Original Paper

The Challenging Experience Questionnaire: Characterization of challenging experiences with psilocybin mushrooms

Psychopharm

Journal of Psychopharmacology
1–17

DOI: 10.1177/0269881116678781
jop.sagepub.com
SAGE

Frederick S Barrett[1], Matthew P Bradstreet[1], Jeannie-Marie S Leoutsakos[1], Matthew W Johnson[1] and Roland R Griffiths[1,2]

Abstract

Acute adverse psychological reactions to classic hallucinogens ("bad trips" or "challenging experiences"), while usually benign with proper screening, preparation, and support in controlled settings, remain a safety concern in uncontrolled settings (such as illicit use contexts). Anecdotal and case reports suggest potential adverse acute symptoms including affective (panic, depressed mood), cognitive (confusion, feelings of losing sanity), and somatic (nausea, heart palpitation) symptoms. Responses to items from several hallucinogen-sensitive questionnaires (Hallucinogen Rating Scale, the States of Consciousness Questionnaire, and the Five-Dimensional Altered States of Consciousness questionnaire) in an Internet survey of challenging experiences with the classic hallucinogen psilocybin were used to construct and validate a Challenging Experience Questionnaire. The stand-alone Challenging Experience Questionnaire was then validated in a separate sample. Seven Challenging Experience Questionnaire factors (grief, fear, death, insanity, isolation, physical distress, and paranoia) provide a phenomenological profile of challenging aspects of experiences with psilocybin. Factor scores were associated with difficulty, meaningfulness, spiritual significance, and change in well-being attributed to the challenging experiences. The factor structure did not differ based on gender or prior struggle with anxiety or depression. The Challenging Experience Questionnaire provides a basis for future investigation of predictors and outcomes of challenging experiences with classic hallucinogens.

Keywords

Psilocybin, challenging experiences, scale development, factor analysis, psychedelics

Introduction

Classic psychedelic hallucinogens, such as psilocybin, lysergic acid diethylamide (LSD), mescaline, and dimethyltryptamine (DMT; contained in the sacramental beverage ayahuasca) share a primary pharmacological site of action at the serotonin 2a (5HT2a) receptor, and display striking similarities in their subjective effects (Halberstadt, 2015; Nichols, 2016). The first era of modern human research on psychedelics, from roughly the 1950s to the 1970s, capitalized on the powerful effects of psychedelics on perception, emotions, and consciousness by investigating the therapeutic and psychotomimetic properties of these substances.

The recent resurgence of empirical research on psychedelics has revived their investigation as therapeutic agents (Bogenschutz et al., 2015; Carhart-Harris et al., 2012, 2016a; Griffiths et al., 2016; Grob et al., 2011; Johnson et al., 2014; Ross et al., 2016) and tools to understand psychosis (Schmid et al., 2015; Vollenweider et al., 1998, 2007). Recent studies have also examined mystical-type (Barrett et al, 2015; Griffiths et al., 2006, 2011) or spiritual (Kometer et al., 2015) experiences, changes in self-referential processing or "ego dissolution" (Carhart-Harris et al., 2014; Tagliazucchi et al., 2016), altered social processing (Preller et al., 2016), visual imagery (Kaelen et al., 2016), changes in emotional experience (Kaelen et al., 2015; Kraehenmann et al., 2015), personality change (Carhart-Harris et al., 2016b; Lebedev et al., 2016; MacLean et al., 2011), positive mental health outcomes (Hendricks et al., 2015; Johansen and Krebs, 2015), and enduring positive changes in attitudes, mood and behavior (Griffiths et al., 2008, 2011) occasioned by these compounds. However, both recent experimental reports (Griffiths et al., 2006, 2011; Johnson et al., 2014; Studerus et al., 2011) and past clinical reports (Cohen, 1960; Strassman, 1984) indicate that challenging psychological experiences during the acute effects of psychedelics are not uncommon.

Phenomenology of challenging psychedelic experiences

An early summary of reports from investigators conducting therapeutic research with LSD and mescaline (for almost 5000

[1]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[2]Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Corresponding author:
Frederick S Barrett, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, 5510 Nathan Shock Drive, Baltimore, MD 21224, USA.
Email: fbarret2@jhmi.edu

individuals and more than 25,000 separate drug administrations) noted instances of fear, delusions, dissociation, depersonalization, and sympathetic nervous system responses during acute effects of these drugs (Cohen, 1960). Reports of psychosis lasting more than 48 h were exceedingly rare (<0.001% of healthy individuals, and <0.002% of individuals undergoing therapy), however this summary must be viewed tentatively, given numerous limitations including, for example, that none of the reporting investigators had conducted follow-up with their patients (Novak, 1997).

A subsequent review, integrating both clinical literature (including emergency department reports) and research literature, presented a summary of acute adverse effects of classic hallucinogens that included frightening illusions and hallucinations, overwhelming anxiety or panic, confusion, aggression and possible violence, depression with suicidal ideation, gestures, or attempts, and fear to the point of paranoid delusions (Strassman, 1984). When encountered, these symptoms typically responded to verbal reassurance and typically subsided within 48 h, often spontaneously. If severe, these symptoms typically responded to medication. However, the review noted that emergency department reports (e.g. McCabe, 1977; Taylor et al., 1970) are limited by uncertainty of substance ingested and unknown mental state of the patient prior to putative ingestion of a psychoactive substance.

Recent research reports that address challenging aspects of experiences with psychedelics (hereby referred to as "challenging experiences" for brevity) are more specific with regards to substance and dose administered.

Approximately 30% of participants in each of three highly controlled experimental studies (total sample: 69 participants and 204 sessions – Griffiths et al., 2006, 2011; Johnson et al., 2014) involving a high dose of psilocybin (0.429 mg/kg psilocybin) experienced marked periods of anxiety or fear, while between 17–39% experienced paranoia (Griffiths et al., 2006, 2011). Only three of 42 (7.1%) participants receiving up to a moderate dose of psilocybin (0.315 mg/kg psilocybin) in a separate series of studies reported marked periods of anxiety, fear, or dysphoria, while only one (2.4%) of those participants reported prolonged symptoms (Studerus et al., 2011).

Clinical and experimental literature on psychedelics suggests a possible profile of challenging experiences that includes the following categories of experience: fear or panic, paranoia, sadness or depressed mood, anger, cognitive effects (e.g. confusion, loss of ego, loss of sanity, delusions, dissociation, depersonalization), perceptual effects (e.g. illusions), and physiological symptoms (e.g. increased heart rate, nausea/emesis, sympathetic system response). These categories are supported both by reviews of the safety of hallucinogens (Frecska and Luna, 2006; Johnson et al., 2008) and published clinical guidelines for the assessment and management of adverse reactions to hallucinogens (McCabe, 1977). While these categories of experience appear in aggregate in the literature, they may not all appear in any single instance of a challenging experience.

Current assessments of challenging experiences

The two most widely used questionnaire instruments assessing the subjective effects of classic hallucinogens are the Hallucinogen Rating Scale (HRS) (Strassman, 1995; Strassman et al., 1994) and the various forms of Dittrich's Altered States of Consciousness questionnaires (Dittrich, 1975, 1998; Studerus et al., 2010), including the Altered States of Consciousness Rating Scale (OAV) and the Five-Dimensional Altered States of Consciousness Rating Scale (5DASC). While some of the sub-scales of these instruments, and many items within these instruments, assess some aspect of cognitive, emotional, or physiological experiences that may be challenging, no questionnaire includes a comprehensive list of clearly defined sub-scales that separately measure the categories of challenging experience that are suggested by previous literature.

The HRS was developed explicitly as a means to quantify the subjective effects of classic hallucinogens (Strassman et al., 1994). Of the six sub-scales of the HRS (i.e. affect, cognition, intensity, perception, somaesthesia, and volition), one might hypothesize that the affect, cognition, and somaesthesia subscales might be most sensitive to challenging experiences. However, although the affect scale includes items that assess negative affect (e.g. anxious, frightened, panic), it also includes items that assess positive affect (e.g. euphoria, love, awe). Items in the cognition and somaesthesia scales include items that assess possibly challenging cognitions and sensations (e.g. "sense of chaos" or "nausea"), but they also include items that assess general but not necessarily "challenging" cognitions and sensations (e.g. "new thoughts" or "change in body temperature"). Thus, scores on the six scales of the HRS do not distinguish between very strong challenging experiences and very strong non-challenging experiences.

The OAV and 5DASC (and the preceding Abnormal Mental States (APZ) questionnaire) have been widely used in behavioral and neuroimaging research of pharmacologically (Gouzoulis-Mayfrank et al., 2005; Griffiths et al., 2006, 2011; Vollenweider et al., 1998) as well as non-pharmacologically induced altered states of consciousness (Hübner, 2007; Kjellgren et al., 2004; Walach and Kaseberg, 1998). The OAV sub-scale "dread of ego dissolution" (DED) covers a wide range of negative experiences, and is generally considered an overall "bad trip" scale (Studerus et al., 2010). This meta-scale of possible negative effects covers many (e.g. panic, loss of ego/control, feelings of insanity) but not all (e.g. sadness/grief/depression) possible categories of challenging experiences. The DED scale also averages responses from a number of proposed categories of experience (panic, loss of ego, insanity) rather than giving an individual score for each. Studerus and colleagues (2010) revealed a rescoring of the 5DASC that includes a separate scale for impaired control and cognition, and for anxiety. While these represent psychometrically justifiable subscales, these two sub-scales do not address shortcomings of the DED scale (e.g. they do not address the wide range of potential dimensions of challenging experience that are suggested by previous literature).

Many other instruments were either used or developed to assess various aspects of subjective experience of hallucinogens, including the Linton-Langs questionnaire (Linton and Langs, 1962), the Abramson questionnaire (Abramson, 1960), the Phenomenology of Consciousness Inventory (Pekala, 1986), the Minnesota Multiphasic Personality Inventory (Belleville, 1956), the States of Consciousness Questionnaire (Griffiths et al., 2006, 2011), and the Addiction Research Center Inventory (Griffiths et al., 2006, 2011; Haertzen, 1966; Riba et al., 2001). These

scales all suffer from at least one of the shortcomings of the 5DASC and HRS, namely: lack of specificity (i.e. lack of separable or independent scales that measure fine-grained facets of a challenging experience), lack of content coverage (missing a proposed dimension of challenging experience), or lack of sensitivity (i.e. not sensitive to the degree of challenge in an experience) to challenging experiences.

Why an assessment of dimensions of challenging experience is needed

With the recognition of the importance of set (participant psychological state) and setting (interpersonal and physical environment), acute adverse effects of psychedelics may generally be minimized and successfully managed in research contexts (Johnson et al., 2008; Metzner et al., 1965; Studerus et al., 2012). Guidelines have been described to minimize adverse reactions in controlled settings (Johnson et al., 2008), and incidence of lasting symptoms (including psychosis) related to administration of classic hallucinogens in controlled settings is quite low (Griffiths et al., 2006, 2011; Strassman, 1984; Studerus et al., 2011). However, challenging experiences can still occur in the presence of substantially controlled and supportive conditions (Griffiths et al., 2006, 2011; Studerus et al., 2011), and occurrence of acute and persisting adverse effects of psychedelics may be greater in uncontrolled settings, or with unscreened individuals (Carbonaro et al., 2016)

Little is known regarding the interdependence, causes, and consequences of any one of the individual categories of challenging experience suggested by the clinical and prior research literature. It is of interest to basic science and clinical application of classic hallucinogens to understand, for example, whether a challenging experience primarily involving panic and physiological distress might lead to persisting effects that differ from those that persist after a challenging experience primarily involving depression or confusion. If that is the case, it would be of value to know whether there are reliable indicators that could be used to predict whether an individual is more likely to experience panic or grief. Various facets of challenging experience may generally covary, but they may not necessarily be interdependent or collinear. Different facets of challenging experiences may lead to different outcomes.

Psychedelics offer a tool for studying the biological basis of emotional experiences using modern neuroscientific tools such as brain imaging. Anger, fear, and sadness have putatively different neural substrates as well as behavioral implications (Panksepp, 1998). To the extent that challenging experiences are encountered in brain imaging studies, it may be necessary to parse inter-individual differences in challenging experience during acute effects of psychedelics in order to be able to fully appreciate the neural basis of psychedelic experiences.

The field is lacking an empirically derived measurement instrument that can assess individual differences in challenging experiences, provide a "profile" of challenging experiences, and be applied to study predictors and persisting effects of challenging experiences. Such a fine-grained knowledge regarding the profile of challenging experiences and the consequences of these experiences will be of value both to optimize our understanding of the neurobiology of psychedelics and to optimize their use in treatment settings.

Aim of the current studies

In the following two studies, a Challenging Experience Questionnaire (CEQ) was constructed. Study 1 consisted of a psychometric analysis of responses to pooled items from a series of questionnaires (the 5DASC, HRS, and an additional measure, the States of Consciousness Questionnaire (SOCQ)) that have been used in an online survey of challenging experiences with psilocybin (Carbonaro et al., 2016) to establish, replicate, and validate the factor structure and factor scores of the CEQ in a pair of stratified sub-samples. In Study 2, administration of the CEQ as a stand-alone instrument in an online survey demonstrated both replication of the CEQ factor structure and validation of the CEQ factor scores.

Study 1

This study is a secondary analysis of data from an online survey of challenging experiences with psilocybin (Carbonaro et al., 2016). Participants in this survey completed a series of questionnaires in reference to a self-identified challenging experience. The survey sample was stratified into two demographically matched groups. Though the HRS, 5DASC, and SOCQ do not have sub-scales that measure all facets of challenging experience suggested by the clinical literature, they all have individual items that assess these facets. Thus, the combined items from the HRS, 5DASC, and SOCQ that address potential aspects of challenging experience form a desirable initial item pool for the development of the CEQ. Exploratory factor analysis was applied to items from the HRS, 5DASC, and SOCQ in the first stratum that could reasonably assess an aspect of challenging experience. The resulting factor structure was validated using confirmatory factor analysis in the second stratum.

Structural equation modeling was applied to the entire sample to demonstrate factorial invariance of the instrument across levels of gender. Clinical literature indicates that lasting psychosis related to challenging experiences may be present in those who have a personal or family history of psychological difficulty. Therefore, factorial invariance of the instrument was also assessed across levels of struggle with a psychiatric disorder that preceded the reported challenging experience. Ratings of the meaningfulness, spiritual significance, and difficulty of the experience, as well as ratings of the impact of the experience on a person's overall well-being, have been used in the past to assess the overall impact of psilocybin experiences on an individual (Griffiths et al., 2006, 2011). Therefore, ratings of the overall impact of the challenging experience in this sample were regressed on CEQ factor scores to assess the relationship between these outcome measures and aspects of challenging experience. A description of measures analyzed in the current report is provided below. A more detailed description of the complete methods, measures, and initial findings of the online survey can be found in the original report (Carbonaro et al., 2016).

Method

Participants

The Bad Trip Survey (Carbonaro et al., 2016) was completed by 2085 participants. Participants were recruited via Internet

advertisements, email invitation, and word of mouth. An Internet link to the survey was posted on websites, such as Erowid (an online information library on psychoactive substances, www.erowid.com), that are frequented by individuals interested in hallucinogens. Participants were not provided compensation for their responses. Inclusion criteria consisted of reading, writing, and speaking fluency in English, having ingested an active dose of psilocybin mushrooms that produced moderate to strong psychoactive effects, having had a difficult or challenging experience (i.e. "bad trip") after ingesting psilocybin mushrooms, age of 18 years or older at the time of completing the survey, and age of between 18–70 years old at the time of the reported challenging experience. Participants who did not meet inclusion criteria after providing demographic information were directed to an abbreviated version of the survey, and their data were not included in the analysis. This approach was taken to obscure to participants the fact that they were excluded, and therefore discourage them from re-taking the survey with false responses in an attempt to be included. Participants were also asked to refrain from completing the survey more than once. Data from completed participants were excluded if they reported that their challenging experience was (a) encountered in the context of a research study, (b) experienced by another person, or (c) attributed to another substance in addition to psilocybin. Participants were also excluded if free-response comments provided at the end of the survey raised concerns about the validity of their reports. Data for 92 completing participants were excluded, yielding a sample of 1993 participants.

Measures

Participants completed an online survey consisting of the below-referenced questionnaires (the HRS, 5-DASC, and SOCQ) among a series of other questions regarding demographics, substance use behaviors, and details of the participant's reported challenging experience. A more detailed description of the full complement of measures administered in the online survey can be found in the original report (Carbonaro et al., 2016). We identified 64 items from the HRS, 5-DASC, and SOCQ (listed in Supplementary Material, Appendix 2) that unambiguously assessed a challenging aspect of experience with classic hallucinogens, and treated these 64 items, as worded and responded to in their original form in the HRS, 5-DASC, and SOCQ, as an initial item pool for the construction of the initial form of the CEQ.

Additional items (including questions regarding the overall impact of the challenging experience, and previous struggle with a psychiatric disorder) were used in regression and factorial invariance analyses, respectively, after the generation of the CEQ. Finally, demographic variables were assessed and used to describe the sample and also to generate two demographically matched strata for separate analyses within this study.

HRS. The HRS is a 99-item instrument that is rated on a five-point scale (1 – "not at all", 2 – "slightly", 3 – "moderately", 4 – " very much", 5 – "extremely"). It assesses the greatest degree to which a respondent encountered different subjective effects during the course of an experience with a specified drug. It consists of six subscales assessing general dimensions of subjective experience (intensity, somaesthesia, affect, perception, cognition, and volition). Twenty-seven items (spanning all six scales of the HRS) that were judged by the authors to assess a potentially challenging aspect of experience with classic hallucinogens were retained for the initial item pool for the CEQ.

5DASC. The 5DASC (Dittrich et al., 2010; Studerus et al., 2010) is the latest version of the Altered States of Consciousness (APZ) questionnaire (Dittrich, 1975), which was designed to assess subjective aspects of a wide range of altered states of consciousness, including but not limited to those occasioned by classic hallucinogens. The 5DASC consists of 94 items that are rated on a 20-point rating scale with two anchors: 1 – "No, not more than usually" and 20 – "Yes, much more than usually". A recent psychometric analysis identified 11 plausible sub-scales of the 5DASC, which include the impaired cognition and control (ICC), and anxiety (ANX) scales (Studerus et al., 2010). The 13 items of the 5DASC that constitute the ICC and ANX sub-scales were retained for the initial item pool for the CEQ. While the 5DASC was originally developed in German and has not been empirically validated in English, this does not undermine the use of items from this questionnaire in the current study, as these items contribute to an initial item pool for validation of a new questionnaire (the CEQ).

SOCQ. The SOCQ is a 100-item questionnaire that is rated on a six-point scale (0 – "none; not at all", 1 – "so slight cannot decide", 2 – "slight", 3 – "moderate", 4 – "strong (equivalent in degree to any previous strong experience or expectation of this description)", 5 – "extreme (more than ever before in my life and stronger than 4)"). The SOCQ contains 43 items from the Mystical Experiences Questionnaire (MEQ43) (Griffiths et al., 2006, 2011; Pahnke et al., 1969; Richards et al., 1977), which was developed to assess several domains of mystical experience. The other 57 items of the SOCQ were based on a wide range of possible subjective effects of classic hallucinogens that were suggested by a sample of clinicians, but these items have not previously been analyzed, and they are typically treated as distractor items. Twenty-four of these distractor items were identified and retained for the initial item pool for the CEQ. These 24 distractor items assess potentially challenging aspects of experiences with classic hallucinogens (such as emotional, social, and physical discomfort, pain, and suffering, disorientation, ego loss, loss of perception of time, isolation, and confusion). No items from the MEQ43 were included in the present analysis. While these items have not been a part of an independently validated questionnaire, this does not undermine the use of these items to contribute to an initial item pool for development and validation of the CEQ.

Questions regarding the overall impact of the challenging experience. These questions comprise a set of items that have been administered in laboratory studies of the acute and persisting effects of psilocybin (Griffiths et al., 2006, 2011; Johnson et al., 2014). Participants were asked to report on the meaningfulness and difficulty of their reported psilocybin experience using the following response options: "no more than routine, everyday experiences", "similar to (difficult or challenging/meaningful) experiences that occur on average once a week", "similar to (difficult or challenging/meaningful) experiences that occur on average once a month", "similar to (difficult or challenging/meaningful) experiences that occur on average once a year",

"similar to (difficult or challenging/meaningful) experiences that occur on average once every 5 years", "among the top 10 most (difficult or challenging/meaningful) experiences of my life", "among the top five most (difficult or challenging/meaningful) experiences of my life", and "the single most (difficult or challenging/meaningful) experience of my life". Participants indicated the degree to which the experience was spiritually significant to them, using the following rating scale: "not at all", "slightly", "moderately", "very much", "among the top five most spiritually significant experiences of my life", and "the single most spiritually significant experience of my life". Participants also responded to a question inquiring about the effect of the challenging experience on their well-being or life satisfaction ("Do you believe that the experience and your contemplation of that experience have led to a change in your current sense of personal well-being or life satisfaction?") using the following rating scale: "Increased very much (+3)", "Increased moderately (+2)", "Increased slightly (+1)", "No change (0)", "Decreased slightly (-1)", "Decreased moderately (-2)", "Decreased very much (-3)". These inquiries conform to previous methods of measuring participants' assessment of the overall impact of a psilocybin session (Griffiths et al., 2006, 2011). Meaningfulness, spirituality, difficulty, and well-being were regressed on latent variable (i.e. factor) scores derived from the CEQ.

Previous struggle with a psychiatric disorder. Participants were asked to indicate whether they had struggled with a psychiatric disorder at some point before their reported challenging experience. Participants were then able to separately endorse having struggled with the following disorders ("Please select all psychiatric disorders you have struggled with:"): anxiety, depression, substance use disorder, or "other psychiatric disorder (e.g. schizophrenia)". This variable was coded "1" to indicate any endorsement of previous struggle with a psychiatric disorder, or "0" to indicate no endorsement of previous struggle with a psychiatric disorder. This coded variable was used to indicate group membership in later factorial invariance analyses.

Demographic questions. These included age, gender, education, race, and total incidence of past hallucinogen use. Age was originally coded as interval data and was converted to an ordinal category variable, with the following five categories:<25, 25–34, 35–44, 45–54, 55–64, and ≥65 years. Participants used the following categories to rate the total number of times of past hallucinogen use: 1, 2–5, 6–10, 11–20, 21–50, 51–100, 101–300, >300. Gender, education, and race were coded as categorical data.

Analysis

Matched exploratory and confirmatory analysis strata. The sample of 1993 participants was stratified into two sub-samples that were matched on age, sex, education, race, and total incidence of past hallucinogen use. Stratification was carried out with the *strata* function of the *sampling* package (Tillé and Matei, 2012) in R (R Core Team, 2012), using simple random sampling without replacement. Before stratification, frequencies were calculated for each cell in the factorial model assumed by the stratification variables to verify that there were no cells with fewer than two observations. A total of 226 observations were removed in this procedure, yielding a final total analysis sample of 1767 (exploratory stratum n=833, confirmatory stratum n=934). After stratification, demographic data were compared between strata using a two-sample t-test for age, and chi-squared independence tests for all other demographic data. Demographic data for the strata are presented in the left-hand portion of Table 1. No variables differed significantly between strata.

Initial scale construction. The first stratum (n=833) was used to identify the latent structure of the items of the CEQ, using item analysis and exploratory factor analysis. Distributional properties of items were assessed to identify and remove items with restricted range or heavily non-normal distribution. The *mixed.cor* function from the *psych* package in R (Revelle, 2013) was then used to estimate the correlations between the remaining response variables in the exploratory stratum. The *mixed.cor* function calculates Pearson product moment correlations between continuous variables, polychoric correlations between polytomous variables, tetrachoric correlations between dichotomous variables, and polyserial or biserial correlations between mixed variables, thus generating a "mixed" correlation matrix. Exploratory factor analysis was conducted on this mixed correlation matrix using the *fa* function in the *psych* toolbox in R (Revelle, 2013), with maximum likelihood factor extraction and oblimin rotation. Latent variables that contribute to the concept of a challenging experience may reasonably be expected to co-vary. Therefore, an oblique factor rotation (oblimin) was chosen over an orthogonal rotation. Visual inspection of the scree plot and parallel analysis with the *fa.parallel* function in the *psych* toolbox in R was used to determine the number of factors to extract.

An initial exploratory factor analysis was fit to the data to identify common factors. Items with no loading above 0.4 on any factor were discarded. The remaining items were entered into iterative scale analyses. An initial scale was created for each factor from the items that loaded most strongly onto that factor. For each scale, at each iteration, a parallel analysis was conducted on the surviving items for that scale to assess dimensionality of the scale. When more than two underlying dimensions were identified for a given scale, McDonald's omega was used to estimate the reliability of the scale. Otherwise, Cronbach's alpha was used to calculate scale reliability. Item-total correlations were calculated at each iteration, as well as change in scale reliability with removal of each item. Items were removed from a scale if they demonstrated low item-total correlation (below 0.4) and if their removal increased scale reliability and average item-total correlation for the scale. The final remaining items were entered into an exploratory factor analysis to determine the initial model for the CEQ.

Model replication. The resulting exploratory factor model was replicated using confirmatory factor analysis in the second stratum (n=934). Confirmatory factor analysis in a separate sample is a conservative test of an exploratory factor model, and it establishes reliability and internal validity of the structure of the CEQ. A combination of fit indices was used to assess confirmatory factor model fit, including the comparative fit index (CFI) (Bentler, 1990) the standardized root mean square residual (SRMR) (Hu and Bentler, 1999), and the root mean square error of approximation (RMSEA) (Browne and Cudeck, 1993). Values of SRMR and RMSEA<0.1, and CFI>0.90 (Browne and Cudeck, 1993; Hu

Table 1. Comparison of demographic variables between stratification groups and studies.

n	Study 1 strata		*t*	df	*p*	Study 1	Study 2	*t*	df	*p*
	Exploratory	Confirmatory				Total	Total			
	833	934				1767	981			
Age			0.8791	1765	0.3794			6.3607	1763	<0.0001
Mean	28.77	29.13				28.96	31.38			
Std	8.36	8.71				8.55	10.06			
			χ^2	df	*p*			χ^2	df	*p*
Education			0.7462	5	0.9803			29.7703	5	<0.0001
<HS	1.3%	1.5%				1.4%	2.9%			
HS diploma	11.0%	11.1%				11.1%	9.8%			
Some college	37.7%	36.9%				37.3%	32.4%			
College	26.7%	26.0%				26.3%	23.5%			
Some grad	6.6%	7.5%				7.1%	10.9%			
Grad/prof	16.7%	16.9%				16.8%	20.5%			
Sex			0.4888	1	0.4845			20.2475	1	<0.0001
Male	81.5%	80.1%				80.8%	73.3%			
Female	18.5%	19.9%				19.2%	26.7%			
Race			0.7345	4	0.9470			48.8493	4	<0.0001
Native American	0.5%	0.5%				0.5%	1.3%			
Asian	0.4%	0.3%				0.3%	1.7%			
White	94.8%	94.0%				94.4%	87.2%			
Other	0.5%	0.6%				0.6%	2%			
Multiple	3.8%	4.5%				4.2%	7.8%			
Ethnicity			0.0096	1	0.9219			21.5225	1	<0.0001
Hispanic	4.6%	4.4%				4.5%	9.1%			
Non-Hispanic	93.6%	94.6%				94.2%	90.9%			
Incidence of serotonergic hallucinogen use			0.6211	7	0.9989			54.2779	7	<0.0001
1×	3.2%	3.6%				3.5%	2.2%			
2–5	18.4%	17.5%				17.9%	11.3%			
6–10	19%	19.3%				19.1%	16.2%			
11–20	20.6%	20.1%				20.4%	18.5%			
21–50	19.7%	19.7%				19.7%	22.7%			
51–100	9.5%	9.7%				9.6%	13.7%			
101–300	7.2%	7.5%				7.4%	10.5%			
>300	2.4%	2.6%				2.5%	4.8%			
Incidence of psilocybin mushroom use			2.7610	7	0.9062			38.2196	7	<0.0001
1×	6.6%	5.9%				6.2%	4.1%			
2–5	30.9%	31.9%				31.4%	25.5%			
6–10	23.3%	24.1%				23.7%	21.7%			
11–20	19.0%	17.6%				18.2%	20.4%			
21–50	12.5%	12.3%				12.4%	14.7%			
51–100	4.4%	5.5%				5.0%	7.8%			
101–300	2.9%	2.2%				2.5%	4.5%			
>300	0.5%	0.5%				0.5%	1.2%			
Relative psychological difficulty of the reported bad trip			11.0153	7	0.1380			14.1258	7	0.0490
Routine, everyday experience	2.3%	1.3%				1.8%	1.4%			
Similar to weekly occurrences	1.2%	0.6%				0.9%	1.7%			
Similar to monthly occurrences	6.6%	4.5%				5.5%	4.7%			
Similar to yearly occurrences	13.0%	14.3%				13.7%	13.6%			
Similar to 5-year occurrences	14.2%	16.5%				15.4%	18.9%			
Top 10 most challenging	23.4%	23.2%				23.3%	21.1%			
Top 5 most challenging	27.6%	29.2%				28.5%	29.8%			
Single most challenging	11.8%	10.3%				11.0%	8.8%			
Would repeat the reported experience, including the challenging portion			0.2740	1	0.6007			8.1741	1	0.0042
No	54.1%	52.8%				53.4%	52.3%			
Yes	45.9%	47.2%				46.6%	47.7%			

Grad: graduate school; prof: professional school; Std: standard deviation; HS: high school.

and Bentler, 1999) indicate acceptable model fit. Consideration of a combination of fit indices, with "good fit" values of SRMR<0.09, and CFI >0.90, have been shown to minimize both Type I and Type II errors (Hu and Bentler, 1999).

Factorial invariance. Factorial invariance analysis is a formal test of whether a set of variables has a similar (or "invariant") factor structure in different groups (i.e. that they show "factorial invariance" across groups). Factorial invariance analysis is conducted by performing a series of multiple-group confirmatory factor analyses, adding more constraints in each sequential model for parameters to be equal across groups and then testing to see whether model fit improves or degrades across sequential models (Kline, 2005; Lomax, 1983; Widaman and Reise, 1997). Configural factorial invariance sets the number of factors between groups to be equal, with the same items loading onto the same factors in each group. Weak factorial invariance adds the additional constraint of the factor loading parameter estimates of each item to be equal across groups. Strong factorial invariance additionally constrains the intercepts of the observed variables to be equal across groups. Strict factorial invariance additionally constrains measurement residuals to be equal across groups.

Demonstrating strong or strict factorial invariance provides evidence that the CEQ is measuring constructs in a similar fashion in each group being compared (men vs women, or having vs not having previous struggle with a psychiatric disorder). Establishing strong or strict factorial invariance also allows direct comparison of latent variable means between groups. This is accomplished by setting the latent variable means for the comparison group to zero, and estimating the latent variable means for the other group. This identifies the model and standardizes latent variable mean estimates for the other group. Latent variable means in the other group that are significantly different from zero indicate that latent variables are significantly different from those in the comparison group.

Factorial invariance was assessed in a series of multiple-group confirmatory factor models using the *cfa* function in the *lavaan* toolbox in R (Rossell, 2012), with maximum likelihood estimation. Model fit in factorial invariance models was assessed using a combination of change in the CFI, SRMR, and RMSEA. These fit indices have been shown through simulation to be sensitive to both measurement invariance and lack of measurement invariance at the three levels (factor loadings, intercepts, and residuals, or weak, strong, and strict invariance) that are tested within the current sample (Chen, 2007). Decrease in CFI>0.01, increase in RMSEA>0.015, and increase in SRMR>0.01 between levels of factorial invariance (i.e. change between modeling steps) indicates noninvariance (Chen, 2007).

CEQ and overall impact of the challenging experience. The confirmatory factor model for the entire sample was then extended to a structural regression model with a measurement component to test the relationship between latent variable scores of the CEQ and ratings of the overall impact of the challenging experience (difficulty of the experience, meaningfulness of the experience, spiritual significance of the experience, and effect of the experience on well-being). Regression of CEQ factor scores on ratings of difficulty of the experience provides evidence for the convergent validity of the CEQ. Structural equation modeling was conducted using the *sem* function in the *lavaan* toolbox in R (Rossell, 2012), with maximum likelihood estimation.

Results

Initial scale construction

Descriptive statistics were calculated for each of the 64 potential CEQ items in the exploratory stratum, and distributional properties of each item (skew and kurtosis) were inspected for each item. Nine items had a skew or kurtosis greater than 1.5. After removing the nominally skewed and kurtotic items, the response distributions of remaining items were visually inspected to verify that the full range of responses had been utilized in each item. The following procedures were repeated after returning these nominally skewed and kurtotic items to the item pool, and there was no change in the outcome.

An initial exploratory factor model was estimated to identify both a general factor structure for the response data and the best-loading items on each factor. A Scree plot and parallel analysis indicated that a six-factor solution is appropriate for the dataset. The seven-factor solution yielded one clear junk factor, and the five-factor solution did not yield a very clear qualitative interpretation. The six-factor solution yielded factors with a cohesive qualitative interpretation: physiological distress (factor 1), grief (factor 2), fear (factor 3), insanity (factor 4), isolation (factor 5), and death (factor 6). The six-factor solution was used as the basis of further item removal. Items with no loading above 0.4 on any factor were discarded. A total of 19 items were removed in this step. Items were removed from a scale if they demonstrated low item-total correlation (below 0.4) and if their removal increased scale reliability and average item-total correlation for the scale. An additional 12 items were removed in this step, yielding a final set of 24 items.

An exploratory factor analysis was conducted on the 24 remaining items. The final exploratory factor structure, factor intercorrelations, and factor reliabilities are reported in Table 2. The factors in this solution explained a cumulative 67% of variance in the observed data. Measures of sampling adequacy, including the Kaiser–Meyer–Olkin test (KMO=0.92) and Bartlett's test of sphericity (χ^2=44,899, df=25, p<0.0001), indicated suitability of the data for factor analysis.

Addition of a paranoia factor. Clinical literature and anecdotal reports indicate that paranoia is a subjective experience frequently represented in challenging experiences with classic hallucinogens (Cohen, 1960; Griffiths et al., 2006, 2011; Strassman, 1984). While a clear paranoia factor did not emerge from the exploratory factor analyses, the initial item pool for construction of the CEQ did contain two items directly related to paranoia (SOCQ item 40 "Feeling that people were plotting against you" and SOCQ item 72 "Experience of antagonism toward people around you"). Interestingly, there are no items in either the HRS or the 5DASC that directly assess paranoia. When conducting confirmatory factor analysis of the second stratum in the current sample, the identified 24-item, six-factor CEQ model was amended by adding the two paranoia items, both loading onto a seventh factor.

Model replication

A mixed correlation matrix was calculated for the confirmatory stratum, using the items from the final exploratory factor model

Table 2. Exploratory factor structure, correlations, and reliabilities in Study 1.

Scale	Item	Fear	Grief	Physical distress	Insanity	Isolation	Death	h2
HRS	26	**0.95**	–0.02	–0.05	–0.01	0.02	0.04	0.88
HRS	27	**0.84**	0.06	0.04	0.07	–0.06	–0.04	0.79
SOCQ	52	**0.72**	0.06	0.00	0.04	0.04	0.12	0.72
HRS	25	**0.58**	0.06	0.19	0.05	0.10	–0.19	0.54
5DASC	63	**0.55**	0.03	0.09	0.02	0.09	0.07	0.48
HRS	39	–0.06	**0.80**	–0.03	–0.03	0.03	–0.04	0.58
HRS	42	0.14	**0.72**	–0.06	0.08	0.09	–0.01	0.78
SOCQ	91	0.02	**0.72**	0.10	–0.07	–0.01	0.12	0.59
HRS	43	–0.02	**0.66**	0.14	–0.04	–0.06	0.00	0.42
SOCQ	16	0.10	**0.61**	–0.05	0.21	0.12	–0.01	0.71
SOCQ	13	0.17	**0.59**	0.01	0.02	0.01	0.10	0.57
HRS	12	0.01	0.02	**0.85**	0.01	–0.05	0.02	0.74
HRS	13	0.04	–0.09	**0.81**	0.04	0.05	–0.03	0.68
HRS	11	–0.06	0.08	**0.67**	0.11	0.02	0.02	0.52
HRS	10	0.00	0.09	**0.64**	0.07	0.06	–0.03	0.50
HRS	9	0.06	–0.01	**0.63**	–0.16	0.00	0.12	0.41
SOCQ	85	0.01	–0.01	0.03	**0.93**	–0.03	0.02	0.90
HRS	92	–0.03	0.05	0.05	**0.75**	0.02	0.03	0.63
5DASC	19	0.11	–0.06	–0.05	**0.68**	0.05	0.03	0.57
5DASC	44	0.03	–0.07	0.01	–0.01	**0.93**	0.01	0.81
HRS	48	–0.01	0.07	0.02	0.00	**0.83**	–0.01	0.74
SOCQ	45	–0.04	0.32	–0.03	0.11	**0.54**	0.08	0.67
SOCQ	70	–0.02	–0.01	–0.02	0.02	0.00	**1.01**	1.00
HRS	74	0.12	0.07	0.11	0.05	0.03	**0.70**	0.74
Fear		*0.88*						
Grief		0.51	*0.89*					
Physical distress		0.36	0.23	*0.87*				
Insanity		0.68	0.40	0.34	*0.84*			
Isolation		0.43	0.55	0.17	0.43	*0.85*		
Death		0.34	0.31	0.23	0.38	0.22	*0.84*	

5DASC: Five-Dimensional Altered States of Consciousness Rating Scale; HRS: Hallucinogen Rating Scale; SOCQ: States of Consciousness Questionnaire.
The upper matrix presents loadings of items onto exploratory factors for the final item set. Factor loadings greater than 0.4 are in bold. The h2 column shows the communality of each item. The lower matrix includes factor intercorrelations below the diagonal in non-italicized typeface, and factor reliabilities on the diagonal in italics.

(Table 2) as well as two additional items loading onto a seventh factor for paranoia. A confirmatory factor model was fit to this correlation matrix, setting positive loadings of each item onto its intended factor, with all other item loadings equal to zero. The model was identified, and factor loadings standardized, by setting the variance of each latent factor to one.

An initial model was fit including the 24-item six-factor structure identified in exploratory analyses of the first stratum. Fit indices for this model (RMSEA=0.070 (90% confidence interval (CI): 0.067–0.074), SRMR=0.054, CFI=0.913) indicate acceptable model fit (Hu and Bentler, 1999). A subsequent model was then fit that also included two additional items (SOCQ items 40 and 72) that loaded onto a seventh factor for paranoia. Fit indices for this model (RMSEA=0.066 (90% CI: 0.063–0.070), SRMR=0.052, CFI=0.912) also indicate acceptable model fit (Hu and Bentler, 1999). Fit indices, as well as factor loadings, factor correlations, and factor reliabilities for the first six factors did not differ substantially from the original model. This confirms both the initial 24-item six-factor structure of the CEQ, and supports inclusion of a seventh factor for paranoia. Factor structure for the 26-item, seven-factor confirmatory model is presented in Table 3. Factor correlations and factor reliabilities for this model are presented in the top portion of Table 4.

Factorial invariance of the CEQ

Factorial invariance was tested separately for two levels of sex and for two levels of previous struggle with a psychiatric disorder (having vs not having had previous struggle). Model fit indices for consecutive factorial invariance models are presented in Table 5. Fit indices showed negligible change between all levels of factorial invariance for sex and previous struggle with a psychiatric disorder, which indicates strict factorial invariance between categories in both sets of groups. Factor correlations for each level of gender and previous struggle with a psychiatric disorder, and differences in factor scores between levels, are presented in the middle and lower portion of Table 4.

Overall impact of the challenging experiences

Model fit indices for the structural equation model regressing overall impact ratings on CEQ factor scores indicated good

Table 3. Confirmatory factor structure of the Challenging Experience Questionnaire in Study 1.

Factor	Scale	Item	Estimate	SE	Item text
Fear					
	HRS	26	0.903	0.026	Frightened
	HRS	27	0.853	0.027	Panic
	SOCQ	52	0.797	0.028	Experience of fear
	HRS	25	0.719	0.029	Anxious
	5DASC	63	0.667	0.030	I had the feeling something horrible would happen
Grief					
	HRS	36	0.690	0.030	Sad
	SOCQ	91	0.650	0.030	Feelings of grief
	HRS	38	0.841	0.027	Despair
	HRS	39	0.636	0.031	Feel like crying
	SOCQ	16	0.801	0.028	Feelings of despair
	SOCQ	13	0.704	0.030	Emotional and/or physical suffering
Physical distress					
	HRS	12	0.618	0.032	Feel heart beating
	HRS	13	0.505	0.033	Feel heart skipping beats or beating irregularly
	HRS	11	0.826	0.029	Feel body shake/tremble
	HRS	10	0.838	0.029	Shaky feelings inside
	HRS	09	0.526	0.033	Pressure or weight in chest or abdomen
Insanity					
	SOCQ	85	0.853	0.028	Fear that you might lose your mind or go insane
	HRS	88	0.813	0.028	Change in sense of sanity
	5DASC	19	0.749	0.029	I was afraid that the state I was in would last forever
Isolation					
	5DASC	44	0.819	0.029	I felt isolated from everything and everyone
	HRS	44	0.816	0.029	Feel isolated from people and things
	SOCQ	45	0.768	0.029	Experience of Isolation and loneliness
Death					
	SOCQ	70	0.736	0.033	Profound experience of your own death
	HRS	70	0.986	0.034	Feel as if dead or dying
Paranoia					
	SOCQ	40	0.677	0.040	Feeling that people were plotting against you
	SOCQ	72	0.711	0.040	Experience of antagonism toward people around you

5DASC: Five-Dimensional Altered States of Consciousness Rating Scale; HRS: Hallucinogen Rating Scale; SE: standard error (of the parameter estimate); SOCQ: States of Consciousness Questionnaire.
The table presents confirmatory factor loadings, in which items were only allowed to load onto their intended factor. All other loadings were set to zero. Scale=parent instrument (SOCQ, HRS, or 5DASC) for a given item; Item=item number from parent instrument; estimate=standardized factor loading.

fit (CFI=0.915, RMSEA=0.060 (90% CI: 0.058–0.062), SRMR=0.045). Regression estimates are presented in Table 6. Scores on the CEQ death factor were associated with a decrease in wellbeing attributed to the experience, and were positively associated with all other ratings of overall impact. Scores on fear were positively associated with ratings of the difficulty of the experience and change in wellbeing, and were negatively associated with the spiritual significance and meaningfulness of the experience. Isolation scores were positively associated with an increase in wellbeing and negatively associated with meaningfulness and spiritual significance of the experience. Scores on both insanity and grief were positively associated with rated difficulty of the experience, meaningfulness and spiritual significance of the experience. Physical distress was positively associated with ratings of spiritual significance and associated with a decrease in wellbeing. Factor scores on paranoia were not associated with any overall impact ratings.

Study 2

The items comprising the CEQ in Study 1 were taken from three separate and extensive instruments (the HRS, with 104 items total, the SOCQ with 100 items total, and the 5DASC with 42 items), each with a different response format. In Study 2, the 26-item CEQ was validated as a stand-alone instrument (i.e. outside of the context of the other items in the HRS, SOCQ, and 5DASC). Item text from the HRS and 5-DASC items was altered to match the prose of the items from the SOCQ, and the six-item SOCQ response format was adopted for all items of the CEQ. Side-by-side comparison of original and altered items is presented in Supplementary Material, Appendix 2. Items of the stand-alone CEQ were administered along with demographic measures and measures of the overall impact of the reported challenging experience used in Study 1 and analyzed in order to replicate the factor structure and regression results obtained in Study 1.

Table 4. Factor correlations, reliabilities, and means for the entire sample, by gender, and by previous struggle with a psychiatric disorder in Study 1.

	Fear	Grief	Physical distress	Insanity	Isolation	Death	Paranoia
Factor correlations and reliabilities for the entire sample							
Fear	*0.89*						
Grief	0.685 (0.016)	*0.87*					
Physical distress	0.469 (0.022)	0.363 (0.024)	*0.80*				
Insanity	0.730 (0.015)	0.535 (0.021)	0.410 (0.024)	*0.84*			
Isolation	0.534 (0.020)	0.677 (0.017)	0.291 (0.026)	0.498 (0.022)	*0.84*		
Death	0.456 (0.020)	0.421 (0.021)	0.347 (0.023)	0.473 (0.020)	0.315 (0.023)	*0.84*	
Paranoia	0.375 (0.028)	0.377 (0.029)	0.247 (0.031)	0.327 (0.030)	0.513 (0.027)	0.143 (0.029)	*0.65*
By gender							
Standardized factor scores							
M[a]	0	0	0	0	0	0	0
F	0.156 (0.077)[b]	0.059 (0.059)	0.066 (0.054)	0.087 (0.102)	0.710 (0.375)	–0.069 (0.089)	0.105 (0.084)
Factor correlations and reliabilities for females (above the diagonal) and males (below the diagonal)							
Fear		0.718 (0.032)	0.467 (0.050)	0.708 (0.034)	0.537 (0.045)	0.449 (0.050)	0.313 (0.061)
Grief	0.675 (0.018)		0.426 (0.052)	0.613 (0.042)	0.694 (0.035)	0.455 (0.050)	0.362 (0.061)
Physical distress	0.471 (0.025)	0.345 (0.028)		0.431 (0.054)	0.390 (0.055)	0.425 (0.053)	0.184 (0.067)
Insanity	0.735 (0.016)	0.513 (0.024)	0.405 (0.027)		0.531 (0.048)	0.486 (0.051)	0.251 (0.065)
Isolation	0.533 (0.023)	0.672 (0.019)	0.266 (0.030)	0.489 (0.025)		0.363 (0.054)	0.518 (0.058)
Death	0.464 (0.022)	0.416 (0.024)	0.333 (0.026)	0.475 (0.023)	0.306 (0.026)		0.117 (0.063)
Paranoia	0.396 (0.032)	0.379 (0.032)	0.259 (0.035)	0.343 (0.034)	0.503 (0.031)	0.154 (0.033)	
By previous struggle with a psychiatric disorder							
Standardized factor scores							
Y[a]	0	0	0	0	0	0	0
N	–0.190 (0.060)[c]	–0.223 (0.047)[c]	–0.109 (0.042)[d]	–0.229 (0.082)[d]	–0.747 (0.299)[b]	–0.144 (0.071)[b]	–0.110 (0.068)
Factor correlations and reliabilities for those with (above the diagonal) and without (below the diagonal) previous struggle							
Fear		0.691 (0.025)	0.443 (0.036)	0.655 (0.028)	0.547 (0.032)	0.435 (0.035)	0.352 (0.046)
Grief	0.676 (0.020)		0.312 (0.041)	0.486 (0.036)	0.692 (0.027)	0.419 (0.036)	0.369 (0.047)
Physical distress	0.481 (0.028)	0.388 (0.031)		0.354 (0.041)	0.225 (0.044)	0.315 (0.039)	0.071 (0.052)
Insanity	0.773 (0.017)	0.562 (0.025)	0.435 (0.030)		0.491 (0.036)	0.458 (0.036)	0.282 (0.050)
Isolation	0.523 (0.026)	0.666 (0.022)	0.322 (0.033)	0.498 (0.028)		0.304 (0.039)	0.479 (0.045)
Death	0.463 (0.028)	0.414 (0.029)	0.355 (0.030)	0.472 (0.028)	0.313 (0.031)		0.125 (0.047)
Paranoia	0.391 (0.036)	0.378 (0.037)	0.348 (0.038)	0.355 (0.038)	0.530 (0.034)	0.150 (0.038)	

Factor reliabilities for the entire sample (calculated as Cronbach's alpha) are presented in italics on the diagonal at the top of the table. Correlations and factor scores are presented with standard error in parentheses.
[a]Comparison group (males, and those with previous struggle with a psychiatric disorder); [b]$p<0.05$; [c]$p<0.001$; [d]$p<0.01$.

Method

Participants

Altogether 1052 participants were recruited through word of mouth and online advertisement (in a fashion similar to Study 1), and completed an online survey (separate from the survey in Study 1) exploring challenging experiences with psilocybin. Participants were informed on the first page of the survey to not continue if they had previously completed a Johns Hopkins survey of "bad trips" or challenging experiences with psilocybin. Participants were asked on the second page of the survey to confirm that they had not previously completed this or any similar survey. Inclusion and exclusion criteria for this survey were identical to those used in Study 1. Participants were not compensated for their participation. Seventy-one completing participants were excluded based on the criteria specified in Study 1, which yielded a final dataset of 981 participants.

Measures

Demographic questions and questions regarding the overall impact of the challenging experience utilized in Study 1 were also presented in this study. Demographics for participants in Study 2 are presented in the right-hand portion of Table 1. Demographic data were compared between Study 1 and Study 2 using a two-sample *t*-test for age, and chi-squared independence tests for all other demographic variables.

CEQ. Twenty-six items identified in Study 1 (including two paranoia items) were presented in fixed pseudorandom order, with the same response format for each item ("0 – none; not at all"; "1 – so slight cannot decide"; "2 – slight"; "3 – moderate"; "4 –strong"; "5 – extreme (more than ever before in my life)"). Since the original items from the 5-DASC, HRS, and SOCQ differed between scales in style and prose, modifications were made to the text of the 5-DASC and HRS items in order to better match the prose of these items in the stand-alone CEQ. Side-by-side comparison of original

Table 5. Model fit indices for tests of factorial invariance.

Model	CFI	ΔCFI	RMSEA	ΔRMSEA	SRMR	ΔSRMR
Gender in Study 1						
1 (configural)	0.911		0.066		0.048	
2 (weak)	0.910	−0.001	0.065	−0.001	0.050	0.002
3 (strong)	0.906	−0.004	0.066	+0.001	0.051	0.001
4 (strict)	0.905	−0.001	0.065	−0.001	0.051	0.000
Struggle with a psychiatric disorder in Study 1						
1 (configural)	0.914		0.065		0.047	
2 (weak)	0.914	0.000	0.063	−0.002	0.049	0.002
3 (strong)	0.914	0.000	0.063	0.000	0.049	0.000
4 (strict)	0.913	−0.001	0.061	−0.002	0.049	0.000
Comparing Study 1 and Study 2						
1 (configural)	0.913		0.066		0.047	
2 (weak)	0.912	−0.001	0.065	−0.001	0.050	+0.003
3 (strong)	0.912	0.000	0.064	−0.001	0.050	0.000
4 (strict)	0.902	−0.010	0.066	+0.002	0.048	−0.002

CFI: comparative fit index; RMSEA: root mean square error of approximation; SRMR: standardized root mean square residual.
Critical values for the rejection of the null hypothesis of factorial invariance at any level is decrease in CFI>0.01, increase in RMSEA>0.015, and increase in SRMR>0.01 (Chen, 2007). No changes in model fit indices exceeded these values.

Table 6. Regression of ratings of overall impact of the experience on Challenging Experience Questionnaire (CEQ) factor scores for Study 1.

CEQ factor	PE	SE	*z*	*p*
Difficulty of the experience				
Fear	0.161	0.043	3.747	**0.000**
Grief	0.132	0.040	3.334	**0.001**
Physical distress	−0.038	0.026	−1.462	0.144
Insanity	0.276	0.038	7.198	**0.000**
Isolation	0.009	0.038	0.228	0.820
Death	0.156	0.027	5.778	**0.000**
Paranoia	−0.085	0.032	−2.626	**0.009**
Meaningfulness of the experience				
Fear	−0.240	0.049	−4.954	**0.000**
Grief	0.162	0.045	3.629	**0.002**
Physical distress	0.021	0.030	0.721	0.471
Insanity	0.277	0.043	6.395	**0.000**
Isolation	−0.194	0.042	−4.576	**0.000**
Death	0.254	0.031	8.250	**0.000**
Paranoia	−0.033	0.036	−0.897	0.370
Spiritual significance of the experience				
Fear	−0.327	0.048	−6.808	**0.000**
Grief	0.190	0.044	4.290	**0.000**
Physical distress	0.106	0.029	3.629	**0.000**
Insanity	0.149	0.043	3.494	**0.000**
Isolation	−0.251	0.042	−5.973	**0.000**
Death	0.319	0.031	10.386	**0.000**
Paranoia	−0.047	0.036	−1.292	0.196
Change in wellbeing attributed to the experience[a]				
Fear	0.282	0.049	5.728	**0.000**
Grief	−0.043	0.045	−0.947	0.343
Physical distress	−0.108	0.030	−3.599	**0.000**
Insanity	−0.073	0.044	−1.674	0.094
Isolation	0.122	0.043	2.848	**0.004**
Death	−0.219	0.031	−7.042	**0.000**
Paranoia	0.065	0.037	1.752	0.080

Note: Bold typeface indicates parameters that are significant $p<0.05$.
PE: parameter estimate; SE: standard error.
[a]Positive scores indicate increased well-being.

Table 7. Factor correlations, factor reliabilities, and factor scores for Study 2.

	Fear	Grief	Physical distress	Insanity	Isolation	Death	Paranoia
				Factor correlations and reliabilities			
Fear	*0.84*						
Grief	0.447 (0.029)	*0.86*					
Physical distress	0.491 (0.030)	0.303 (0.034)	*0.81*				
Insanity	0.654 (0.023)	0.416 (0.031)	0.403 (0.034)	*0.76*			
Isolation	0.298 (0.032)	0.634 (0.022)	0.211 (0.035)	0.398 (0.031)	*0.77*		
Death	0.434 (0.029)	0.283 (0.032)	0.346 (0.033)	0.463 (0.030)	0.213 (0.032)	*0.85*	
Paranoia	0.261 (0.038)	0.246 (0.038)	0.240 (0.040)	0.207 (0.040)	0.321 (0.036)	0.091 (0.038)	*0.7*
				Standardized factor scores			
Study 1[a]	0	0	0	0	0	0	0
Study 2	0.616 (0.057)[b]	0.052 (0.043)	0.333 (0.043)[b]	−1.408 (0.052)[b]	−0.070 (0.065)	−0.010 (.061)	−0.180 (0.050)[b]

Factor reliabilities (calculated as Cronbach's alpha) are presented in italics on the diagonal.
[a]Comparison group; [b]$p<0.001$.

and altered items is presented in Supplementary Material, Appendix 2. The stand-alone CEQ is provided for reference and for research use in Supplementary Material, Appendix 1.

Analysis

A series of multiple-group confirmatory factor analyses were estimated to test for factorial invariance of the 26 items of the CEQ between the sample in Study 1 and the sample in Study 2. Factorial invariance procedures and assessment of model fit follow those procedures used in Study 1. Factor means were compared between Study 1 and Study 2. Finally, a structural equation model was fit in the Study 2 sample, regressing ratings of the overall impact of the challenging experience (difficulty, meaningfulness, spiritual significance, and change in well-being attributed to the challenging experience) on CEQ latent variables.

Results

Study 2 differed from Study 1 on all demographic variables (Table 1, right-hand portion). Participants in Study 2 were older, more highly educated, and more racially and ethnically diverse. Study 2 contained a greater percentage of females than Study 1. Participants in Study 2 also reported a greater number of experiences with classic hallucinogens in general and more specifically psilocybin mushrooms than participants in Study 1. Finally, participants in Study 2 rated their experiences as overall slightly less difficult, and slightly more individuals on average in Study 2 indicated that they would repeat their reported experience, including the challenging portion, if given the opportunity.

Measures of sampling adequacy, including the KMO test (KMO=0.89) and Bartlett's test of sphericity (χ^2=538.01, df=26, $p<0.0001$), indicated suitability of the data for factor analysis. Model fit indices for factorial invariance models are presented in the lower portion of Table 5. Change in fit indices for each level of factorial invariance between Study 1 and Study 2 did not exceed critical values for noninvariance. This supports strict factorial invariance of the 26-item CEQ between Study 1 and Study 2. Factor correlations and reliabilities for the stand-alone CEQ are presented in the top portion of Table 7. Factor score differences between Study 1 and Study 2 are presented in the bottom portion of Table 7. Factor scores were significantly higher in Study 2 on fear and physical distress factors, and significantly lower on insanity and paranoia factors than in Study 1.

Model fit indices for the structural equation model regressing overall impact ratings on CEQ factor scores indicated good fit (CFI=0.912, RMSEA=0.064 (90% CI: 0.061–0.067), SRMR=0.048). Regression estimates are presented in Table 8. Scores on the CEQ death factor were negatively associated with ratings of change in wellbeing attributed to the experience, and positively associated with all other ratings of overall impact. Scores on fear were positively associated with ratings of the difficulty of the experience and change in wellbeing, and negatively associated with the spiritual significance of the experience. Scores on insanity were positively associated with rated difficulty of the experience and meaningfulness of the experience. Grief and physical distress were positively and negatively associated with ratings of the difficult of the experience, respectively, and paranoia factor scores were positively associated with change in wellbeing attributed to the experience, but neither factor score were associated with other impact ratings. Factor scores on isolation were not associated with any overall impact ratings.

General discussion

The CEQ was developed from responses to an Internet survey of challenging experiences, and the stand-alone CEQ was validated in responses to a separate Internet survey of challenging experiences. The seven factors of the CEQ display a simple factor structure, have high face validity, and represent a wide sampling of challenging aspects of psychedelic experience (referred to as "challenging experiences" for brevity) that aligns well with previous reports of presenting symptoms of acute adverse reactions to hallucinogenic substances (Cohen, 1960; Strassman, 1984; Ungerleider et al.,1966, 1968). This includes affective (fear and grief), physiological (physical distress), and cognitive/affective (feelings of isolation, paranoia, feelings of insanity or loss of sanity, and the subjective experience of death) categories of subjective experience.

Internal and external validity of the CEQ

The factor structure of the CEQ was identified and confirmed in Study 1 within separate sub-samples of the data, and was shown to be invariant across levels of gender and struggle with psychiatric illnesses before the reported challenging experience. The factor

Table 8. Regression of ratings of overall impact of the experience on Challenging Experience Questionnaire (CEQ) factor scores for Study 2.

DV	PE	SE	*z*	*p*
Difficulty of the experience				
Fear	0.164	0.047	3.456	**0.001**
Grief	0.177	0.043	4.081	**0.001**
Physical distress	–0.099	0.039	–2.543	**0.011**
Insanity	0.165	0.049	3.393	**0.001**
Isolation	0.005	0.042	–0.126	0.900
Death	0.214	0.037	5.834	**0.000**
Paranoia	–0.038	0.037	–1.030	0.303
Meaningfulness of the experience				
Fear	–0.080	0.052	–1.549	0.121
Grief	0.013	0.047	0.270	0.787
Physical distress	–0.024	0.043	–0.572	0.567
Insanity	0.126	0.053	2.379	**0.017**
Isolation	–0.010	0.046	–0.225	0.822
Death	0.256	0.040	6.360	**0.000**
Paranoia	0.009	0.041	0.231	0.814
Spiritual significance of the experience				
Fear	–0.116	0.052	–2.229	**0.026**
Grief	–0.048	0.047	–1.004	0.315
Physical distress	0.025	0.043	0.586	0.558
Insanity	0.045	0.053	0.855	0.393
Isolation	–0.018	0.046	–0.396	0.692
Death	0.303	0.040	7.495	**0.000**
Paranoia	–0.017	0.041	–0.412	0.680
Change in wellbeing attributed to the experience				
Fear	0.116	0.053	2.200	**0.028**
Grief	–0.052	0.048	–1.073	0.283
Physical distress	–0.053	0.043	–1.211	0.226
Insanity	–0.061	0.054	1.130	0.258
Isolation	0.027	0.047	–0.570	0.569
Death	–0.185	0.041	–4.530	**0.000**
Paranoia	0.105	0.041	2.528	**0.011**

Note: Bold typeface indicates parameters that are significant $p<0.05$.
PE: parameter estimate; SE: standard error.

structure of the stand-alone CEQ was validated in Study 2. While demographic variables did not differ in extreme ways between Study 1 and Study 2, they all differed significantly. Factorial invariance between Study 1 and Study 2 in spite of demographic differences provides evidence for internal validity of the CEQ and makes a case for the resilience of the CEQ factor structure.

Participants provided ratings of the degree of difficulty of the challenging experience that they reported. Scores on the fear, grief, insanity, and death factors of the CEQ were consistently positively associated with these difficulty ratings in both Study 1 and Study 2, providing evidence for external validity of the CEQ. The experiences represented by the items in these four factors may constitute the core of what may be considered a challenging experience, while the remaining three factors may represent less reliable aspects of challenging experience.

Differences in factor scores between samples in four scales (fear, physical distress, insanity, and paranoia) were observed (Table 7), and should be explored further in future studies. Greater fear and physical distress scores, and lower paranoia and insanity scores, were observed in Study 2 when compared to Study 1. These differences may be attributable to differences between Study 1 and Study 2 samples in demographics, but it is equally if not more compelling to consider that additional factors that were not controlled in these samples, such as psilocybin dose and details of set and setting, may be associated with differences in CEQ scores.

An accepted clinical definition of "paranoia" is "unfounded fears that others intend harm to the individual" (Freeman et al., 2015). While one item of the paranoia scale of the CEQ is consistent with this definition ("feeling that people were plotting against you"), the other ("experience of antagonism toward people around you"), while likely related, is not closely consistent with this definition. Thus, the CEQ paranoia scale may be viewed as a crude measure of the clinical construct of "paranoia", and the external validity of the paranoia scale may be somewhat restricted by this limitation.

Challenging experiences and the overall impact of experiences with psilocybin

Previous studies of the subjective effects of psilocybin have used ratings of the meaningfulness and spiritual significance of the experience to provide a very general characterization of the effects of psilocybin, in relation to other experiences that a volunteer has had (Griffiths et al., 2006, 2008, 2011). A consistent finding has been that moderate to high doses of psilocybin (20+ mg/70 kg) occasion experiences that are frequently rated in the top five most meaningful and spiritually significant experiences of a participant's life. Participants have also attributed positive change in well-being to their psilocybin experience, and this change in well-being has been shown to correlate positively with mystical experience (Garcia-Romeu et al., 2015) and ego-dissolution (Nour et al., 2016). Preliminary research further suggests positive psilocybin-occasioned behavior change in the context of addiction treatment (Bogenschutz et al., 2015; Garcia-Romeu et al., 2015; Johnson et al., 2014), providing a general characterization of the longer-term impact of these experiences. The associations between CEQ factor scores and the overall impact ratings of meaningfulness, spiritual significance, and change in well-being were assessed in the current samples.

Scores on the insanity and death factors of the CEQ were positively associated with ratings of the meaningfulness of the reported experience in both Study 1 and Study 2. Scores on the fear factor were negatively associated with spiritual significance, while scores on the death factor were positively associated with spiritual significance of reported experiences in both studies. To the extent that individuals might construe or relate the loss of self-referential processing that is often reported during mystical experiences as feeling as though they are losing a sense of sanity or experiencing their own death, an encounter with this facet of challenging experience may be expected to covary with both the meaningfulness and the spiritual significance of an experience. The subjective experience of one's own death and loss of control of the mind might somehow allow for the type of unity experience that leads to spiritual and meaningful experiences. However, scores on the death factor of the CEQ were negatively associated with change in well-being attributed to challenging experiences. While the experience of ego dissolution (Nour et al., 2016) and mystical experience (Barrett et al., 2015) are positively associated with well-being, it is not completely clear that the items of the 'death' subscale of the CEQ ("Profound experience of your own death" and "Feel as if dead or dying") are collinear with either mystical experience or ego dissolution. The positive association between wellbeing and both mystical experience and ego-dissolution, contrasted against the negative association between wellbeing and the 'death' scale of the CEQ in both Study 1 and Study 2, suggests that there may be something unique about the subjective experience described as 'death' or 'dying' during a challenging experience that may detract from wellbeing, and this may have implications for therapeutic efficacy of psychedelics in clinical trials. Thus, future work may benefit from further elucidating the relationship between ego dissolution or mystical experience and the 'death' factor of the CEQ. Scores on the fear factor of the CEQ were associated with an increase in well-being attributed to challenging experiences and negatively predicted meaningfulness and spiritual significance of challenging experiences. Fear may generally detract from a spiritual experience, but the crucible of panic during a challenging experience might still lead to positive outcomes.

Some discrepancies between Study 1 and Study 2 are noted in the significant predictors of meaningfulness, spiritual significance, and change in well-being attributed to experiences. Scores on the paranoia factor were negatively associated with the rated difficulty of the experience in Study 1 but not Study 2, while scores on the physical distress factor were negatively associated with the rated difficulty of the experience in Study 2 but not Study 1, and scores on the isolation factor were not associated with rated difficulty of the experience in either study. While fear and isolation were negatively associated and grief was positively associated with the rated meaningfulness of the experience in Study 1, they were not significantly associated with this rating in Study 2. Grief, physical distress, and insanity were positively associated, and isolation was negatively associated, with spiritual significance of an experience in Study 1, but not Study 2. Further, physical distress was negatively associated and isolation was positively associated with change in well-being in Study 1 but not Study 2, while paranoia was positively associated with change in wellbeing in Study 2 but not Study 1.

It is possible that demographic differences between the samples in each study or the difference between studies in the relative rated difficulty of challenging experiences may have contributed to observed discrepancies in the associations between CEQ factor scores and ratings of the overall impact of the experiences (difficulty, meaningfulness, spiritual significance, and change in well-being attributed to the experience). It may also be that aspects of experience that were not controlled in these studies, such as particular features of set and setting, psilocybin dose ingested (which cannot be precisely known, especially from retrospective reports of ingested fungal matter), or traits of individuals such as personality or attachment type, interact with the relationship between factor scores and ratings of the overall impact of the experiences. Another limitation of this study is that individuals who have a positive attitude towards psychedelics were probably more likely to complete the survey than those with a negative attitude towards psychedelics, given that the main sources of recruitment were websites that are typically frequented by those with a positive attitude towards psychedelics. It may be that the relationships between CEQ scores and measures of the overall impact of the experiences could be different for those with a different regard towards psychedelics. While these questions deserve prospective scrutiny in controlled environments, it is encouraging that consistencies were found between studies (e.g. the relationship between fear, grief, insanity, and death factor scores and rated difficulty of the experience).

Potential risk factors for challenging experiences

Factor scores in Study 1 did not differ by gender (except for slightly greater scores on the fear factor for women compared to men). In contrast to gender, factor scores were significantly greater in those who had previously struggled with a psychiatric disorder, compared to those who had not, for all factors except for paranoia. While this may seem to suggest that individuals with a history of psychiatric disorders may be more

prone to a greater degree of challenge with psilocybin, it is not clear that such a generalization is warranted. The current survey does not inform or address the frequency or likelihood of a challenging experience, but rather the potential profile of subjective experience when a challenging experience is encountered. Also, the designation of having previously struggled with a psychiatric disorder is a coarse designation at best. This categorization does not differentiate well among many important and distinct disorders, including mood disorders or psychosis. Those who suffer from disorders on the psychotic spectrum, which can be accompanied by delusions, ideas of reference, and paranoia, may be at increased risk of experiencing paranoia while experiencing the effects of a classic hallucinogen. Also, reliance on self-identified struggle with a psychiatric disorder is imperfect for diagnostic purposes.

While recommendations for safe conduct of hallucinogen research indicate a contraindication for individuals with personal or family history of psychosis (Johnson et al., 2008), non-psychotic depression may not be contraindicated. Recent empirical work has shown the potential value of psilocybin as a treatment for depression (Carhart-Harris et al., 2016a; Grob et al., 2011) and anxiety secondary to a life-threatening cancer diagnosis (Griffiths et al., 2016; Ross et al., 2016), and recent large-sample surveys have demonstrated a lower risk for psychological distress and suicidality in those who have endorsed having consumed a classic hallucinogen in the past (Hendricks et al., 2015; Johansen and Krebs, 2015). Future studies should explore a more fine-grained understanding of the relationship of previous history with various psychiatric disorders to challenging experiences with psilocybin or other classic hallucinogens and clinical outcomes of psychedelic therapies.

Towards a model of challenging experiences

With the CEQ, one may begin to consider an empirical model for challenging experiences. A full model of challenging experiences should include the acute effects of challenging experiences, the conditions that may predict a challenging experience, and the persisting effects of that experience. The CEQ provides a structure to consider the acute effects of challenging experiences that may be used in prospective studies. Additional categories of challenging experience have been suggested in the literature, such as the experience of karmic or "astral" experiences, experience of "kundalini" energy, or the experience of such phenomena as age regression (McCabe, 1977). Some of these experiences may be culturally bound or framed explanations of intense emotional or physiological responses to drug conditions. Moreover, many such experiences are non-specfiic to challenging experiences, as they may arise during both challenging and non-challenging experiences. In this case, they may not be uniquely informative of challenging experiences.

There are many questions that can be asked regarding challenging experience, including the nature of challenging experiences, the predictors of emergence of challenging experiences, and the consequences of having had a challenging experience. The current findings deal primarily with challenging aspects of experiences with psilocybin, as reported by the study volunteers regarding their primary subjective experience. Future studies may benefit from this understanding, and will be able to use the CEQ to investigate predictors and consequences of challenging experiences.

Conclusion

The CEQ as validated in this article may serve as a valuable tool for characterizing psychologically difficult aspects of experiences occasioned by psilocybin and, very likely, by other classic hallucinogens. Better understanding of challenging experiences with classic hallucinogens may increase the precision of our understanding of both the psychological nature of and neural mechanisms underlying the effects of these drugs. Understanding challenging experiences with classic hallucinogens may also facilitate the optimization of therapeutic application of drugs within this class. Thus, the CEQ developed in this article makes a significant contribution to methods of assessment of acute psychologically adverse reactions to psilocybin and, with further validation, will likely make a significant contribution to methods of assessment of such experiences with other classic hallucinogens.

Acknowledgements

The authors wish to thank the many anonymous survey respondents who generously gave their time and effort to complete this survey. They also thank Earth and Fire from www.erowid.org for their generous assistance in facilitating survey recruitment

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: RRG is a member of the Board of Directors of the Heffter Research Institute.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by NIH grants RO1DA03889 and 5T32 DA007209 and grants from the Heffter Research Institute and the Council on Spiritual Practices.

References

Abramson HA (1960) Lysergic acid diethylamide (LSD-25). XXXI. comparison by questionnaire of psychotomimetic activity of congeners on normal subjects and drug addicts. *J Ment Sci* 106: 1120–1123.

Barrett FS, Johnson MW and Griffiths RR (2015) Validation of the revised mystical experience questionnaire in experimental sessions with psilocybin. *J Psychopharmacol* 29: 1182–1190.

Belleville RE (1956) MMPI score changes induced by lysergic acid diethylamide (LSD-25). *J Clin Psychol* 12: 279–282.

Bentler PM (1990) Comparative fit indexes in structural models. *Psychol Bull* 107: 238–246.

Bogenschutz MP, Forcehimes AA, Pommy JA, et al. (2015) Psilocybin-assisted treatment for alcohol dependence: A proof-of-concept study. *J Psychopharmacol* 29: 289–299.

Browne MW and Cudeck R (1993) Alternative ways of testing model fit. In: Bollen KA and Long JS (eds) *Testing Structural Equation Models*. Newbury Park, CA: Sage, pp.136–162.

Carbonaro TM, Bradstreet MP, Barrett FS, et al. (2016) Survey study of challenging experiences after ingesting psilocybin mushrooms: Acute and enduring positive and negative consequences. *J Psychopharmacol*. Epub ahead of print 30 August 2016. DOI: 0269881116662634.

Carhart-Harris RL, Bolstridge M, Rucker J, et al. (2016a) Psilocybin with psychological support for treatment-resistant depression: An open-label feasibility study. *Lancet Psychiatry* 3: 619–627.

Carhart-Harris RL, Kaelen M, Bolstridge M, et al. (2016b) The paradoxical psychological effects of lysergic acid diethylamide (LSD). *Psychol Med* 46: 1379–1390.

Carhart-Harris RL, Leech R, Hellyer PJ, et al. (2014) The entropic brain: A theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front Hum Neurosci* 8: 1–22.

Carhart-Harris RL, Leech R, Williams TM, et al. (2012) Implications for psychedelic-assisted psychotherapy: Functional magnetic resonance imaging study with psilocybin. *Br J Psychiatry* 200: 238–244.

Chen FF (2007) Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Modeling* 14: 464–503.

Cohen S (1960) Lysergic acid diethylamide: Side effects and complications. *J Nerv Ment Dis* 130: 30–40.

Dittrich A (1975) Zusammenstellung eines fragebogens (APZ) zur erfassung abnormer psychischer zustände [Construction of a questionnaire (APZ) for assessing abnormal mental states]. *Z Klin Psychol Psychiatr Psychother* 23: 12–20.

Dittrich A (1998) The standardized psychometric assessment of altered states of consciousness (ASCs) in humans. *Pharmacopsychiatry* 31: 80–84.

Dittrich A, Lamparter D and Maurer M (2010) *5D-ASC: Questionnaire for the Assessment Of Altered States of Consciousness. A Short Introduction*. Zürich, Switzerland: PSIN PLUS.

Freeman D, Dunn G, MurrayRM, et al. (2015) How cannabis causes paranoia: using the intravenous administration of 9-tetrahydrocannabinol (THC) to identify key cognitive mechanisms leading to paranoia. *SchizophrBull* 41: 391–399.

Frecska E and Luna LE (2006) The adverse effects of hallucinogens from intramural perspective. *Neuropsychopharmacol Hung* 8: 189–200.

Garcia-Romeu A, Griffiths RR and Johnson MW (2015) Psilocybin-occasioned mystical experiences in the treatment of tobacco addiction. *Curr Drug Abuse Rev* 7: 157–164.

Gouzoulis-Mayfrank E, Heekeren K, Neukirch A, et al. (2005) Psychological effects of (S)-ketamine and N,N-dimethyltryptamine (DMT): A double-blind, cross-over study in healthy volunteers. *Pharmacopsychiatry* 38: 301–311.

Griffiths RR, Johnson MW, Carducci MA, et al. (2016) Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial. *J Psychopharmacol* 30: XXX–XXX.

Griffiths RR, Johnson MW, Richards WA, et al. (2011) Psilocybin occasioned mystical-type experiences: Immediate and persisting dose-related effects. *Psychopharmacology* 218: 649–665.

Griffiths RR, Richards W, Johnson MW, et al. (2008) Mystical-type experiences occasioned by psilocybin mediate the attribution of personal meaning and spiritual significance 14 months later. *J Psychopharmacol* 22: 621–632.

Griffiths RR, Richards WA, McCann U, et al. (2006) Psilocybin can occasion mystical-type experiences having substantial and sustained personal meaning and spiritual significance. *Psychopharmacology* 187: 268–283; discussion 284–292.

Grob CS, Danforth AL, Chopra GS, et al. (2011) Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer. *Arch Gen Psychiatry* 68: 71–78.

Haertzen CA (1966) Development of scales based on patterns of drug effects, using the addiction research center inventory (ARCI). *Psychol Rep* 18: 163–194.

Halberstadt AL (2015) Recent advances in the neuropsychopharmacology of serotonergic hallucinogens. *Behav Brain Res* 277: 99–120.

Hendricks PS, Thorne CB, Clark CB, et al. (2015) Classic psychedelic use is associated with reduced psychological distress and suicidality in the united states adult population. *J Psychopharmacol* 29: 280–288.

Hu L and Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling* 6: 1–55.

Hübner C (2007) EnTrance: Entrance to wider worlds, or mystifcation of mere relaxation? *Music Therapy Today* 8: 257–293.

Johansen PO and Krebs TS (2015) Psychedelics not linked to mental health problems or suicidal behavior: A population study. *J Psychopharmacol* 29: 270–279.

Johnson MW, Garcia-Romeu A, Cosimano MP, et al. (2014) Pilot study of the 5-HT2AR agonist psilocybin in the treatment of tobacco addiction. *J Psychopharmacol* 28: 983–992.

Johnson MW, Richards W and Griffiths RR (2008) Human hallucinogen research: Guidelines for safety. *J Psychopharmacol* 22: 603–620.

Kaelen M, Barrett FS, Roseman L, et al. (2015) LSD enhances the emotional response to music. *Psychopharmacology* 232: 3607–3614.

Kaelen M, Roseman L, Kahan J, et al. (2016) LSD modulates music-induced imagery via changes in parahippocampal connectivity. *Eur Neuropsychopharmacol* 26: 1099–1109.

Kjellgren A, Sundequist U, Sundholm U, et al. (2004) Altered consciousness in flotation-REST and chamber-REST: Experience of experimental pain and subjective stress. *Soc Behav Pers* 32: 103–115.

Kline RB (2005) *Principles and Practice of Structural Equation Modeling*. New York: Guilford.

Kometer M, Pokorny T, Seifritz E, et al. (2015) Psilocybin-induced spiritual experiences and insightfulness are associated with synchronization of neuronal oscillations. *Psychopharmacology* 232: 3663–3676.

Kraehenmann R, Preller KH, Scheidegger M, et al. (2015) Psilocybin-induced decrease in amygdala reactivity correlates with enhanced positive mood in healthy volunteers. *Biol Psychiatry* 78: 572–581.

Lebedev AV, Kaelen M, Lovden M, et al. (2016) LSD-induced entropic brain activity predicts subsequent personality change. *Hum Brain Mapp* 37: 3203–3213.

Linton HB and Langs RJ (1962) Placebo reactions in a study of lysergic acid diethylamide (LSD-25). *Arch Gen Psychiatry* 6: 369–383.

Lomax RG (1983) A guide to multiple-sample structural equation modeling. *Behav Res Methods Instrum Comput* 15: 580–584.

McCabe OL (1977) Psychedelic drug crises: Toxicity and therapeutics. *J Psychedelic Drugs* 9: 107–121.

MacLean KA, Johnson MW and Griffiths RR (2011) Mystical experiences occasioned by the hallucinogen psilocybin lead to increases in the personality domain of openness. *J Psychopharmacol* 25: 1453–1461.

Metzner R, Litwin G and Weil G (1965) The relation of expectation and mood to psilocybin reactions: A questionnaire study. *Psychedelic Review* 5: 3–39.

Nichols DE (2016) Psychedelics. *Pharmacol Rev* 68: 264–355.

Nour MM, Evans L, Nutt D, et al. (2016) Ego-dissolution and psychedelics: Validation of the ego-dissolution inventory (EDI). *Front Hum Neurosci* 10: 269.

Novak SJ (1997) LSD before Leary. Sidney Cohen's critique of 1950s psychedelic drug research. *Isis* 88: 87–110.

Pahnke WN, Kurland AA, Goodman LE, et al. (1969) LSD-assisted psychotherapy with terminal cancer patients. *Curr Psychiatr Ther* 9: 144–152.

Panksepp J (1998) *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York: Oxford University Press.

Pekala RJ (1986) Measurement of phenomenological experience: Phenomenology of consciousness inventory. *Percept Mot Skills* 63: 983–989.

Preller KH, Pokorny T, Hock A, et al. (2016) Effects of serotonin 2A/1A receptor stimulation on social exclusion processing. *Proc Natl Acad Sci U S A* 113: 5119–5124.

R Core Team (2012) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Revelle W (2013) *Psych: Procedures for personality and psychological research*. Available at: http://CRAN.R-project.org/package=psych (accessed 29 August 2015).

Riba J, Rodriguez-Fornells A, Urbano G, et al. (2001) Subjective effects and tolerability of the South American psychoactive beverage ayahuasca in healthy volunteers. *Psychopharmacology* 154: 85–95.

Richards WA, Rhead JC, DiLeo FB, et al. (1977) The peak experience variable in DPT-assisted psychotherapy with cancer patients. *J Psychedelic Drugs* 9: 1–10.

Ross S, Bossis A, Guss J, et al. (2016) Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: A randomized controlled trial. *J Psychopharmacol* 30: XXX–XXX.

Rossell Y (2012) Lavaan: An R package for structural equation modeling. *J Stat Softw* 48: 1–36.

Schmid Y, Enzler F, Gasser P, et al. (2015) Acute effects of lysergic acid diethylamide in healthy subjects. *Biol Psychiatry* 78: 544–553.

Strassman RJ (1984) Adverse reactions to psychedelic drugs. A review of the literature. *J Nerv Ment Dis* 172: 577–595.

Strassman RJ (1995) Hallucinogenic drugs in psychiatric research and treatment. perspectives and prospects. *J Nerv Ment Dis* 183: 127–138.

Strassman RJ, Qualls CR, Uhlenhuth EH, et al. (1994) Dose-response study of N,N-dimethyltryptamine in humans. II. Subjective effects and preliminary results of a new rating scale. *Arch Gen Psychiatry* 51: 98–108.

Studerus E, Gamma A and Vollenweider FX (2010) Psychometric evaluation of the altered states of consciousness rating scale (OAV). *PloS One* 5: e12412.

Studerus E, Gamma A, Kometer M, et al. (2012) Prediction of psilocybin response in healthy volunteers. *PloS One* 7: e30800.

Studerus E, Kometer M, Hasler F, et al. (2011) Acute, subacute and long-term subjective effects of psilocybin in healthy humans: A pooled analysis of experimental studies. *J Psychopharmacol* 25: 1434–1452.

Tagliazucchi E, Roseman L, Kaelen M, et al. (2016) Increased global functional connectivity correlates with LSD-induced ego dissolution. *Curr Biol* 26: 1043–1050.

Taylor RL, Maurer JI and Tinklenberg JR (1970) Management of "bad trips" in an evolving drug scene. *JAMA* 213: 422–425.

Tillé Y and Matei A (2012) *Sampling: Survey sampling*. Available at: http://CRAN.R-project.org/package=sampling (accessed 30 June 2015).

Ungerleider JT, Fisher DD and Fuller M (1966) The dangers of LSD. analysis of seven months' experience in a university hospital's psychiatric service. *JAMA* 197: 389–392.

Ungerleider JT, Fisher DD, Fuller M, et al. (1968) The "bad trip"–the etiology of the adverse LSD reaction. *Am J Psychiatry* 124: 1483–1490.

Vollenweider FX, Csomor PA, Knappe B, et al. (2007) The effects of the preferential 5-HT2A agonist psilocybin on prepulse inhibition of startle in healthy human volunteers depend on interstimulus interval. *Neuropsychopharmacology* 32: 1876–1887.

Vollenweider FX, Vollenweider-Scherpenhuyzen MF, Babler A, et al. (1998) Psilocybin induces schizophrenia-like psychosis in humans via a serotonin-2 agonist action. *Neuroreport* 9: 3897–3902.

Walach H and Kaseberg E (1998) Mind machines: A controlled study on the effects of electromagnetic and optic-acoustic stimulation on general well-being, electrodermal activity, and exceptional psychological experiences. *Behav Med* 24: 107–114.

Widaman KF and Reise SP (1997) Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In: Bryant KJ, Windle M and West G (eds) *The Science of Prevention: Methodological Advances from Alcohol and Substance Abuse Research*. Washington, DC: American Psychological Association, pp.281–324.