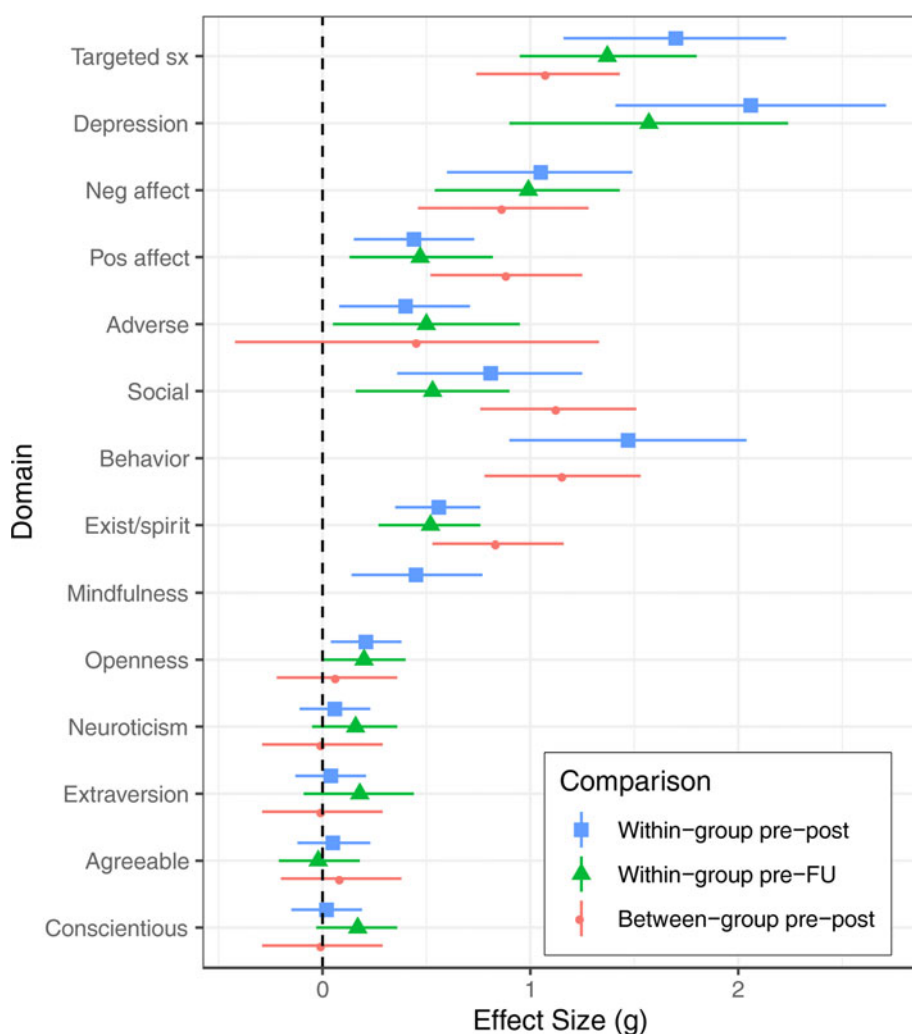
Fig. 3. Forest plots displaying effects of classical psychedelics across psychological outcome domains. Each point represents an effect size estimates (Hedges' g units) and a corresponding 95% confidence interval. Targeted sx, targeted symptoms within psychiatric samples; Depression, depression outcomes restricted to samples with depression; Neg, negative; Pos, positive; Exist/spirit, existential/spiritual; ES, effect size in Hedges' g units; FU, follow-up; NA, not available.

important limitation that must be addressed (Michaels, Purdon, Collins, & Williams, 2018). While we attempted to aggregate effects in conceptually coherent ways, there remained methodological heterogeneity (e.g. psychedelic dose, provision of behavioral support) that was either not modeled or tested in underpowered ways. This makes it impossible to provide recommendations regarding the specific treatment characteristics most strongly linked to beneficial effects. Similarly, although results generally did not change when accounting for publication bias, trim-and-fill analyses were also likely underpowered.

A broader potentially more pernicious limitation is risk of bias within the available studies. As noted, obviously psychoactive substances may be particularly difficult to adequately double blind. However, several studies included features that may increase the strength of the placebo condition (e.g. using methylphenidate or other psychoactive agents, making specific treatment conditions and study aims ambiguous; Griffiths et al., 2006). Two potential sources of bias that would be relatively straightforward to address are risks associated with attrition and selective reporting. None of the included studies explicitly used an intention-to-treat analysis, although this would be a straightforward way to address attrition bias. Of note, studies rated here as low on attrition bias generally had no attrition. Selective reporting could be reduced through more consistent pre-registration of study hypotheses. While several included studies were pre-registered (e.g. clinicaltrials.gov),

many were not, making it difficult to ascertain the degree to which the reported outcomes were specified *a priori v.* drawn from a larger number of unpublished outcomes (i.e. increasing risk for opportunistic bias; DeCoster, Sparks, Sparks, Sparks, and Sparks, 2015). It did not appear that any of the included studies published their hypotheses using the Open Science Framework or similar platforms (e.g. AsPredicted.Org). While perhaps unsurprising given these platforms are relatively new (Foster & Deardorff, 2017) and some contemporary research on psychedelics has been exploratory in nature and may not have had *a priori* hypotheses, explicit pre-registration of study hypotheses and analysis plans could help reduce selective reporting bias and increase confidence in this body of literature.

These limitations notwithstanding, the current study joins the two previous meta-analyses (Goldberg et al., 2020; Krebs & Johansen, 2012) suggesting that psychedelics are a class of substances worthy of further exploration.[2] Careful, large-scale, placebo-controlled RCTs are especially needed to clarify the empirical status for specific clinical conditions (e.g. depression) as well as for non-clinical applications. Particularly promising applications may include the use of psilocybin for the treatment of anxiety and depression (Goldberg et al., 2020), although

---

[2]Since the time of submission, Romeo et al. (2020) published a meta-analysis of eight studies focused on the effects of classical psychedelics on depression symptoms.

ayahuasca and LSD may also prove beneficial for these indications. While based on only one study each in the contemporary period, the use of psilocybin for smoking cessation and LSD for alcohol use are also promising avenues for future exploration, given the prevalence, health burden, and recalcitrance associated with both nicotine and alcohol use disorders. Future studies could pursue the pairing of psychedelics with behavioral interventions and non-psychotherapeutic approaches (e.g. meditation retreats; Smigielski et al., 2019) to enhance well-being and support flourishing in both clinical and non-clinical samples.

However, it is crucial that future work investigating clinical and non-clinical applications of psychedelics carefully evaluate adverse effects. While we found no clear evidence of persistent adverse effects, many of the included studies excluded individuals with personal or family histories of psychiatric conditions (e.g. bipolar disorder, psychotic disorders). Future studies using alternative designs (e.g. naturalistic and population-based surveys, case reports); extending long-term follow-up to measure protracted effects and naturalistic use in trial participants; and examining safety in previously excluded samples (e.g. contraindicated family histories; personality disorder) may help clarify potential risks.

References

Aday, J. S., Mitzkovitz, C. M., Bloesch, E. K., Davoli, C. C., & Davis, A. K. (2020). Long-term effects of psychedelic drugs: A systematic review. *Neuroscience & Biobehavioral Reviews*, 113, 179–189. doi: 10.1016/j.neubiorev.2020.03.017.

Anderson, B., Danforth, A., Daroff, R., Stauffer, C., Diller, J, & Woolley, J. (2019). Psilocybin-assisted group therapy for demoralization in long-term AIDS survivors. *Biological Psychiatry*, 85, S182.

Barrett, F. S., Doss, M. K., Sepeda, N. D., Pekar, J. J., & Griffiths, R. R. (2020). Emotions and brain function are altered up to one month after a single high dose of psilocybin. *Scientific Reports*, 10(1), 1–14. doi: 10.1038/s41598-020-59282-y.

Becker, B. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41, 257–278.

Bershad, A. K., Schepers, S. T., Bremmer, M. P., Lee, R., & de Wit, H. (2019). Acute subjective and behavioral effects of microdoses of lysergic acid diethylamide in healthy human volunteers. *Biological Psychiatry*, 86(10), 792–800. doi: 10.1016/j.biopsych.2019.05.019.

Bogenschutz, M. P., Forcehimes, A. A., Pommy, J. A., Wilcox, C. E., Barbosa, P. C. R., & Strassman, R. J. (2015). Psilocybin-assisted treatment for alcohol dependence: A proof-of- concept study. *Journal of Psychopharmacology*, 29 (3), 289–299. doi: 10.1177/0269881114565144.

Bonson, K. R. (2018). Regulation of human research with LSD in the United States (1949–1987). *Psychopharmacology*, 235(2), 591–604. doi: 10.1007/s00213-017-4777-4.

Bouso, J. C., dos Santos, R. G., Alcázar-Córcoles, MÁ, & Hallak, J. E. (2018). Serotonergic psychedelics and personality: A systematic review of contemporary research. *Neuroscience & Biobehavioral Reviews*, 87, 118–132. doi: 10.1016/j.neubiorev.2018.02.004.

Callaway, J. C., McKenna, D. J., Grob, C. S., Brito, G. S., Raymon, L. P., Poland, R. E., … Mash, D. C. (1999). Pharmacokinetics of Hoasca alkaloids in healthy humans. *Journal of Ethnopharmacology*, 65(3), 243–256.

Carhart-Harris, R. L. (2019). How do psychedelics work? *Current Opinion in Psychiatry*, 32(1), 16–21. doi: 10.1097/YCO.0000000000000467.

Carhart-Harris, R. L., Bolstridge, M., Day, C. M. J., Rucker, J., Watts, R., Erritzoe, D. E., … Nutt, D. J. (2018a). Psilocybin with psychological support for treatment-resistant depression: Six-month follow-up. *Psychopharmacology*, 235(2), 399–408. doi: 10.1007/s00213-017-4771-x.

Carhart-Harris, R. L., Bolstridge, M., Rucker, J., Day, C. M., Erritzoe, D., Kaelen, M., … Nutt, D. J. (2016a). Psilocybin with psychological support for treatment-resistant depression: An open-label feasibility study. *The Lancet Psychiatry*, 3(7), 619–627. doi: 10.1016/ S2215-0366(16)30065-7.

Carhart-Harris, R. L., Kaelen, M., Bolstridge, M., Williams, T. M., Williams, L. T., Underwood, R., … Nutt, D. J. (2016b). The paradoxical psychological effects of lysergic acid diethylamide (LSD). *Psychological Medicine*, 46(7), 1379–1390. doi: 10.1017/S0033291715002901.

Carhart-Harris, R. L., Leech, R., Williams, T. M., Erritzoe, D., Abbasi, N., Bargiotas, T., … Wise, R. G. (2012). Implications for psychedelic-assisted psychotherapy: Functional magnetic resonance imaging study with psilocybin. *The British Journal of Psychiatry*, 200(3), 238–244. doi: 10.1192/bjp.bp.111.103309.

Carhart-Harris, R. L., Roseman, L., Haijen, E., Erritzoe, D., Watts, R., Branchi, I., & Kaelen, M. (2018b). Psychedelics and the essential importance of context. *Journal of Psychopharmacology*, 32(7), 725–731. doi: 10.1177/0269881118754710.

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290.

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., … Geddes, J. R. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: A systematic review and network meta-analysis. *The Lancet*, 391(10128), 1357–1366. doi: 10.1016/ S0140-6736(17)32802-7.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H. M., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage Foundation.

DeCoster, J., Sparks, E. A., Sparks, J. C., Sparks, G. G., & Sparks, C. W. (2015). Opportunistic biases: Their origins, effects, and an integrated solution. *American Psychologist*, 70(6), 499–514. doi: 10.1037/a0039191.

Del Re, A. C., & Hoyt, W. T. (2014). MAd: Meta-analysis with mean differences. R package version 0.8–2, http://CRAN.R-project.org/package=MAd.

dos Santos, R. G., Bouso, J. C., Alcázar-Córcoles, MÁ, & Hallak, J. E. (2018). Efficacy, tolerability, and safety of serotonergic psychedelics for the management of mood, anxiety, and substance-use disorders: A systematic review of systematic reviews. *Expert Review of Clinical Pharmacology*, 11(9), 889–902. doi: 10.1080/17512433.2018.1511424.

Foster, E. D., & Deardorff, A. (2017). Open science framework (OSF). *Journal of the Medical Library Association*, 105(2), 203–206. doi: 10.5195/jmla.2017.88.

Gasser, P., Holstein, D., Michel, Y., Doblin, R., Yazar-Klosinski, B., Passie, T., & Brenneisen, R. (2014). Safety and efficacy of lysergic acid diethylamide-assisted psychotherapy for anxiety associated with life-threatening diseases. *The Journal of Nervous and Mental Disease*, 202(7), 513–520. doi: 10.1097/NMD.0000000000000113.

Giacomelli, S., Palmery, M., Romanelli, L., Cheng, C. Y., & Silvestrini, B. (1998). Lysergic acid diethylamide (LSD) is a partial agonist of D2 dopaminergic receptors and it potentiates dopamine-mediated prolactin secretion in lactotrophs in vitro. *Life Sciences*, 63(3), 215–222.

Goldberg, S. B., Pace, B. T., Nicholas, C. R., Raison, C. L., & Hutson, P. R. (2020). The experimental effects of psilocybin on symptoms of anxiety and depression: A meta- analysis. *Psychiatry Research*, 284, 1–4. doi: 10.1016/j.psychres.2020.112749.

Griffiths, R. R., Johnson, M. W., Carducci, M. A., Umbricht, A., Richards, W. A., Richards, B. D., … Klinedinst, M. A. (2016). Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial. *Journal of Psychopharmacology*, *30*(12), 1181–1197. doi: 10.1177/0269881116675513.

Griffiths, R. R., Johnson, M. W., Richards, W. A., Richards, B. D., Jesse, R., MacLean, K. A., … Klinedinst, M. A. (2018). Psilocybin-occasioned mystical-type experience in combination with meditation and other spiritual practices produces enduring positive changes in psychological functioning and in trait measures of prosocial attitudes and behaviors. *Journal of Psychopharmacology*, *32*(1), 49–69. doi: 10.1177/0269881117731279.

Griffiths, R. R., Johnson, M. W., Richards, W. A., Richards, B. D., McCann, U., & Jesse, R. (2011). Psilocybin occasioned mystical-type experiences: Immediate and persisting dose- related effects. *Psychopharmacology*, *218*(4), 649–665. doi: 10.1007/s00213-011-2358-5.

Griffiths, R. R., Richards, W. A., McCann, U., & Jesse, R. (2006). Psilocybin can occasion mystical-type experiences having substantial and sustained personal meaning and spiritual significance. *Psychopharmacology*, *187*(3), 268–283. doi: 10.1007/s00213-006-0457-5.

Grob, C.S., Danforth, A.L., Chopra, G.S., Hagerty, M., McKay, C.R., Halberstadt, A.L, & Greer, G.R. (2011). Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer. *Archives of General Psychiatry*, *68*(1), 71–78.

Guerra-Doce, E. (2015). Psychoactive substances in prehistoric times: Examining the archaeological evidence. *Time and Mind*, *8*(1), 91–112. doi: 10.1080/1751696X.2014.993244.

Guzmán, G. (2008). Hallucinogenic mushrooms in Mexico: An overview. *Economic Botany*, *62*(3), 404–412.

Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). Doing Meta-Analysis in R: A Hands-on Guide. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/.

Higgins, J. P. T., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions*. London: John Wiley & Sons.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560.

Hofmann, A. (1980). *LSD: My problem child*. New York: McGraw-Hill.

Hoyt, W. T., & Del Re, A. C. (2018). Effect size calculation in meta-analyses of psychotherapy outcome research. *Psychotherapy Research*, *28*, 379–388. doi: 10.1080/10503307.2017.1405171.

Johnson, M. W., Garcia-Romeu, A., Cosimano, M. P., & Griffiths, R. R. (2014). Pilot study of the 5-HT2AR agonist psilocybin in the treatment of tobacco addiction. *Journal of Psychopharmacology*, *28*(11), 983–992. doi: 10.1177/0269881114548296.

Jungaberle, H., Thal, S., Zeuch, A., Rougemont-Bücking, A., von Heyden, M., Aicher, H., & Scheidegger, M. (2018). Positive psychology in the investigation of psychedelics and entactogens: A critical review. *Neuropharmacology*, *142*, 179–199. doi: 10.1016/j.neuropharm.2018.06.034.

Krebs, T. S., & Johansen, PØ (2012). Lysergic acid diethylamide (LSD) for alcoholism: Meta- analysis of randomized controlled trials. *Journal of Psychopharmacology*, *26*(7), 994–1002. doi: 10.1177/0269881112439253.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd. ed.). Sydney, Australia: Psychology Foundation.

Luoma, J. B., Chwyl, C., Bathje, G. J., Davis, A. K., & Lancelotta, R. (2020). A meta-analysis of placebo-controlled trials of psychedelic-assisted therapy. *Journal of Psychoactive Drugs*, *52*(4), 289–299. doi:10.1080/02791072.2020.1769878.

MacLean, K. A., Johnson, M. W., & Griffiths, R. R. (2011). Mystical experiences occasioned by the hallucinogen psilocybin lead to increases in the personality domain of openness. *Journal of Psychopharmacology*, *25*(11), 1453–1461. doi: 10.1177/0269881111420188.

Madsen, M.K., Fisher, P.M., Stenbæk, D.S., Kristiansen, S., Burmester, D., Lehel, S., & Knudsen, G.M. (2020). A single psilocybin dose is associated with long-term increased mindfulness, preceded by a proportional change in neocortical 5-HT2A receptor binding. *European Neuropsychopharmacology*, *33*, 71–80.

McKenna, D. J. (2004). Clinical investigations of the therapeutic potential of ayahuasca: Rationale and regulatory challenges. *Pharmacology & Therapeutics*, *102*(2), 111–129.

Michaels, T. I., Purdon, J., Collins, A., & Williams, M. T. (2018). Inclusion of people of color in psychedelic-assisted psychotherapy: A review of the literature. *BMC Psychiatry*, *18*(1), 245. doi: 10.1186/s12888-018-1824-6.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., and the PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269.

Muttoni, S., Ardissino, M., & John, C. (2019). Classical psychedelics for the treatment of depression and anxiety: A systematic review. *Journal of Affective Disorders*, *258*, 11–24. doi: 10.1016/j.jad.2019.07.076.

Nicholas, C. R., Henriquez, K. M., Gassman, M. C., Cooper, K. M., Muller, D., Hetzel, S., … Hutson, P. R. (2018). High dose psilocybin is associated with positive subjective effects in healthy volunteers. *Journal of Psychopharmacology*, *32*(7), 770–778. doi: 10.1177/0269881118780713.

Nutt, D. J., King, L. A., & Nichols, D. E. (2013). Effects of schedule I drug laws on neuroscience research and treatment innovation. *Nature Reviews Neuroscience*, *14*(8), 577–585.

Ott, J. (1999). Pharmahuasca: Human pharmacology of oral DMT plus harmine. *Journal of Psychoactive Drugs*, *31*(2), 171–177.

Palhano-Fontes, F., Barreto, D., Onias, H., Andrade, K. C., Novaes, M. M., Pessoa, J. A., … Tófoli, L. F. (2019). Rapid antidepressant effects of the psychedelic ayahuasca in treatment-resistant depression: A randomized placebo-controlled trial. *Psychological Medicine*, *49*(4), 655–663. doi: 10.1017/S0033291718001356.

Preller, K. H., Herdener, M., Pokorny, T., Planzer, A., Kraehenmann, R., Stämpfli, P., … Vollenweider, F. X. (2017). The fabric of meaning and subjective effects in LSD-induced states depend on serotonin 2A receptor activation. *Current Biology*, *27*(3), 451–457. doi: 10.1016/j.cub.2016.12.030.

Project MATCH Research Group (1998). Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. *Alcoholism: Clinical and Experimental Research*, *22*(6), 1300–1311.

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org/.

Reiche, S., Hermle, L., Gutwinski, S., Jungaberle, H., Gasser, P., & Majić, T. (2018). Serotonergic hallucinogens in the treatment of anxiety and depression in patients suffering from a life-threatening disease: A systematic review. *Progress in Neuropsychopharmacology & Biological Psychiatry*, *81*, 1–10. doi: 10.1016/j.pnpbp.2017.09.012.

Reiff, C. M., Richman, E. E., Nemeroff, C. B., Carpenter, L. L., Widge, A. S., & Rodriguez, C. I., … & Work Group on Biomarkers and Novel Treatments, a Division of the American Psychiatric Association Council of Research. (2020). Psychedelics and psychedelic- assisted psychotherapy. *American Journal of Psychiatry*, *177*, 391–410. doi: 10.1176/appi.ajp.2019.19010035.

Romeo, B., Karila, L., Martelli, C., & Benyamina, A. (2020). Efficacy of psychedelic treatments on depressive symptoms: A meta-analysis. *Journal of Psychopharmacology*, *34*(10), 1079–1085. doi:10.1177/0269881120919957.

Rosenberg, M. S. (2005). The file-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution*, *59*(2), 464–468.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641.

Ross, S. (2018). Therapeutic use of classic psychedelics to treat cancer-related psychiatric distress. *International Review of Psychiatry*, *30*(4), 317–330. doi: 10.1080/09540261.2018.1482261.

Ross, S., Bossis, A., Guss, J., Agin-Liebes, G., Malone, T., Cohen, B., … Schmidt, B. L. (2016). Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: A randomized controlled trial. *Journal of Psychopharmacology*, *30*(12), 1165–1180. doi: 10.1177/0269881116675512.

Samorini, G. (2019). The oldest archeological data evidencing the relationship of Homo sapiens with psychoactive plants: A worldwide overview. *Journal of Psychedelic Studies*, *3*, 63–80. doi: 10.1556/2054.2019.008.

Sampedro, .F, de La Fuente Revenga, M., Valle, M., Roberto, N., Domínguez-Clavé, E., Elices, M., & Friedlander, P. (2017). Assessing the psychedelic "after-glow" in ayahuasca users: post-acute neurometabolic and functional

connectivity changes are associated with enhanced mindfulness capacities. *International Journal of Neuropsychopharmacology*, 20(9), 698–711.

Sanches, R.F., de Lima Osório, F., dos Santos, .G.R, Macedo, L.R., Maia-de-Oliveira, J.P., Wichert-Ana, L., & Hallak, J.E. (2016). Antidepressant effects of a single dose of ayahuasca in patients with recurrent depression: a SPECT study. *Journal of Clinical Psychopharmacology*, 36(1), 77–81.

Schenberg, E. E. (2018). Psychedelic-assisted psychotherapy: A paradigm shift in psychiatric research and development. *Frontiers in Pharmacology*, 9, 733. doi: 10.3389/fphar.2018.00733.

Schmid, Y., Enzler, F., Gasser, P., Grouzmann, E., Preller, K. H., Vollenweider, F. X., … Liechti, M. E. (2015). Acute effects of lysergic acid diethylamide in healthy subjects. *Biological Psychiatry*, 78(8), 544–553. doi: 10.1016/j.biopsych.2014.11.015.

Schmid, Y., & Liechti, M.E.. (2018). Long-lasting subjective effects of LSD in normal subjects. *Psychopharmacology*, 235(2), 535–545.

Schultes, R. E. (1969). Hallucinogens of plant origin. *Science (New York, N.Y.)*, 163(3864), 245–254.

Smigielski, L., Scheidegger, M., Kometer, M., & Vollenweider, F. X. (2019). Psilocybin- assisted mindfulness training modulates self-consciousness and brain default mode network connectivity with lasting effects. *NeuroImage*, 196, 207–215. doi: 10.1016/j.neuroimage.2019.04.009.

Soler, J., Elices, M., Dominguez-Clavé, E., Pascual, J. C., Feilding, A., Navarro-Gil, M., … Riba, J. (2018). Four weekly ayahuasca sessions lead to increases in "acceptance" capacities: A comparison study with a standard 8-week mindfulness training program. *Frontiers in Pharmacology*, 9, 224. doi: 10.3389/fphar.2018.00224.

Soler, J., Elices, M., Franquesa, A., Barker, S., Friedlander, P., Feilding, A., … Riba, J. (2016). Exploring the therapeutic potential of Ayahuasca: acute

intake increases mindfulness-related capacities. *Psychopharmacology*, 233(5), 823–829. doi:10.1007/s00213-015-4162-0.

Strassman, R. J., Qualls, C. R., Uhlenhuth, E. H., & Kellner, R. (1994). Dose-response study of N, N-dimethyltryptamine in humans: II. Subjective effects and preliminary results of a new rating scale. *Archives of General Psychiatry*, 51(2), 98–108.

Studerus, E., Gamma, A., & Vollenweider, F. X. (2010). Psychometric evaluation of the altered states of consciousness rating scale (OAV). *PloS One*, 5(8), e12412. doi:10.1371/journal.pone.0012412.

Studerus, E., Kometer, M., Hasler, F., & Vollenweider, F. X. (2011). Acute, sub-acute and long- term subjective effects of psilocybin in healthy humans: A pooled analysis of experimental studies. *Journal of Psychopharmacology*, 25(11), 1434–1452. doi: 10.1177/0269881110382466.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, 35(2), 215–247. doi: 10.3102/1076998609346961.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–49.

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta- analysis. *Research Synthesis Methods*, 1(2), 112–125. doi: 10.1002/jrsm.11.

Vollenweider, F. X., Vollenweider-Scherpenhuyzen, M. F., Bäbler, A., Vogel, H., & Hell, D. (1998). Psilocybin induces schizophrenia-like psychosis in humans via a serotonin-2 agonist action. *NeuroReport*, 9(17), 3897–3902.

Wampold, B., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). New York: Routledge.