



Published in final edited form as:

*J Psychoactive Drugs*. 2020 ; 52(4): 289–299. doi:10.1080/02791072.2020.1769878.

## A Meta-Analysis of Placebo-Controlled Trials of Psychedelic-Assisted Therapy

Jason B. Luoma<sup>1,\*</sup>, Christina Chwyl<sup>1</sup>, Geoff J. Bathje<sup>2</sup>, Alan K. Davis<sup>3,4</sup>, Rafael Lancelotta<sup>5</sup>

<sup>1</sup>Portland Psychotherapy Clinic, Research, & Training Center

<sup>2</sup>Adler University, 17 N Dearborn St, Chicago, IL 60602

<sup>3</sup>College of Social Work, The Ohio State University, 1947 College Rd N, Columbus, OH 43210 USA

<sup>4</sup>Behavioral Pharmacology Research Unit, Department of Psychiatry, Johns Hopkins School of Medicine, 5510 Nathan Shock Drive, Baltimore, MD 21224 USA

<sup>5</sup>Innate Path Psychotherapy Clinic, 1445 Holland St, Lakewood, CO 80215 USA

### Abstract

After a two-decade hiatus in which research on psychedelics was essentially halted, placebo-controlled clinical trials of psychedelic-assisted therapy for mental health conditions have begun to be published. We identified nine randomized, placebo-controlled clinical trials of psychedelic-assisted therapy published since 1994. Studies examined psilocybin, LSD (lysergic acid diethylamide), ayahuasca (which contains a combination of N,N-dimethyltryptamine and harmala monoamine oxidase inhibitor alkaloids), and MDMA (3,4-methylenedioxymethamphetamine). We compared the standardized mean difference between the experimental and placebo control group at the primary endpoint. Results indicated a significant mean between-groups effect size of 1.21 (Hedges *g*), which is larger than the typical effect size found in trials of psychopharmacological or psychotherapy interventions. For the three studies that maintained a placebo control through a follow-up assessment, effects were generally maintained at follow-up. Overall, analyses support the efficacy of psychedelic-assisted therapy across four mental health conditions—post-traumatic stress disorder, anxiety/depression associated with a life-threatening illness, unipolar depression, and social anxiety among autistic adults. While study quality was high, we identify several areas for improvement regarding the conduct and reporting of trials. Larger trials with more diverse samples are needed to examine possible moderators and mediators of effects, and to establish whether effects are maintained over time.

---

\* jbluoma@portlandpsychotherapy.com.

#### Declaration of Interest Statement

AKD and RL are members of the board of directors at Source Research Foundation, a non-profit educational grant-making organization providing financial assistance to students who conduct research related to psychedelics. This organization had no input into the present study. The other authors have no conflicts of interest to declare.

## Introduction

Clinical research on classical psychedelics, mostly LSD (lysergic acid diethylamide) and psilocybin, was once widespread in the 1950s and 1960s (Grinspoon and Bakalar 1997). The early evidence for the therapeutic use of classical psychedelics was promising for non-psychotic mental health difficulties (Krebs and Johansen 2012; Rucker Illiff Nutt 2018), but studies lacked the stringent methods typical of modern clinical trials, such as placebo conditions, validated measures, blinding, and randomization (Carhart-Harris and Goodwin 2017). Research was largely halted in the 1970s after the drugs were reclassified in the USA as Schedule I substances (substances which have no established medical purpose and have a high risk of abuse) and subsequently banned in most countries (Drug Enforcement Administration Intelligence Division 2016). Before MDMA (3,4-methylenedioxymethamphetamine) was banned in 1985 (Drug Enforcement Administration Intelligence Division 2019), it was widely used for therapeutic purposes, with early reports suggesting that MDMA had psychotherapeutic benefits (Greer and Tolbert 1998; Metzner and Adamson 2001). However, the first systematic study of MDMA-assisted psychotherapy was not published until 2008 (Bouso Doblin Farré Alcázar and Gómez-Jarabo 2008). This study was part of a resurgence of research on the therapeutic application of classical psychedelics, such as psilocybin, LSD, and ayahuasca (which contains a combination of N,N-dimethyltryptamine and harmala monoamine oxidase inhibitor alkaloids) and non-classic psychedelic compounds, such as MDMA. These newer trials include several randomized, placebo-controlled clinical trials that address many of the weaknesses of older studies.

To our knowledge, this paper is the first meta-analysis of modern placebo-controlled clinical trials of psychedelic-assisted therapy. Although the biological and psychological mechanisms of action for MDMA, psilocybin, ayahuasca, and LSD may vary, we review the outcomes of these psychedelic-assisted therapies together because they share several similarities. First, the studies reviewed have comparable study designs, wherein the substance is typically administered between 1-3 times. Second, in most of the trials, patients received concurrent psychotherapeutic support that follows a similar structure. For example, psychotherapeutic support typically occurs during a preparation period, during dosing sessions, and after dosing sessions, during an integration period. Third, these substances are typically only taken in the context of psychotherapeutic support, with the understanding that their effectiveness depends on the individual using them, the context in which they are used, and the psychotherapy that enables a focus on behaviour change post-drug administration. This is typically referred to as the “set and setting”, with “set” referring to attributes of the individual (e.g., their intention, mind state, personality) and “setting” referring to factors outside the individual (e.g., presence of a therapist, culture, and the environment in which the substance is used; Hartogssohn 2016). We refer to these interventions collectively, in the current paper, as psychedelic-assisted therapies.

A meta-analysis is timely for several reasons. First, research into the effects of psychedelic-assisted therapy is rapidly increasing. Multiple organizations are progressing toward clinical trials that could result in MDMA and psilocybin being approved for medical use in the USA and Europe within the next several years. Compounds with novel mechanisms of action,

such as psychedelics, are needed. Psychiatric drug development is widely acknowledged to have slowed with most new drugs considered “me too” drugs that represent small modifications to existing drugs, such as small molecular changes or re-packaging the same drug and changing its mode of administration (O'Brien Thomas Hodgkin Levit and Mark 2014). In addition, political efforts related to psychedelics have begun. In California, in 2018, there was an unsuccessful campaign for the decriminalisation of psilocybin. In 2019, Denver, Colorado and Oakland, California reduced the local enforcement of drug laws against certain psychedelic containing plants and mushrooms to very low levels. In addition, a pending ballot initiative in Oregon would provide a legal mechanism for therapeutic psilocybin mushroom services. Thus, a scientific review of current findings could inform ongoing legislative initiatives, further the development of clinical trials with psychedelic-assisted therapies, and further drug development of psychedelic medicines.

In this meta-analysis, we synthesized the results of all randomized, placebo-controlled trials on psychedelic-assisted therapy published after 1993. We compared the pooled effect size between the experimental and placebo (control) group at the primary endpoint. As an exploratory aim, we also examined effect sizes for classic psychedelics (LSD, ayahuasca, and psilocybin) versus MDMA due to their differing biological and phenomenological effects. To study the durability of effects, we analysed data for those studies that maintained a placebo control through a long-term follow-up assessment. To determine the extent to which trials converge on a typical effect size, we examined between-study variability. We also characterized the risk of bias (Higgins and Green 2011) in these studies so as to suggest avenues for methodological improvements in future clinical trials. Finally, although statistical power was limited, we examined whether moderators might predict effect sizes. Overall, we aimed to synthesize the best available clinical evidence on these therapies and suggest directions for future research.

## Method

### Literature Searches

We sought to identify all published randomized, placebo-controlled clinical trials of psychedelic-assisted therapy since the publishing of the DSM-IV (American Psychiatric Association 1994). Our systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher Liberati Tetzlaff and Altman 2009). Inclusion criteria were: a) original research published in a peer-reviewed journal, b) randomized placebo-controlled clinical trial, and c) assesses the effect of MDMA, psilocybin, ayahuasca, N,N-dimethyltryptamine (DMT), or LSD on symptoms of a diagnosed psychiatric condition listed in either the DSM-IV or DSM-V (American Psychiatric Association 2013). Non-English language was an exclusion criterion. We searched PsycInfo, ERIC, Medline, Academic Search Premiere and CINHALL from January 1994 to March 2019 with terms specifying the active substance administered in the study ('psilocybin', 'MDMA', '3,4-methylenedioxymethamphetamine', 'ayahuasca', 'DMT', '5-MeO-DMT', 'Banisteriopsis caapi', or 'LSD') in pairwise combination with terms specifying the trial's clinical focus ('depressive', 'depression', 'anxiety', 'distress', 'trauma', 'post-traumatic stress disorder', 'obsessive compulsive disorder', 'PTSD', 'OCD', 'alcohol',

'drinking', 'nicotine', 'smoking', 'cigarette', 'marijuana', 'cocaine', 'stimulant', 'opiate', 'heroin', 'methamphetamine', 'depressant', 'inhalant' or 'injection').

### Study Selection

The process of arriving at the final set of papers is outlined in the PRISMA diagram (Figure 1). The search yielded 931 records. Four articles were duplicates within each database. At least one author reviewed the title of each remaining article and excluded 733 articles whose titles did not meet our inclusion criteria. Two authors then independently reviewed the abstracts of the remaining studies and excluded articles if both authors agreed upon exclusion ( $n = 131$ ). For the remainder of the 63 articles, two authors independently reviewed the full texts, with disagreements resolved through discussion until 100% agreement was reached. We excluded 54 articles that were duplicates, not placebo-controlled, or did not treat a mental health condition, leaving 9 studies. We contacted the corresponding author of each selected article to ask if they knew of additional studies (published or unpublished/in progress) that we had not yet included. From these correspondences, we included one additional article (Palhano-Fontes et al. 2019). During the data extraction phase, an additional paper was excluded, because it did not present data in a manner which allowed for meta-analysis and the author did not supply the needed data when contacted. This left us with nine studies included in the meta-analysis (Table 1).

### Data Extraction and Risk of Bias Assessment

All effect sizes were calculated as standardized mean differences between-groups on primary outcome measures, comparing the active treatment groups to the control (placebo) groups. The primary endpoint for each study (reported in Table 1) was identified in the following manner. If authors identified a primary endpoint, this was typically used in our meta-analysis as long as it occurred before a cross-over. Because a cross-over eliminates the placebo-controlled arm, for those studies that used cross-over designs, only data collected before the cross-over was included. For these studies, the primary endpoint was identified as the first available data point following the last psychedelic session. We extracted the following data: sample (i.e. clinical diagnosis), primary outcome measure, sample size, study design, drug and dose administered, number of therapy only sessions, number of dosing sessions before the primary endpoint, type of control group, primary endpoint, percentage of participants who identified as women, participants' average age, and percentage of white participants. A senior investigator checked the extracted data. Disagreements were resolved by discussion. Missing data was obtained by electronic mail (Ross et al. 2016). Risk of bias was assessed using the method prescribed by Cochrane's Handbook for Systematic Reviews and Interventions (Higgins and Green 2011; see online Appendix 1). A senior investigator checked the data regarding risk of bias, with disagreements resolved by discussion.

### Data Synthesis and Analyses

**Computation and analysis of effect sizes.**—Effect sizes were calculated as a standardized mean difference (Cohen's  $d$ ), which were then transformed to Hedges  $g$  to adjust for bias due to sample size (Hedges and Olkin 1985). Effect sizes were calculated by

subtracting the average score of the placebo group from the average score of the psychedelic-assisted therapy group and dividing by the pooled standard deviations of the two groups. Effect sizes of .2, .5, .8, and 1.2 are typically considered small, medium, large, and very large, respectively (Cohen 1992; Sawilowsky, 2009). In calculating effect sizes, we identified primary outcome measures from registration information at [clinicaltrials.gov](https://clinicaltrials.gov) or in the published reports. In cases where multiple outcomes were designated as primary (Gasser et al. 2014; Griffiths et al. 2016), effect sizes were calculated for each variable and then averaged to create a single effect size for each study (Borenstein Hedges Higgins and Rothstein 2009). We selected a random-effects model based upon the assumption that our data would have variable population parameters and average effect sizes (Hunter and Schmidt 2004) and that the studies are not functionally equivalent in design and methods (Borenstein et al. 2009). We conducted meta-analyses using Meta-Essentials (Suurmond van Rhee and Hak 2017).

**Moderator analyses.**—We tested six continuous and two categorical moderating variables using meta-regressions. We tested the following continuous moderating variables: gender (operationalized as the percentage of participants who identified as men versus women), race/ethnicity (operationalized as the percentage of White participants versus people of colour), mean age, number of dosing sessions, number of psychotherapy sessions, and sample size. Due to the relatively small number of people of colour across multiple countries, we did not have enough power to test moderations for multiple ethnic or racial categories. We tested the following categorical moderating variables: inert placebo (versus other controls) and low dose psychedelic (versus other controls).

## Results

### Characteristics of Studies

Characteristics of the nine studies are detailed in Table 1. A total of 211 people participated in these studies. Race was reported in seven studies (85% reported being White) and gender was reported in all studies (53% identified as women). Five studies examined MDMA, two psilocybin, one ayahuasca, and one LSD. The average number of dosing sessions before the primary endpoint (as defined in this meta-analysis) was 1.9 ( $SD = 0.8$ ) with a range of 1-3 dosing sessions across studies. Placebo controls consisted of very low doses of psychedelic in five studies, inert placebo in three, and an active placebo (i.e., niacin) in one. Three studies included placebo-controlled long-term follow-up assessments.

Online Appendix 1 summarizes our risk of bias findings. All studies were preregistered at [clinicaltrials.gov](https://clinicaltrials.gov). In most of the trials, the randomization sequence was not clearly described. Most studies did not describe the methods for concealing the allocation sequence in sufficient enough detail so as to allow for quality determinations. We judged five of the nine studies as either high or unclear risk in terms of whether patients, research staff, and clinicians were blind to treatment assignment. We judged three of the studies as high risk in terms of blinding of outcome assessment. Attrition rates across all studies were excellent, resulting in a low risk of bias due to study dropout. All studies analysed pre-determined primary outcomes, resulting in a very low risk of bias due to selective outcome reporting.

## Overall Effect Sizes

The analysis of the pooled effect of psychedelic-assisted therapy in these nine studies resulted in a mean between-groups effect of Hedges  $g = 1.21$  (see Figure 2), which was significantly different from zero,  $z = 9.48$ ,  $p < .001$ . We assessed heterogeneity using the  $Q_w$  statistic. A nonsignificant Q test suggested that there was no significant heterogeneity beyond that which could be accounted for by sampling error,  $Q_w(8) = 5.61$ ,  $p = .69$ . The  $I^2$  value, which describes the percentage of variation across studies that is due to heterogeneity rather than chance, could not be calculated due to a negative numerator, which further suggests very low levels of heterogeneity. We used Rosenthal's failsafe N to assess the number of additional studies with an effect size of zero that would be needed to bring statistically significant associations down to the  $p = .05$  level (Rosenthal 1991). The failsafe N was 193, suggesting that the observed effects are tolerant to null results. The funnel plot (see Figure 3) also suggests low levels of heterogeneity and a lack of publication bias (Sterne et al. 2011).

## Subgroup Analyses

We conducted subgroup analyses comparing the four trials of classic psychedelic-assisted therapy to the five trials of MDMA-assisted therapy (see Figure 2). There was no significant difference in the overall effect size between these two subgroups,  $z = .08$ ,  $p = .94$ . The mean effect size for the four studies of classic psychedelic-assisted therapy was Hedges  $g = 1.20$  and the mean effect size for the five studies of MDMA-assisted therapy was Hedges  $g = 1.22$ .

## Moderator Analyses

Pre-planned tests of moderation were calculated individually for the following variables: sample size ( $p = .85$ ), number of dosing sessions ( $p = .97$ ), number of psychotherapy sessions ( $p = .64$ ), percent of sample White ( $p = .60$ ), percent of sample women ( $p = .61$ ), average sample age ( $p = .93$ ), and type of control condition (inert placebo vs other controls,  $p = .81$  and low dose psychedelic vs other controls,  $p = .82$ ). None of these variables were statistically significant moderators.

## Effects at Long-Term Follow-Up

Three studies reported long-term follow-up data with placebo control beyond the primary endpoint (see Figure 2). Long-term follow-up endpoints were six months (Danforth et al. 2018), two months (Mithoefer Wagner Mithoefer Jerome and Doblin 2011), and seven weeks (Ross et al. 2016) after the primary endpoint assessment. The average effect size for these three studies at the primary endpoint was Hedges  $g = 1.47$ . At the long-term follow-up point, the average effect size was Hedges  $g = 1.36$ . This represents a decrease of 7.5% in average effect size by the follow-up point, indicating that effects were largely sustained over the follow-up period.

## Discussion

This meta-analysis of nine placebo-controlled randomized clinical trials of psychedelic-assisted therapy provides support for their efficacy across four mental health conditions:



PTSD, anxiety/depression associated with life-threatening illness, unipolar depression, and social anxiety among autistic adults. The overall between-group effect size at the primary endpoint for psychedelic-assisted therapy compared to placebo was very large (Hedges  $g = 1.21$ ). This effect size reflects an 80% probability that a randomly selected patient undergoing psychedelic-assisted therapy will have a better outcome than a randomly selected patient receiving a placebo (McGraw and Wong 1992).

The effect sizes reported were obtained using a therapeutic agent that participants only received a total of 1-3 times, as opposed to most psychopharmacological agents which patients receive at least daily. This suggests the possibility of a novel mechanism of action wherein effects extend beyond the direct biological effects of the psychedelic substance. Aside from the observed effect size, other strengths of these trials included low levels of attrition, a documented lack of serious adverse effects, and all trials not deviating from their preregistered primary outcomes. Overall, results suggest that psychedelic-assisted therapy is effective with minimal adverse effects.

In order to understand the magnitude of the observed effect, we identified other meta-analyses of randomized clinical trials of pharmacological and psychological interventions for mental health conditions. Meta-analyses comparing pharmacotherapies to placebo have found small to moderate effect sizes: .30 for PTSD (Cipriani et al. 2018), .32 for major depressive disorder (Leucht Hierl Kissling Dold and Davis 2012), .31 for generalized anxiety disorder (Huhn et al. 2014), .55 for social anxiety disorder (Huhn et al. 2014), and .40 for mental health problems overall (Huhn et al. 2014). Meta-analyses comparing gold standard psychotherapy treatments to control treatments have found moderate effect sizes: .53 for depression (Cuijpers et al. 2013), .56 for social anxiety disorder (Mayo-Wilson et al. 2014), and .65 for PTSD (Powers Halpern Ferenschak Gillihan and Foa 2010). It's possible that the effect size observed in our review was larger partially because psychedelic-assisted therapy is a combined treatment. Meta-analyses have found, for example, that combined psychotherapy and pharmacotherapy treatments for depression were larger than for pharmacotherapy (Huhn et al. 2014; Cuijpers et al. 2013) or psychotherapy (Huhn et al. 2014) alone. Although we compared the effect size in our study to those reported in prior reviews and meta-analyses, future research would benefit from directly comparing the effectiveness of psychedelic-assisted therapy to more traditional pharmacological, psychotherapeutic, or combined interventions. Future studies would also benefit from examining cost-effectiveness, given that psychedelic-assisted therapy, as currently delivered, may be relatively expensive compared to single-modality treatments.

Further research is needed to examine the reliability of the effect sizes observed. Previous reviews have shown that effect sizes may diminish as larger trials are conducted (Johnsen and Friberg 2015). The extent to which this may occur in future trials is unknown. Alternatively, it's possible that the effect size may be larger in future trials, as five of the nine trials in our meta-analysis included chronic or treatment-resistant populations and some trials only had one dosing session as we analysed data before cross-over. Additionally, the type of placebo may have influenced outcomes: five of the studies used placebos consisting of putatively therapeutically inactive doses of the substance under investigation, two used

inactive placebos, and two used active placebos. The effects of these varying placebos require examination in future studies.

We found no significant moderators of treatment effects on planned moderation analysis, including comparisons of the effect sizes for classical psychedelics to those of MDMA. This was perhaps not surprising given the low level of heterogeneity in the sample and small number of studies in the meta-analysis resulting in low statistical power. Results suggested comparable efficacy across the two classes of treatments (MDMA versus classical psychedelics). The p-value did not approach the  $p = .05$  level ( $p = .67$ ), suggesting that increased power may not have yielded significant results. The lack of a statistically significant difference could also be affected by the fact that the studies treated different mental health conditions; while the MDMA studies treated PTSD and social anxiety, the classical psychedelic studies treated end-of-life depression/anxiety and major depression. Future research would benefit from having larger sample sizes that enable moderation and mediation analyses.

The placebo-controlled long-term follow-up data available suggested that treatment effects were maintained after the primary endpoint assessment, with effect sizes decreasing an average of 7.5% during the follow-up period (Danforth et al. 2018; Mithoefer et al. 2011; Ross et al. 2016). Nevertheless, these results should be interpreted with caution as only three studies included placebo-controlled follow-up data and additional psychotherapy occurred during this follow-up period in one of the trials (Mithoefer et al. 2011). Studies that maintain placebo controls through a long-term follow-up assessment (when ethical) could help clarify the durability of these interventions. Future meta-analyses focusing on within-group effect sizes could provide additional information about the durability of effects because analyses could include additional studies that did not have a placebo control.

The observed effect sizes are most indicative of the likely effects of psychedelics among those with end-of-life distress and those with PTSD, and among White volunteers. Only four mental health conditions were studied across all nine studies: PTSD ( $n = 4$ ), anxiety/depression associated with life-threatening illness ( $n = 3$ ), unipolar depression ( $n = 1$ ), and social anxiety among autistic adults ( $n = 1$ ). Despite good representation by women (53%), people of color were largely absent from these studies of psychedelic-assisted therapy and more work is needed to include people of color in these trials. Approaches such as participatory research could help recruit more diverse samples (Israel Eng Schulz Parker and Satcher 2005).

Through our risk of bias assessment (see online Appendix 1), we identified several potential sources of bias, including possible failures in randomization, failures of concealment of allocation, lack of blinding of treatment staff, patients, and assessors, and lack of blinding of outcome assessment in some studies. More research is needed into how to effectively achieve blinding when investigating compounds that frequently have obvious acute subjective effects, a well-known methodological challenge in pharmacological research (Lee et al. 2016). This may require study designs that depart from standard clinical trial methods to assess the influence of different blinding procedures and placebo conditions. One option could be to vary the expectancy effects of the pharmacological agents used in the trial, and



to randomize the allocation of the placebo session. Some of the measures that authors in the current studies took to achieve blinding included using a low-dose psychedelic (Gasser et al. 2014; Griffiths et al. 2016; Mithoefer et al. 2011; Oehen et al. 2013; Ot'alora et al. 2018) or a substance that mimicked some of the psychoactive substances' effects (Palhano-Fontes et al. 2019; Ross et al. 2016) as a control. So that blinding success can be quantitatively evaluated, we recommend that future investigators directly assess whether the blind was broken by asking assessors, treatment staff, and patients to guess whether they received the active treatment.

Relatedly, studies could be improved by better describing, controlling for, and studying concurrent psychotherapy practices especially given the role of set and setting in outcomes from psychedelic-assisted therapy. For example, trials that randomize type of or amount of psychotherapy, while maintaining consistency of psychedelic dosing, would be helpful in understanding the role of concurrent psychotherapy practices. Because it is generally acknowledged that psychedelics cannot be ethically administered for treatment purposes without concurrent psychotherapeutic support (Carhart-Harris and Goodwin 2017), the role of this aspect of treatment merits further study. In several reports, concurrent psychotherapy and supportive contact was not well described or quantified. In addition, it was not clear whether the amount of contact with therapists was balanced among conditions. Given research showing a dose-response effect of greater amounts of psychotherapy (Falkenström Josefsson Berggren and Holmqvist 2016), the levels of concurrent contact should be balanced between conditions and reported.

We did not find any statistical indication of reporting bias. At the same time, a recent report of six Phase 2 randomized trials of MDMA-assisted psychotherapy for the treatment of PTSD demonstrated a combined between-groups effect size of Cohen's  $d = .8$  (Mithoefer et al. 2019), which is lower than what we calculated based on the four MDMA trials that had been published and were thereby included in this report (Hedges  $g = 1.21$ ). It is possible that our effect size was larger than that observed in the recent report (Mithoefer et al. 2019) because we calculated effect sizes differently or because the two unpublished trials that we did not review had less favorable findings than the four that were published. This could reflect publication bias that might have affected our estimates.

We suggest that future investigators conform with the CONSORT statement (Schulz Altman and Moher 2010) for reporting clinical trials and report the information needed to allow categorization on risk of bias (Higgins and Green 2011). In particular, improvements compared to current studies could be made in terms of reporting of randomization procedures, allocation procedures, condition blinding procedures and success, methods for blinding assessors, and the reporting of adverse events (Carhart-Harris and Goodwin 2017).

## Conclusion

This meta-analytic review provides support for ongoing research on psychedelic-assisted therapy. The effect sizes of these randomized placebo-controlled trials were large, when compared to the effect sizes obtained with more well-researched stand-alone pharmacological agents and psychotherapy interventions. The quality of the reviewed

studies was high, with low levels of attrition, a documented lack of serious adverse effects, and no deviations from preregistered primary outcomes. However, a number of weaknesses were identified and improvements are needed regarding the conduct and reporting of trials in the future. Despite these limitations, large effect sizes are rare in psychiatry and psychology and our findings suggest that psychedelic-assisted therapy presents a promising new direction in mental health treatment.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Brandon Strickland for his help during the literature search. During his work on this study, Dr. Davis was supported by a NIDA T32 under grant #DA07209. The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. J.B.L., G.B., and A.K.D. conceived the meta-analysis and facilitated data extraction. J.B.L. and C.C. conducted risk of bias assessment and assisted in data analysis. C.C. and R.L. conducted literature review and extracted data. All authors assisted in drafting and editing the manuscript.

## References

- American Psychiatric Association. 1994 Diagnostic and Statistical Manual of Mental Disorders 4 Edition (Text Revision). American Psychiatric Publishing.
- American Psychiatric Association. 2013 Diagnostic and Statistical Manual of Mental Disorders 5th Edition American Psychiatric Publishing.
- Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, and Keane TM 1995 The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress* 8: 75–90. [PubMed: 7712061]
- Borenstein M, Hedges LV, Higgins JPT, & Rothstein HR 2009 *Introduction to Meta-Analysis*. John Wiley & Sons.
- Bouso JC, Doblin R, Farré M, Alcázar MÁ, & Gómez-Jarabo G 2008 MDMA-assisted psychotherapy using low doses in a small sample of women with chronic posttraumatic stress disorder. *Journal of Psychoactive Drugs* 40: 225–236. [PubMed: 19004414]
- Bruss GS, Gruenberg AM, Goldstein RD and Barber JP 1994 Hamilton Anxiety Rating Scale Interview guide: joint interview and test-retest methods for interrater reliability. *Psychiatry research* 53: 191–202. [PubMed: 7824679]
- Carhart-Harris RL & Goodwin GM 2017 The therapeutic potential of psychedelic drugs: Past, present and future. *Neuropsychopharmacology* 42: 2105–2113. [PubMed: 28443617]
- Cipriani A et al. 2018 Comparative efficacy and acceptability of pharmacological treatments for post-traumatic stress disorder in adults: A network meta-analysis. *Psychological Medicine* 48: 1975–1984. [PubMed: 29254516]
- Cohen J 1992 A power primer. *Psychological Bulletin* 112: 155–159. [PubMed: 19565683]
- Cuijpers P, et al. 2013 A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry* 58: 376–385. [PubMed: 23870719]
- Danforth AL et al. 2018 Reduction in social anxiety after MDMA-assisted psychotherapy with autistic adults: A randomized, double-blind, placebo-controlled pilot study. *Psychopharmacology* 235: 3137–3148. [PubMed: 30196397]
- Drug Enforcement Administration Intelligence Division. Title 21 United States Code (USC) Controlled Substances Act. 2016 At <https://www.deadiversion.usdoj.gov/21cfr/21usc/> (2016).
- Drug Enforcement Administration Intelligence Division. Controlled substance schedules. Accessed May 8, 2020 <https://www.deadiversion.usdoj.gov/schedules/> (2019).

- Falkenström F, Josefsson A, Berggren T, & Holmqvist R 2016 How much therapy is enough? Comparing dose-effect and good-enough models in two different settings. *Psychotherapy* 53: 130–139. [PubMed: 26928273]
- Gasser P et al. 2014 Safety and efficacy of lysergic acid diethylamide-assisted psychotherapy for anxiety associated with life-threatening diseases. *Journal of Nervous and Mental Disease* 202: 513–520.
- Greer GR, & Tolbert R 1998 A method of conducting therapeutic sessions with MDMA. *Journal of Psychoactive Drugs* 30: 371–379. [PubMed: 9924843]
- Griffiths RR et al. 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: 1181–1197. [PubMed: 27909165]
- Grinspoon L, & Bakalar JB 1997 *Marijuana, the Forbidden Medicine*. Yale University Press.
- Grob CS et al. 2011 Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer. *Archives of General Psychiatry* 68: 71–78. [PubMed: 20819978]
- Hamilton M 1960 Hamilton Depression Rating Scale (HAM-D) instructions for the clinician: HAM-D scoring instructions. *Journal of Neurology* 23: 56–62.
- Hartogsohn I 2016 Set and setting, psychedelics and the placebo response: An extra-pharmacological perspective on psychopharmacology. *Journal of Psychopharmacology* 30: 1259–1267. [PubMed: 27852960]
- Hedges LV & Olkin I 1985 *Statistical Methods for Meta-Analysis*. Academic Press.
- Higgins J, & Green S 2011 *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration.
- Huhn M et al. 2014 Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: A systematic overview of meta-analyses. *JAMA Psychiatry* 71: 706–715. [PubMed: 24789675]
- Hunter JE, & Schmidt FL 2004 *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*.
- ISCDD. 2003 GRID-HAMD-17. Structured Interview Guide. International Society for CNS Drug Development.
- Israel BA, Eng E, Schulz AJ, Parker EA, & Satcher D 2005 Introduction to Methods in Community-Based Participatory Research for Health In *Methods in Community-Based Participatory Research for Health*, eds. Israel BA, Eng E, Schulz AJ, and Parker EA Jossey-Bass.
- Johnsen TJ, & Friborg O 2015 The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin* 141: 747–768. [PubMed: 25961373]
- Krebs TS, & Johansen PO 2012 Lysergic acid diethylamide (LSD) for alcoholism: Meta-analysis of randomized controlled trials. *Journal of Psychopharmacology* 26: 994–1002. [PubMed: 22406913]
- Lee DJ et al. 2016 Psychotherapy versus pharmacotherapy for posttraumatic stress disorder: Systemic review and meta-analyses to determine first-line treatments. *Depression and Anxiety* 33: 792–806. [PubMed: 27126398]
- Leucht S, Hierl S, Kissling W, Dold M, & Davis JM 2012 Putting the efficacy of psychiatric and general medicine medication into perspective: Review of meta-analyses. *British Journal of Psychiatry* 200: 97–106.
- Liebowitz MR, Gorman JM, Fyer AJ, & Klein DF 1985 Social phobia: Review of a neglected anxiety disorder. *Archives of General Psychiatry* 42: 729–736. [PubMed: 2861796]
- Mayo-Wilson E et al. 2014 Psychological and pharmacological interventions for social anxiety disorder in adults: A systematic review and network meta-analysis. *Lancet Psychiatry* 1: 368–376. [PubMed: 26361000]
- McGraw KO, & Wong SP 1992 A common language effect size statistic. *Psychological Bulletin* 111: 361–365.
- Metzner R, & Adamson S 2001 *Using MDMA in Healing, Psychotherapy and Spiritual Practice In Ecstasy: The Complete Guide*, ed. Holland J Inner Traditions, 2001.
- Mithoefer MC et al. 2018 3, 4-methylenedioxymethamphetamine (MDMA)-assisted psychotherapy for post-traumatic stress disorder in military veterans, firefighters, and police officers: A randomised,

- double-blind, dose-response, phase 2 clinical trial. *Lancet Psychiatry* 5: 486–497. [PubMed: 29728331]
- Mithoefer MC, et al. 2019 MDMA-assisted psychotherapy for treatment of PTSD: Study design and rationale for phase 3 trials based on pooled analysis of six phase 2 randomized controlled trials. *Psychopharmacology* 236: 1–11. [PubMed: 30604181]
- Mithoefer MC, Wagner MT, Mithoefer AT, Jerome L, & Doblin R 2011 The safety and efficacy of  $\pm$ 3,4-methylenedioxymethamphetamine-assisted psychotherapy in subjects with chronic, treatment-resistant posttraumatic stress disorder: The first randomized controlled pilot study. *Journal of Psychopharmacology* 25: 439–452. [PubMed: 20643699]
- Moher D, Liberati A, Tetzlaff J, & Altman DG 2009 Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine* 151: 264–269. [PubMed: 19622511]
- O'Brien PL, Thomas CP, Hodgkin D, Levit KR, & Mark TL 2014 The diminished pipeline for medications to treat mental health and substance use disorders. *Psychiatric Services* 65: 1433–1438. [PubMed: 25178309]
- Oehen P, Traber R, Widmer V, & Schnyder U 2013 A randomized, controlled pilot study of MDMA ( $\pm$ 3, 4-Methylenedioxymethamphetamine)-assisted psychotherapy for treatment of resistant, chronic Post-Traumatic Stress Disorder (PTSD). *Journal of Psychopharmacology* 27: 40–5. [PubMed: 23118021]
- O'Alora GM et al. 2018 3, 4-Methylenedioxymethamphetamine-assisted psychotherapy for treatment of chronic posttraumatic stress disorder: A randomized phase 2 controlled trial. *Journal of Psychopharmacology* 32: 1295–1307. [PubMed: 30371148]
- Palhano-Fontes F et al. 2019 Rapid antidepressant effects of the psychedelic ayahuasca in treatment-resistant depression: A randomized placebo-controlled trial. *Psychological Medicine* 49: 655–663. [PubMed: 29903051]
- Powers MB, Halpern JM, Ferenschak MP, Gillihan SJ, & Foa EB 2010 A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical Psychology Review* 30: 635–641. [PubMed: 20546985]
- Rosenthal R 1991 *Meta-Analytic Procedures for Social Research*. Sage, 1991.
- Ross S, Bossis A, Guss J, Agin-Liebes G, Malone T, Cohen B, ... Schmidt BL 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: 1165–1180. [PubMed: 27909164]
- Rucker JJ, Iliff J, & Nutt DJ *Psychiatry & the psychedelic drugs*. 2018. Past, present & future. *Neuropharmacology* 142: 200–218. [PubMed: 29284138]
- Sawilowsky SS 2009 New effect size rules of thumb. *Journal of Modern Applied Statistical Methods* 8: 597–599.
- Schulz KF, Altman DG, & Moher D 2010 CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. *Annals of Internal Medicine* 152: 726–732. [PubMed: 20335313]
- Schnyder U & Moergeli H 2002 German version of clinician-administered PTSD scale. *Journal of Traumatic Stress* 15: 487–492. [PubMed: 12482188]
- Shear MK et al. 2001 Reliability and validity of a structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-A). *Depression and Anxiety* 13: 166–178. [PubMed: 11413563]
- Spielberger CD, Gorsuch RL, Lushene RE, Vagg PR, & Jacobs GA 1970 *State-Trait Anxiety Inventory (STAI) for Adults-Manual*. Mind Garden Inc..
- Sterne JA et al. 2011 Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal* 343: d4002. [PubMed: 21784880]
- Suurmond R, van Rhee H, & Hak T 2017 Introduction, comparison and validation of MetaEssentials: A free and simple tool for meta-analysis. *Research Synthesis and Methods* 8: 537–553.
- Weathers FW, Keane TM, & Davidson JR 2001 Clinician-Administered PTSD Scale: A review of the first ten years of research. *Depression and Anxiety* 13: 132–156. [PubMed: 11387733]

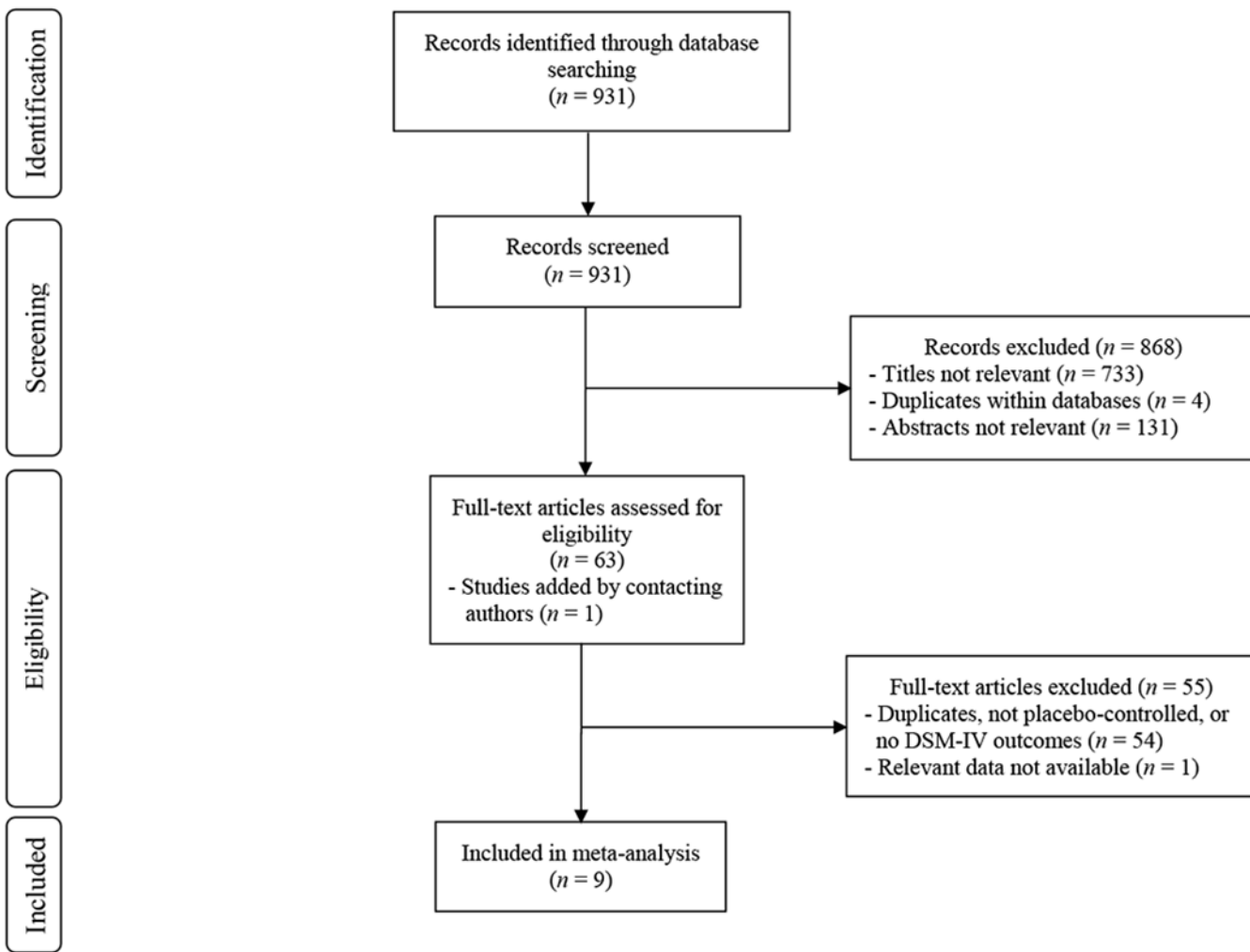
- Williams JB, Kobak KA, Bech P, Engelhardt N, Evans K, Lipsitz J, Olin J, Pearson J and Kalali A 2008 The GRID-HAMD: standardization of the Hamilton depression rating scale. *International clinical psychopharmacology* 23: 120–129. [PubMed: 18408526]
- Zigmond AS, & Snaith RP 1983 The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica* 67: 361–370. [PubMed: 6880820]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 1.**  
RISMA Diagram Outlining Study Selection

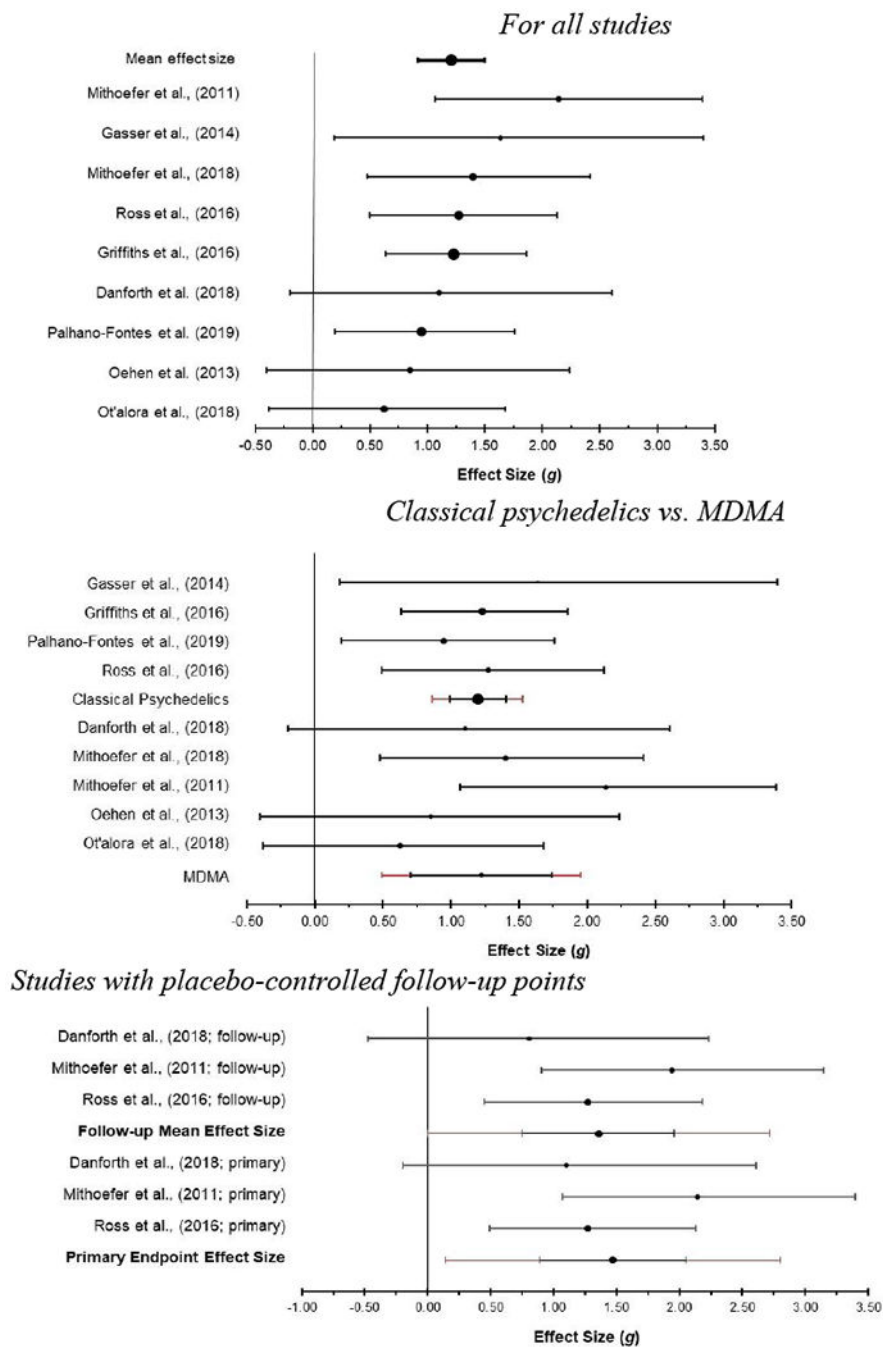


Author Manuscript

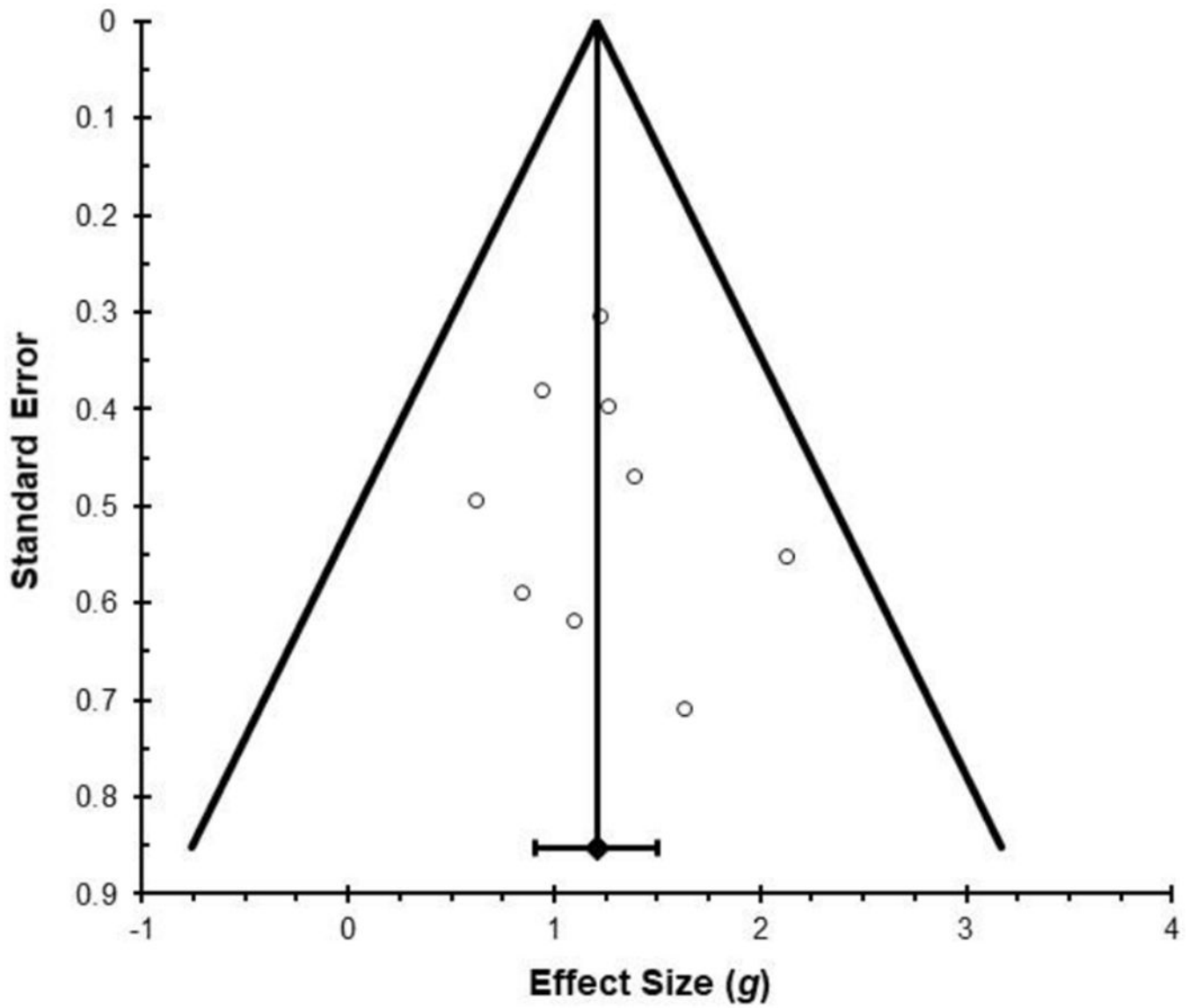
Author Manuscript

Author Manuscript

Author Manuscript



**Figure 2.** Forest plots with mean weighted effect sizes and 95% confidence intervals For all studies



**Figure 3.**  
Funnel Plot Depicting Effect Size (Hedge's  $g$ ) as a Function of Standard Error

Table 1

Descriptive Summary of Included Studies

Study	Sample	Primary Outcome	N <sup>a</sup>	Study Design	Drug Dose	Therapy Only Sessions (# / hours) <sup>b,c</sup>	Dosing Sessions <sup>b</sup>	Control Group	Primary End-point	% Women	Mean Age	% White
Danforth et al. (2018)	Socially anxious adults with autism	Social Anxiety (LSAS)	11 (MDMA n = 7; placebo n = 4)	Double-blind, randomized, placebo-controlled, parallel arm	MDMA (75-125 mg)	9 / Mean = 11.3	3	Inert placebo	1-month after 2nd experimental session	16.7	31.3	50.0
Gasser et al. (2014)	Anxiety associated with life-threatening illness	Anxiety (STAI-state and trait)	11 (LSD n = 8; placebo n = 3)	Double-blind, randomized, placebo-controlled, open-label cross-over	LSD (200 µg)	Unclear	2	Low-dose psychedelic (20 µg LSD)	2-month after 2nd blinded session	36.4	51.7	Not reported
Griffiths et al. (2016)	Life-threatening cancer, depression and/or anxiety	Depression (HAM-D) & Anxiety (HAM-A)	50 (High dose first n = 25; low dose first n = 25)	Double-blind, randomized, placebo-controlled, cross-over	Psilocybin (22 mg or 30 mg/70 kg)	Mean = 4.0 / Mean = 9.1	1	Low-dose psychedelic (1 or 3 mg/70 kg psilocybin)	5-weeks after session 1	51.0	56.3	94
Mithoefer et al. (2011)	Chronic PTSD	PTSD (CAPS-1V)	20 (MDMA n = 12; placebo n = 8)	Double-blind, randomized, placebo-controlled, open-label cross-over	MDMA (125 mg; optional 65 mg supplemental)	10 / 15.0	2	Inert placebo	3-5 days after session 2	85.0	40.4	100.0
Mithoefer et al. (2018)	Chronic PTSD	PTSD (CAPS-1V)	26 (MDMA 125/75 mg combined n = 19; MDMA 30 mg n = 7)	Double-blind, randomized, placebo-controlled, open-label cross-over	MDMA (75 mg or 125 mg)	9 / 13.5	2	Low-dose psychedelic (30 mg MDMA)	1-month after experimental session 2	26.9	37.2	84.6
Oehen et al. (2013)	Chronic PTSD	PTSD (CAPS-1V)	12 (MDMA n = 8; placebo n = 4)	Double-blind, randomized, placebo-controlled, open-label cross-over	MDMA (125 mg, 62.5 mg supple-mental)	Mean = 13.30 / Unclear	3	Low-dose (25 mg) psychedelic; supplemental (12.5 mg) dose	3-weeks after experimental session 3	83.3	41.4	Not reported
Of'alora et al. (2018)	Chronic PTSD	PTSD (CAPS-1V)	23 (MDMA 125/100 mg combined n = 18; MDMA 40 mg n = 5)	Double-blind, randomized, open-label cross-over	MDMA (125 or 100 mg)	Unclear	2	Low-dose psychedelic (40 mg MDMA)	1-month after 2nd blinded session	32.1	42.0	92.9

Study	Sample	Primary Outcome	N <sup>a</sup>	Study Design	Drug Dose	Therapy Only Sessions (# / hours) <sup>b,c</sup>	Dosing Sessions <sup>b</sup>	Control Group	Primary End-point	% Women	Mean Age	% White
Palhano-Fontes et al. (2019)	Treatment-resistant major unipolar depressive disorder	Depression (HAM-D)	29 (Ayahuasca n = 14; placebo n = 15)	Double-blind, randomized, placebo-controlled, parallel arm	Ayahuasca (containing 0.36 mg/kg of N, N-DMT)	Unclear	1	Liquid causing mild gastric distress; matched for taste/color	1-week	72.4	42.0	58.6
Ross et al. (2016)	Life-threatening cancer; anxiety diagnosis	Anxiety and Depression (HADS)	29 (Psilocybin 1st n = 14; niacin 1st n = 15)	Double-blind, randomized, placebo-controlled, cross-over	Psilocybin (0.3 mg/kg)	6 / 12.0	1	Niacin	1-day after dose one	62.1	56.3	89.7

Note.

<sup>a</sup> = sample size at primary endpoint;

<sup>b</sup> = data before primary endpoint;

<sup>c</sup> = estimates of sessions/hours are best estimates as this information was not always clearly specified within papers;

CAPS-IV = Clinician Administered PTSD Scale – IV (Blake Weathers Nagy Kaloupek Gusman Charney and Keane 1995); GRID-HAMD = GRID-Hamilton Depression Rating Scale (Williams et al. 2008), HADS = Hospital Anxiety and Depression Scale (Zigmond and Snaith 1983); HAM-A = Hamilton Anxiety Rating Scale (Bruss Gruenberg Goldstein and Barber 1994); HAM-D = Hamilton Depression Rating Scale (Hamilton 1960); LSAS = Liebowitz Social Anxiety Scale (Liebowitz Gorman Fyer and Klein 1985)