Microdosing psychedelics and its effect on creativity: Lessons learned from three double-blind placebo controlled longitudinal trials

Luisa Prochazkova*,
Michiel van Elk*
Josephine Marschall*,
Ben D. Rifkin*,
Neil R. Schon*,
Donatella Fiacchino*,
George Fejer*,
Martin Kuchar*,
Bernhard Hommel*

1 Institute of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK, Leiden, The Netherlands
2 Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 129B, 1018WT, Amsterdam, The Netherlands.
4 Forensic Laboratory of Biologically Active Substances, Department of Chemistry of Natural Compounds, University of Chemistry and Technology Prague, Prague, Czech Republic
5 Department of Experimental Neurobiology, National Institute of Mental Health, Kłecany, Czech Republic.

Acknowledgements: We would like to thank the Microdosing Institute and Psychedelic Society of The Netherlands for providing us with the opportunity to run the experiments at several of their events as well as MagicTruffles.com for donating the truffles to participants. Furthermore, we would like to give a special acknowledgment to Dr. Zsuzsika Sjoerds for her invaluable help and advice through-out the study. Finally, we would like to acknowledge Katerina Hajokova, Pascal Lempe, Maria J. Maraver, Dominique P. Lippelt, Lorenza S. Colzato, Samo Zlinicík, Archisman Dash, Eliska Prochazkova and Kenny Elteren for their help in carrying out the experiment. This research was partly funded by an Advanced Grant of the European Research Council (ERC-2015-AdG-694722) to BH, and by the MICR (grant no. VI20172020056) to MK.

Author contributions: The design of the study was developed by LP, BH & MvE. The study was set up and run together by all authors, who also contributed to data interpretation and manuscript reviews and approved the final version of the manuscript. LP had the original idea to investigate effects of microdosing on creativity and wrote a first draft of the article.

Corresponding author
L.Prochazkova
Leiden University, Faculty of Social and Behavioural Sciences, Cognitive Psychology Unit
Wassenaarseweg 52, 2333 AK Leiden, The Netherlands
e-mail: l.prochazkova@fsw.leidenuniv.nl
Tel: +31640296999
ABSTRACT

Introduction: Microdosing refers to the repetitive administration of tiny doses of psychedelics (LSD, Psilocybin) over an extended period of time. This practice has been linked to alleged cognitive benefits, such as improved mood and creativity, potentiated by targeting serotonergic 5-HT2A receptors and facilitating cognitive flexibility. Nonetheless, in the absence of robust, quantitative and double-blind research on the effect of microdosing, such claims remain anecdotal.

Methods: Here, our main aim was to quantitatively explore the effect of microdosing psychedelic truffles on two creativity tasks assumed to rely on separable processes: the Picture Concept Task assessing convergent thinking and the Alternative Uses Task assessing divergent thinking. We present results from 3 double-blind placebo-controlled longitudinal trials (of which one was pre-registered) conducted in a semi-naturalistic setting. Furthermore, we controlled for expectation and learning biases, and the data were mega-analyzed across trials with a pooled sample of 175 participants in order to maximize statistical power.

Results: In the final analyses we found that active microdosing increased the ratio of original responses (originality/fluency), indicating higher quality of divergent answers in the active microdosing condition. The unadjusted originality score was significantly more pronounced in the active microdosing condition, but only when relative dosage (dose/weight of participants) was considered. These effects were present after controlling for expectation and demographic biases. No effects of active microdosing were found for convergent thinking or any other divergent-thinking score. The results suggest that the effects of truffle microdosing are limited to divergent quality and are more subtle than initially anticipated. Our findings furthermore highlighted the importance of controlling for expectation biases, placebo effects, and prior psychedelic experience in microdosing practice and research.

Keywords: Microdosing, Psychedelics, Double-blind, Placebo-controlled, Field study, Convergent thinking, Divergent thinking, Creativity
1. INTRODUCTION

Humour, metaphors and poetry are all expressions of creativity that rely on connecting remote concepts through an uncommon, unexpected link or association. As high psychedelic doses were shown to stimulate hyper-associative thinking, increase mental imagery, intensify emotions and alter sense of meaning (Baggott, 2015; Carhart-Harris, 2018a; Lu et al., 2015; Nichols, 2016; Zabelina & Robinson, 2010), the question arises whether psychedelics can promote creativity (Girn et al., 2020; Sessa, 2008).

Creativity is not a unitary construct or a single process (Sadler-Smith, 2015, Wallas, 1926). Rather, it is a multi-layered set of processes consisting of dissociable and, to some degree opposing sub-processes, including convergent and divergent thinking (Guilford, 1967). Convergent thinking is the ability to combine remote concepts in order to derive a single logical solution. This ability can be assessed by means of the Remote Associates Test (Mednick, 1968), which requires deriving a single logical connection by connecting three words (e.g., "...Cottage", "...Swiss", and "...Cake"; with the connection being "Cheese"; e.g., Bowden & Jung-Beeman, 2003). In contrast, divergent thinking is the ability to generate many loosely associated ideas triggered by a single open-ended question. This ability can be assessed by means of the Alternative Uses Task (AUT), which requires reporting as many possible creative uses for an ordinary object (e.g., creative uses of a bottle, towel) within a time limit.

Mechanistically, divergent thinking is associated with "cognitive flexibility" and low top-down inhibition allowing for the parallel co-activation of mental representations (Boot et al., 2017; Cools & D'Esposito, 2011; Dreisbach & Goschke, 2004; Hommel, 2015). In contrast, convergent thinking is associated with high top-down control, increasing mutual competition between representations, and steering cognitive search towards a single correct answer. Conversely, the creative output involves distinct and possibly antagonistic sets of cognitive control sub-processes underpinned by specific neural network (Zhang et al., 2020) and underlying psycho-pharmacology (Akbari Chermahini and Hommel, 2010). Psychedelics were hypothesized to possibly induce state of enhanced cognitive flexibility (Carhart-Harris & Nutt (2017) and as such we would anticipate to observe effects on divergent domain rather than observing effects on overall creativity. Accordingly, we will systematically distinguish between convergent and divergent thinking.

Specifically, psilocybin (e.g., active compound in magic truffles) belongs to a group of tryptamines that exert their primary psychedelic effects through agonizing serotonergic 5-HT2A receptors (Kraehenmann et al., 2017; Nichols, 2016; Vollenweider and Geyer, 2001). Pertinently, 5-HT2A signalling has been linked to enhanced cognitive flexibility (Carhart-Harris & Nutt, 2017), improved associative and reversal learning (Aloyo et al., 2001; Boulougouris et al., 2008; Harvey, 1995, 2003; Zhang & Stackman, 2015), and functional neuro-plasticity (Catlow et al., 2013; Hutten et al., 2021), which can be theoretically related to greater divergent creativity.

However, research into the acute effects of psychedelics on creativity has been thus far challenging. Most research assessing effects of psychedelics has been carried out with moderate-large doses (Mason et al. 2019; 2021), which are known to render cognition highly disorganized (Robin L.
State of highly dis-organized cognition makes measurement of high level cognition and its subprocesses difficult as participants struggle to comply with experimental instructions and show impairments in basic memory and executive functions (Pokorny et al., 2020). For instance, Schmidt et al. (2018) showed that higher-order inhibitory performance is impaired under moderate doses (100µg) LSD and the inhibition deficits positively correlated with the level of visual hallucinations. Similarly, a recent study by Mason et al. (2021) showed that acute effects of moderate doses of psilocybin decrease task-dependent creativity (convergent thinking and divergent fluency), while seven days after treatment the number of novel ideas increased. Fortunately, novel opportunities to study cognitive enhancement via psychedelics arose recently. The regular administration of very small psychedelic doses, referred to as 'microdosing', has become a popular practice in western culture (Fadiman and Korb, 2019). Taking a fraction (1/10th - 1/20th) of psychedelic substances was anecdotally associated with various benefits, including improvements in creativity (Szigeti et al., 2021; Ona and Bouso, 2020), while avoiding strong psychedelic effects, which are generally undesirable in routine life. As such, microdosing practice presents a promising approach for cognitive enhancement as well as avenue to studying the interaction between psychedelics and cognition in a more controlled manner.

1.1. Previous studies
Considering the novelty of the microdosing phenomena, the efficacy of microdosing on cognition remains relatively unexplored. The majority of published research on microdosing is based on online prospective surveys. Relevant to the current study, enhanced in creativity and mood were shown to be among the most commonly reported benefits associated with microdosing (Anderson, et al., 2019a; Anderson, et al., 2019b; Fadiman & Korb, 2019; Lea et al., 2020; Ona and Bouso, 2020). Online qualitative research is an effective way to collect data from many participants in a naturalistic setting, but this method is limited in validity due to participants' false memories and (un)intentional distortions of subjective experiences (Schwarz, 1999). Moreover, psychedelics are known to enhance suggestibility and sense of meaning (Carhart-Harris et al., 2015; Hartogsohn, 2018), potentially making participants susceptible to overestimating their creativity.

In our previous research (Prochazkova et al., 2018), we utilized quantitative measures of creativity and intelligence to substantiate the effect of microdosing on cognition beyond self-reports. The results indicated that both convergent and divergent thinking performance improved during the acute effects of a microdose, whereas performance on an intelligence test remained unaffected. However, these findings could only be considered as preliminary because we employed an open-label prospective design, which is subject to expectation bias. Similarly, Anderson and colleagues (2019b) conducted an open-label quasi-experimental online study with participants who were microdosing and compared these to a non-microdosing control group. The authors also utilized quantitative measures of creativity (e.g., Unusual Uses Task) and responses were evaluated on three dimensions by three independent
researchers following instructions by Silvia et al. (2008). In line with our findings, they reported that microdosing participants showed more uncommon, remote and subjectively perceived as clever answers than the control participants with non-microdosing experience. However, open-label and quasi-experimental research does not control for placebo effects and expectation biases, making these results difficult to interpret.

The gold standard in quantitative clinical trials is to employ a double-blind placebo-control design. However, restriction in drug policy renders placebo-controlled research with psilocybin financially and administratively challenging. This is especially true for microdosing, which requires the regular administration of psychedelics for a prolonged period of time. To our knowledge, only one clinical placebo-controlled trial with healthy volunteers examined the effects of microdosing on creativity. Specifically, Bershad et al. (2019) assessed performance on convergent creativity measured by Remote Association Task in a dose-dependent LSD study (0 µg; 6.5 µg; 13 µg; 26 µg) in 20 healthy adults. Results showed no effect of active microdosing on the convergent thinking, yet more attempted trials were reported for the RAT in the active microdosing condition. Bershad et al. (2019) did not assess divergent thinking and did not control for expectation bias, which presents an incomplete picture of the microdosing effects on creativity.

Furthermore, existing clinical trials assessing the effects of microdosing on cognition generally suffer from low sample sizes (Bershad et al., 2019, 2020; Family et al., 2020; Hutten et al., 2020, 2021; Yanakieva et al., 2019). Small sample sizes may present statistical power issues, especially if interacting factors are included in the analyses (expectation effects, previous psychedelic experience). Psychoactive microdosing effects are very subtle, and thus well-powered research is necessary to detect small cognitive effects. Also, a clinical environment does not ensure high ecological validity in terms of setting (e.g. synthetically sourced psilocybin, long testing hours, clinical supervision). A clinical setting may interfere with microdosing effects, especially considering that psychedelics are infamously susceptible to environmental factors (Carhart-Harris & Nutt, 2017).

1.2. Current study

In the current research, we aimed to address these limitations by conducting three double-blind placebo-controlled longitudinal trials, allowing for well-powered mega-analyses of a pooled sample of 175 participants. The study’s design was unique considering psychedelic truffles’ legal status in the Netherlands, permitting us to combine well-controlled laboratory assessment with a naturalistic microdosing setting. Similar to our previous study (Prochazkova et al., 2018), the three experiments were carried out around public microdosing workshops organized by the Microdosing institute (MI) and the Psychedelic Society of the Netherlands (PSN) — which put some constraints on the degree of experimental control. We aimed to extend our previous approach by implementing a placebo-controlled longitudinal design with the help of collaborating parties (MI, PSN, MagicTruffles.com). Microdoses were obscured in opaque capsules, and experimental randomization was carried out by an external
charity organization (PSN) which concealed the blinding manipulation from researchers until the end of the analyses. Psilocybin-containing doses (or placebos) were self-administered by participants in naturalistic settings, and creativity assessments were carried out in a well-controlled laboratory setting. The doses were post-hoc analysed for psychedelic content. Subjective drug effects and expectation effects were systematically assessed throughout all trials.

2. GENERAL METHOD

2.1. General design

Three randomized, double-blind, placebo-controlled longitudinal trials took place at two experimental testing sites at Leiden University (Experiment 1 and Experiment 2) and the University of Amsterdam (Experiment 3). Data was collected from November 2018 to March 2019. Considering the legal status of psychedelic truffles in the Netherlands and the fact that psychedelic doses were self-administered in the field by the participants, the protocols were approved by the local ethics committees of Leiden University and University of Amsterdam. The experimental procedure complies with ethical standards from the Helsinki Declaration of 1975, as revised in 2008.

Experiments 1 (Fig. 1a) and 2 (Fig. 1b) were between-participant placebo-controlled designs during which approximately half of the participants self-administered placebos (the other half administered active doses) in a naturalistic setting. Experiments 1 and 2 lasted approximately three and four weeks, respectively. The last experiment, which was pre-registered on the Open Science Framework used a within-participant crossover design lasting approximately eight weeks (Fig. 1c). Lastly, data was mega-analyzed across the three trials with frequentist and Bayesian approaches, while controlling for expectation and demographic factors. The mega-analysis was carried out to overcome possible power issues, sampling differences, and to control for differences in study designs (e.g., dose size, time spent on trial) across experiments.
**Figure 1.** Overall flowcharts of the experimental procedures. (A) Outline of the Experiment 1, between-subject design (B) Outline of the Experiment 2, between-subject design (C) Outline of Experiment 3, within-subject cross-over design. Meta-analyses was conducted across the three studies at the 4 weeks of dosing (between-subject comparison). The dosing days indicate either placebo or psilocybin active microdose.

**Figure 2. Dosing and experimental schedule.** (A) An example of a possible dosing schedule that participants could receive during a microdosing workshop. Participants were asked to microdose on the days marked in red. Participants were also asked to note any changes they may have made in their dosing throughout the trial. (B) The dependent measures (DV) and control measures (Control) used in Experiment 1 & 2 administered at different time-points of the trials. All measures were administered in the lab apart from Post-acute (Post-Ac) measures which were administered online.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Experiment 1</th>
<th></th>
<th>Experiment 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Dose 2</td>
<td>Dose 6</td>
<td>Post-Ac.</td>
</tr>
<tr>
<td><strong>Outcome measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluency</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Flexibility</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Originality</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Elaboration</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Ratio</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Convergent</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Control measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous psych. exp.</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microdosing strength</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Condition expectation</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>
2.2. General Procedure

All three trials were organized around public microdosing workshop events organized by the MI & PSN - an external organization promoting psychedelic education. Participants who passed PSN’s initial screening of mental health were invited to attend the microdosing workshop and could volunteer in a placebo-controlled study organized by university researchers. Workshop participants interested in taking part in the experimental trials were asked to fill in an additional screening form created by the university researchers and to attend a baseline testing session at the university labs (in Experiment 1 and 2). Only healthy applicants free from contra-indications, including a prior diagnosis or family problems with schizophrenia, psychosis, mania, or borderline disorder, were invited to take part in the microdosing event and related study.

After screening and baseline measures, participants attended the MI & PSN workshops. The public workshops always consisted of the same program, involving a short lecture about psychedelics, dose preparation, randomization to placebo and control conditions, and first dose self-administration. During each workshop, participants put precisely pre-determined amounts and packed psilocybin-containing truffles into opaque capsules. Participants were then randomly assigned by members of the PSN to receive either active microdoses or non-psychoactive placebos. Before the workshops, participants were informed that they may receive either a placebo or an active compound. Condition allocation was concealed by the PSN team to both participants and researchers until the analyses were finalized.

Participants were asked to follow a regular microdosing schedule approximately every 3 days, which is an interval that has been recommended by previous, qualitative research (Fadiman & Korb, 2019). To ensure regular dosing, participants received a dosing schedule with prescribed dosing days (see Fig. 2a), where they could also note any changes they may have made in their dosing throughout the trial. Participants were reminded to self-administer their doses on the dosing days by MI & PSN through reminders sent online. The dosing schedules were collected at the end of the trial by the researchers to provide additional screening information. Participants were requested to take their microdoses up to 1 hour before every testing session, since the effects of psilocybin contained in the truffles are reported to peak approximately 60-90 minutes after ingestion with a plateau phase lasting 2-4 hours (Tylš et al., 2014).

Participants were tested 2 to 3 times in the university labs under the acute effects of the microdose. The test battery was always kept short-lasting (up to ~ 1 hour and 45 minutes) to prevent fatigue and to assess the peak effect of microdosing. For the overview of dependent variables at each testing session see (Fig. 2b and Fig 4a). Other tasks, including creativity tasks, were administered within this time limit, yet creativity tasks were consistently performed as one of the first tasks, to avoid fatigue and potential priming effects after prolonged cognitive testing. Each of the three trials was part of a larger collaborative effort among several independent researchers with different research interests.
Experimental sessions took place in front of a computer screen in a university laboratory and participants were free to take a short break between tasks. During every experimental session in the university lab, participants were asked to reflect on their subjective experience and guess which group they believed to be ('placebo', 'not sure', 'active'). Participants found out their group allocation (placebo/active) after the last online assessment from a PSN representative. However, participants’ blinding stayed concealed to the researchers until the analyses were finalized. Participants were also invited to attend an integration workshop organized by MI & PSN representatives, yet the number of participants attending the integration workshop is not known.

Due to the heterogeneity of experimental questions posed by each researcher, the limited scope of this paper and our primary goal of the current trial to replicate the pilot effect of microdosing on creativity (Prochazkova et al., 2018), results of other tasks will be reported elsewhere.

2.3. The truffle dosing
Psilocybin containing sclerotia (e.g., magic truffles) were donated to participants by certified Magic truffle manufacturer (MagicTruffles.com, Netherlands). All the truffles were the same truffle brand and strain (i.e., Psilocybe Galindoii). The size of microdoses were pre-determined by the PSN team ahead of the workshop. The doses were estimated based on the accepted definition of a microdose, which typically lies around 1/10th to 1/20th of a regular dose (Kim P.C. Kuypers et al., 2019). Considering that the weight of a recreational dose of magic truffles sold in Smart-shops consists of 10 grams of fresh truffles, a mid-size microdose would then equal 1 gram of fresh truffles. All participants consumed the same dose size (regardless of participant weight) within each of the experimental trials. Yet, the weight of the participants was recorded for post-hoc analyses.

The microdosing dose size was up-tiered at every experimental trial to explore possible dose-dependent effects on outcome measures in the final mega-analyses. Specifically, in Experiment 1 participants microdosed with ~ 0.65 grams of fresh truffles equivalent lower range microdose ~ 1/15th, in the Experiment 2 participants microdosed with ~ 1 gram of fresh truffles equivalent mid-range microdose ~ 1/10th and in Experiment 3 participants microdosed with ~ 1.5 grams of fresh truffles equivalent higher-range microdose ~ 1/7th.

Furthermore, the samples of truffles that participants consumed were post-hoc analyzed at the University of Chemistry and Technology Prague to estimate microdosing strength. This measurement followed the same procedure as in our previous study Prochazkova et al., (2018) and its results and procedure are described in detail in the Supplementary materials. Analyses of truffles showed that 1 gram of dehydrated truffles contained on average 1.14mg of active psychedelic alkaloids (psilocybin and psilocin combined) which is similar to previous studies (Pellegrini et al., 2013). Previous clinical trials utilized 10 mg of synthetic psilocybin as a safe clinical dose (Carhart-Harris et al., 2016). Thus, after conversion to the active compounds participants consumed approximately ~ 1/19th of the safe
clinical dose in the first experiment; ~1/13th in the second experiment and ~1/8th in the third experiment. Participants were required to take their microdoses up to 1 hour before every testing session, since the effects of psilocybin contained in the truffles are reported to peak approximately 60-90 minutes after ingestion with a plateau phase lasting 2-4 hours (Tylš et al., 2014). For further details regarding the analysis’s technique and description of the specific truffle alkaloids present in the samples, please see the Supplementary materials (Section S2.1).

2.4. Data pre-processing

The creativity outcomes and subjective measures were collected in Qualtrics software (Qualtrics, Provo, UT). As standard, in all experiments the divergent measures were assessed by two independent raters, and inter-rater reliability was measured. Statistical analysis of interrater reliability were conducted in IBM SPSS Statistics 24. The Reliability scores (Cronbach's Alpha) ranged from excellent (i.e., fluency) to acceptable (i.e., elaboration) confirming satisfactory data quality. Please see the Supplementary materials (Section S2.2), for the reliability analyses of each parameter across all sessions and the three experiments. Bayesian and regression analyses were conducted in JASP (Version 0.14.1). A power calculation was ran with G*power version 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) and it’s input and output is shown in the Supplementary materials (Section S2.3).

3. EXPERIMENT 1

3.1. Method

The goal of the first experiment was to provide quantitative support for the cognitive-enhancing properties of microdosing psychedelics on creativity. Based on our previous preliminary research (Prochazkova et al., 2018), we hypothesized that microdosing psychedelics would either improve divergent thinking and decrease convergent thinking (which would indicate that the drug affects the direct competition between processes involved in convergent and divergent thinking) or improve both convergent and divergent thinking through the optimization of balance in metacontrol policies (Hommel, 2015a). To this end, participants were asked to perform two creativity tasks tapping into convergent and divergent thinking in a single experimental session.

3.1.1. Study design overview

Experiment 1 (Fig. 1a): This first experiment was set-up as a randomized, double-blind, placebo-controlled, longitudinal, design with three assessment time-points carried out in the labs of Leiden University. The design was between-participants with two parallel groups microdosing either with active doses or placebos. Participants were tested once at baseline before the MI & PSN workshop and twice acutely approximately one hour after participants took a microdose. Specifically, participants were tested after the 2nd (e.g., acute 1: short-term effects) and 6th microdose (e.g., acute 2: longer-term additive effects). Furthermore, participants were tested once more sub-acutely - 2 days after the last
microdose to explore possible sub-acute build-up effects of microdosing. For the overview of dependent variables at each testing session see Table 1. Participants were scheduled to attend testing sessions at similar times across the three testing moments to control for possible fluctuations in arousal in a day. During every experimental session, participants were asked to reflect on their subjective experience and guess which condition they were in (placebo vs active). Participants in the placebo condition received identical opaque capsules as in the active condition yet contained 0.65g of non-psychedelic cellulose.

3.1.2. Participants
A total of 103 participants signed up for the psychedelic workshop by means of social media and MI & PSN contact details and 80 participants filled in the initial screening forms. Only 77 healthy participants free from contra-indication, including a prior diagnosis or family problems with schizophrenia, psychosis, mania, or borderline disorder, were invited to attend the microdosing event and baseline measurement at the Leiden University. In total, 72 participants attended the baseline session, 70 attended the acute 1 session (after 2nd dose) and 61 participants attended the acute 2 session (after 6th dose). Two participants had to be excluded from the final data because one took other psycho-active drugs during the trial and the other skipped more than 2 doses. The final AUT sample consisted of 59 healthy participants (30 in placebo and 29 in the experimental group) all of which completed baseline and both follow-up sessions. The final PCT sample consisted only of 57 participants (28 in placebo and 29 in the experimental group) as 2 participants were excluded due to incorrect interpretation of the task. The post-acute AUT measurement which was administered online approximately 2 days after the end of the trial was filled in by 62 participants, yet 3 participants were excluded because they failed to attend previous testing sessions and thus their dosing behavior could not be determined. The first MI & PSN workshop took place in Leiden and the related study was advertised by word-of-mouth at Leiden University. As such the majority of participants were Leiden University students. All participants provided informed consent before starting the study procedure, which was approved by the local ethics committee (Leiden University, Institute for Psychological Research).

3.1.3. Instruments
Alternate Uses Task (AUT). Divergent creativity was assessed by means of the AUT (Guilford 1967) in Experiment 1 and 2. The participants were asked to report as many creative uses of an everyday object. Participants had 5 minutes per item to think of as many possible uses for the object (e.g., towel, pen, brick, mop or bottle). Different items were used for each session, but the order of items presented across time was consistent. By keeping the order consistent we eliminated possible co-founding factors of item difficulty (e.g., availability heuristics) to interact with the result. Conversely, the same pair of AUT items were presented to all participants at each testing sessions and thus the test battery remain equally challenging for both condition across time.
As usual, the AUT reports were rated by two different raters according to five different variables. 

**Fluency** = the total number of responses;  
**Flexibility** = the number of different categories of responses;  
**Elaboration** = the degree to which the person elaborates on their response (each 'elaboration' receives one point, so that the response "using a brick to prevent a door from slamming shut (1), when it is windy (2)" earns two elaboration points);  
**Originality** = the uniqueness of a response in that sample (Responses that have also been mentioned by 1% or less of the other participants receive 2 points for originality, responses that have been mentioned by 1-5% of the participants receive 1 point); and the ratio between originality and fluency (originality/fluency) as an additional index of divergent thinking that was previously suggested as a more parsimonious measure of divergent quality (Hocevar & Michael, 1979.; Runco & Albert, 1985) and was shown to be affected by psychedelics in previous research (Kuypers et al., 2016; Mason et al., 2019).

**Picture Concept Task (PCT).** The (PCT) (Wechsler, 2003, Hurks et al., 2010) is a visual creativity task commonly used as a measure of convergent thinking. In this task, 17 stimuli consisting of a matrix of pictures (either 2x3, 3x3 or 3x4) were presented. For each stimuli participants had to find single common association between one picture from each row by clicking on the items and writing down the common connection (please see Supplementary materials, Section S3.1). In line with previous studies participants had 30 seconds per item to find the solution (Mason et al., 2021). Because the task assumes there is only one correct solution to each trial, the task has been previously compared to the RAT assessing convergent thinking (Hurks et al., 2010). This is because, to complete the task, one should converge on the correct solution while inhibiting inappropriate or less obvious associations or incorrect solutions. The final PCT scores were calculated as the sum of all correct responses.

**Subjective effects.** Subjective drug effects were measured, utilizing several tailored questions. Specifically, participants were asked about how psychoactive they found the microdoses to be (i.e., *How intensively do you perceive microdosing effect at the present moment?*). The intensity of their experience was ranked on a Likert-type scale with ratings ranging from zero to one hundred where zero referred to "no effects", and one hundred referred to "extremely strong psychedelic effects". Secondly, participants reported their guesses regarding their condition allocation at every test session (i.e., *Do you think you are using active psilocybin doses or a placebo?).* Participants were asked to select one from three possible responses: i) active psilocybin, ii) not sure iii) placebo.

**3.1.4. Main analyses**

First, independent samples t-tests and chi-square tests were conducted on demographic data to see whether randomization to placebo and active condition was successful and groups were comparable before the start of the interventions. Next, the subjective effects of microdosing were analyzed. The differences in subjective beliefs concerning condition allocation were analyzed with χ2 tests at every session. Group differences in the perceived psychoactive strength of the microdose were analyzed by Mixed-design rmANOVAs with the session (baseline, acute 1, acute 2) entered as the within-
participant factor and group (placebo vs active microdose) as the between-participant factor. Furthermore, we explored to what degree previous psychedelic experience (naïve vs. experienced) played a role in the ratings of the subjective effects. To this end, differences in subjective ratings of psychedelic strength between psychedelically naïve and experienced participants were analyzed with independent samples t-tests at every session.

To examine whether psilocybin, compared to placebo, altered creativity scores, AUT and PCT data were analyzed by a set of mixed-design rmANOVAs. Assumptions regarding outliers, sample distribution and sphericity were examined, and in case of assumption violation, an appropriate non-parametric correction was applied. Specifically, the acute effect of microdosing on AUT was tested by 2x3 Mixed-design rmANOVAs carried out for each of the five divergent measures (e.g., fluency, flexibility, originality, elaboration, originality ratio) separately. The session (e.g., baseline, acute 1, acute 2) was entered as the within-participant factor and condition (placebo vs active) as the between-participant factor. The post-acute effects on AUT were tested by separate 2x2 rmANOVAs with session (e.g., baseline, post-acute session) entered as within-participant factor and condition (placebo vs active) as the between-participant factor. Since we carried out five univariate tests, we corrected for multiple comparisons with Bonferroni correction, adjusting alpha level to $p < 0.01$. Similarly, the effect of microdosing on PCT was tested with 2x2 Mixed-design rmANOVA with the session (baseline vs acute 2) entered as the within-participant factors and group (placebo vs active microdose) as the between-participant factor. Of particular interest were the two-way interactions between condition and session, as it is these interactions that are diagnostic of the active microdosing effects. Significant effects in the Mixed-design rmANOVAs were followed-up by the appropriate univariate tests and were again Bonferroni corrected according to contrasts of interest.

### 3.1.5. Exploratory analyses of subjective effects

Furthermore, a set of exploratory analyses were carried out to examine what role guessed condition allocation (expect active condition, not sure or placebo) and previous psychedelic experience (naïve vs. experienced) had on the change in creativity measures. To this end, Mixed-design rmANOVAs on AUT and PCT measures were re-run but this time including expectation and previous psychedelic use as a between-participant factor. Furthermore, Pearson's correlations were run to investigate the association between the subjective microdosing strength, and the creativity change from baseline for each session. This analysis aimed to clarify to what degree perceived psychoactive strength may have interacted with the significant results beyond pure drug effects. However, results of the exploratory analyses were reported in the main text only when significant and meaningful results were found, otherwise the analyses were moved to the Supplementary materials to improve readability of the current manuscript.

### 3.2. Results

#### 3.2.1. Sample
Randomized groups did not differ in terms of age, weight, gender, BMI, time of the dosing and previous psychedelic experience at first and second acute session. Please see Table 3. (Experiment 1), which provides an overview of descriptive information regarding the sample statistic. The power calculation for the acute AUT scores (administered 3 times) indicated that a total of 78 participants (with two equal-sized groups of n = 39) should be included within the analyses in order to detect a significant interaction with a moderate effect size (Effect size $f = 0.25$), given alpha = 0.05 and power = 0.80. A power calculation for the PCT scores (administered 2 times) indicated that a total of 82 participants (with two equal-sized groups of n = 41) should be included within the analyses in order to detect a significant interaction with a medium effect size (Cohen’s $f = 0.25$), given alpha = 0.05 and power = 0.80.

Table 3: Demographics and other baseline characteristics, by treatment (randomized set), in all 3 Experiments. Data are presented as mean (SD) unless specified otherwise.
3.2.2. **Subjective effects**

Interestingly, participants in the active psilocybin condition did not differ from the placebo condition regarding their estimates of group allocation at the first ($\chi^2 (3, N = 59) = 0.587, p = 0.746, \text{Cramer’s V} = 0.1$) and the second follow-up session ($\chi^2 (3, N = 59) = 6.21, p = 0.1, \text{Cramer’s V} = 0.324$). The majority of participants reported being uncertain about their group allocation and showed a similar distribution of false positive and negative estimates (see Table 4: Experiment 1). More specifically, 26.6% of participants in the placebo condition believed to be in the active condition at the end of the two-week dosing period. Notably, two participants who were allocated in the control group, experienced very strong placebo effects, and requested smaller dosage from PSN during the trial. This further suggests presence of salient placebo effects as well as the success of the blinding manipulation.

Next, the group differences in strength of the perceived psychoactive effects were analyzed. The rmANOVA indicated a main effect of session ($F (2,110) = 4.499, p = 0.013, \eta^2_p = 0.08$) suggesting that the subjective microdosing strength changed over time while controlling for experimental condition. The follow up t-test indicated that the diminishing microdosing effects were observed between the initial workshop ($M = 23.77, SD = 21.82$) and the last experimental session ($M = 15.68, SD = 16.86$); $t(56) = 2.91, p = 0.005$, for both placebo and active microdosing condition. Nevertheless, the main effect of group ($F (1,55) = 0.97, p = 0.75, \eta^2_p = 0.002$) and the interaction between group and session ($F (2,110) = 0.94, p = 0.39, \eta^2_p = 0.017$) were not-significant, suggesting that participants in active psilocybin condition did not differ in ratings of psychedelic strength from placebo. As can be seen in Table 5. (Experiment 1) participants rated the strength of the psychedelic around 20% (100% referred to extremely strong psychedelic effects) and the psychedelic ratings diminished over time. The result again confirmed that the blinding manipulation was successful and the presence of placebo effects.

### Table 4. Participants' subjective estimation regarding their own group allocation by treatment condition for all three experiments. Total N indicates total sample size for each cell. Note: *=p<0.05; **=p<0.01.

<table>
<thead>
<tr>
<th>Gender % (F/M)</th>
<th>Placebo (N=30)</th>
<th>Active (N=31)</th>
<th>X2</th>
<th>p</th>
<th>Cramer’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>53.3%/46.6%</td>
<td>50%/50%</td>
<td>0.04</td>
<td>0.84</td>
<td>0.04</td>
<td>[Reference]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acute 1</th>
<th>Expect Active</th>
<th>Not sure</th>
<th>Expect Placebo</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5: Ratings of subjective microdosing strength at all three experiments divided by treatment (placebo vs active) measured at different time points. Data are presented in mean (SD). The intensity was measured on the Likert-type scale, zero referred to "no effects" and hundred referred to "extremely strong psychedelic effects". Note: *=p<0.05; **=p<0.01.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Workshop</th>
<th>Dose 2</th>
<th>Dose 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td>21.1 (21.35)</td>
<td>20.5 (21.23)</td>
<td>16.2 (18.53)</td>
</tr>
<tr>
<td>Active</td>
<td>26.74 (22.36)</td>
<td>19.7 (17.55)</td>
<td>15.11 (15.13)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>Workshop</th>
<th>Dose 6</th>
<th>Dose 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td>35.6 (32.07)</td>
<td>21.17 (28.02)</td>
<td>17.1 (19.95)</td>
</tr>
<tr>
<td>Active</td>
<td>38.61 (32.17)</td>
<td>20.77 (24.8)</td>
<td>16.6 (21.6)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>Workshop</th>
<th>Dose 7</th>
<th>Dose 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td>19.38 (18.22)</td>
<td>10.75 (13.38)</td>
<td>4.28 (8.01)</td>
</tr>
<tr>
<td>Active</td>
<td>36.7 (19.28)</td>
<td>25.6 (27.77)</td>
<td>33.23 (30.09) **</td>
</tr>
</tbody>
</table>

### 3.2.3. Previous experience and subjective effects
We further assessed the effect of previous psychedelic experience on subjective drug ratings. Interestingly, independent sample t-tests (with Welsh correction) indicated that participants, who had previous psychedelic experience rated the microdosing drug effect stronger at the acute 2 session, \((M = 19.87, SD = 10.56)\) compared to participants who did not have prior psychedelic experience, \((M = 8.81, SD = 10.56)\), \(t(49.74) = 2.69, p = 0.01, d = 0.713\), regardless of their actual microdosing condition. No effect of previous psychedelic experience was found at baseline and first acute session \((Ts \leq 1.13, Ps \geq 0.28)\). Follow-up analyses indicated that this effect was mainly driven by participants in the placebo condition (see Figure 3). Specifically, participants who had previous psychedelic experience and were in the placebo condition reported significantly higher placebo effects in terms of microdosing strength, \((M = 22.6, SD = 21.5)\) in the acute 2 session than those who had no prior experience with psychedelics, \((M = 6.9, SD = 8.9)\), \(t(21.44) = 2.6, p = 0.016, d = 0.951\). No significant differences for ratings of psychedelic strength were found in the active condition with regards to previous psychedelic experience \((Ts \leq 1.08, Ps \geq 0.29)\).

Figure 3. The graphs show ratings of subjective microdosing strength in Experiment 1. The line plot (a) shows participants allocated to the placebo condition and indicates with participants previous psychedelic experience rated the microdosing strength significantly more salient at the last testing session compared to naive participants. This result suggests that placebo effects are more salient and enduring for people with the previous psychedelic experience. The line plot (b) shows that the microdosing strength was rated as comparable in naive and experienced participants in the active condition. Note: *=p<0.05; **=p<0.01.

### 3.2.4. Divergent thinking

The effect of microdosing on AUT was tested by 2 x 3 Mixed-design rmANOVAs (with Greenhouse-Geisser correction wherever needed) carried out for each of the five divergent measures (see Fig. 4, Experiment 1). The main effects of session were significant for all measures \((Fs \geq 16.3, Ps \leq 0.001, \eta^2_p \geq 0.22)\) when controlling for the effect of condition. The unstandardized results across sessions (see Supplementary materials, Section S3.4) indicated that the main effect reflected mixture of practice effects and variations in AUT items difficulty. The main effects of condition were significant for originality, \((F (1,57) = 4.26, p = 0.044, \eta^2_p = 0.070)\) and originality ratio, \((F (1,57) = 12.87, p < 0.001, \eta^2_p = 0.184)\), and non-significant for fluency, flexibility, and elaboration, \((Fs \leq 0.747, Ps \geq 0.391, \eta^2_p\)
Importantly, the analyses showed far-from-significant two-way interactions for fluency, $(F(2,114) = 0.844, p = 0.433, \eta^2_p = 0.015)$ and flexibility, $(F(2,114) = 0.628, p = 0.535, \eta^2_p = 0.011)$. 

Crucially, significant interactions between session and condition were found for elaboration, $(F(2,114) = 4.924, p = 0.009, \eta^2_p = 0.080)$; originality, $(F(1.64, 93.8) = 5.741, p = 0.007, \eta^2_p = 0.092)$ and originality ratio $(F(1.63, 93) = 8.907, p = 0.001, \eta^2_p = 0.135)$. To clarify where in the experimental procedure the differences occur, contrast analyses were carried. Specifically, we were interested in the between-participant effects at the two follow-up sessions (acute 1 and acute 2) for the five divergent scores. The initial significance level of $p < 0.05$ was corrected to $p < 0.005$, considering we had ten contrasts of interest. As depicted in Fig. 3 (A - E), the active psilocybin group improved in originality as compared to placebo after the last microdose ($p = 0.002, d = 0.837$), although not after the second microdose ($p = 0.162, d = 0.369$). Originality/fluency was higher for the microdosing group in both the first ($p = 0.003, d = 0.804$) and the second experimental session ($p < 0.001, d = 1.11$). Independent sample t-test showed a non-significant difference between groups for elaboration in any of the two follow-up sessions ($p \geq 0.104, d \leq 0.43$). The initial interaction was driven by a significant increase in elaboration from baseline to the first follow-up session in the active condition ($p = 0.009, d = 0.520$), which was not of experimental interest.

3.2.5. Interactions with subjective effects

We explored to what degree the change in creativity scores was influenced by participant’s beliefs regarding their own group allocation and subjectively perceived strength of the acute microdose effect. The exploratory analyses indicated that the expectations regarding group allocation and the ratings of acute microdosing strength did not significantly interact with the change in outcome measures. Similarly, we found a non-significant correlation between the ratings of subjective microdosing strength and the change in outcome measures, $(r_s \leq 0.096, ps \geq 0.478)$. This result suggests that the significant effects could not be explained by subjective effects alone (see Supplementary materials for full analyses, Section S3).

Next, we explored to what degree prior psychedelic experience moderated the drug effects and creativity scores. Interestingly, a significant interaction between condition and previous psychedelic experience was found for the originality ratio, $(F(1, 49) = 5.52, p = 0.023, \eta^2_p = 0.101)$. As shown in Supplementary materials (Section S3.4), the drug effects on creativity were mostly driven by participants with prior psychedelic experience for originality and originality ratio scores. Notably, significant baseline difference in elaboration score were found for psychedelic experienced subjects ($M=4.5, SD=2.6$) as compared to naïve participants ($M = 2.6, SD = 2.8$); $t(51)= 2.38, p=0.021; d=0.66$.

3.2.6. Post-acute effects on divergent thinking

The post-acute effects of microdosing on AUT were non-significant as indicated by $2x2$ rmANOVAs. The main effects of condition $(Fs \leq 0.155, Ps \geq 0.695, \eta^2_p \leq 0.003)$ and interaction between time and
condition were non-significant \((F_s \leq 3.236, P_{\alpha} \geq 0.078, \eta^2_{p} \leq 0.056)\). For visualization of the post-acute effects see Supplementary materials (Section S3.5).

3.2.7. Convergent thinking

Independent sample t-test indicated no differences at baseline performance on PCT between the two conditions \((t (55) = 0.76, p = 0.88)\). The \(rmANOVA\) indicated significant main effect of time \((F (1,55) = 0.35, p = 0.040, \eta^2_{p} = 0.075)\), showing that performance on PCT improved slightly at the second experimental session regardless of the condition. The practice effect could explain this effect as participants were more familiar with the task's rules at the follow-up session. More importantly, results showed no main effect of a condition, \((F (1,55) = 0.426, p = 0.517, \eta^2_{p} = 0.008)\), when controlling for the effect of session and crucially, the interaction effect between time and the condition was also insignificant, \((F (1,55) = 0.35, p = 0.556, \eta^2_{p} = 0.006)\), meaning that performance on PCT in the experimental group did not significantly differ from the placebo (See Figure 3F). This suggests there was no effect of microdosing psilocybin on convergent thinking. The effects remain insignificant even after controlling for expectation effects (see Supplementary materials for full analyses, Section S3).

3.3. Discussion

The goal of the experiment was to provide quantitative support, for the cognitive-enhancing properties of microdosing psychedelics on creativity in a placebo-controlled semi-naturalistic setting. We examined both acute effects (e.g., after second and sixth dose) and sub-acute effects of microdosing on convergent and divergent creativity. Analyses of participants' expectations regarding their group allocation confirmed that the blinding manipulation was successful, and a majority of participants reported uncertainty about their group allocation.

Crucially, we observed that taking microdoses of active psychedelic truffles enhanced performance on two out of five divergent measures compared to placebo. Specifically, the originality and originality ratio scores were significantly higher in the active microdosing condition at the last acute testing session, by which point, participants took six psychedelic microdoses. A significant difference between groups was already observed acutely after the second microdose, but only for the originality ratio score. These effects were present even after controlling for expectations, but there was a moderating impact of previous psychedelic experience. Specifically, the result indicated that microdosing effects were significantly more pronounced for participants with prior psychedelic experience. Similarly, participants with prior psychedelic experience showed significantly higher placebo effects in their ratings of microdosing strength and were more likely to believe to be in the active microdosing condition while actually dosing with placebos. Moreover, the perceived microdosing strength diminished over time regardless of experimental conditions.
These findings could be interpreted as a by-product of state-conditioning account of placebo effects (for review please see Schwarz, Pfister and Büchel, 2016). and essentially, shows the relevance of previous psychedelic experience as an important moderator in microdosing effects. As such, prior psychedelic experience should be accounted for by future research. Finally, the analyses of post-acute effects carried out two days after the last microdose showed non-significant effects. This effect indicates no evidence for sometimes anecdotally reported residual effects of microdosing (Fadiman & Korb, 2019).

The observation of increased originality after active microdosing is consistent with our previous findings (Prochazkova et al., 2018). However, in contrast to our previous research, the current results did not show a significant increase in fluency, flexibility, and the convergent score (PCT measure) in the active microdosing condition as compared to placebo. Some inconsistencies in results can be expected, considering the differences in experimental designs across the two studies. In the former study, an open-label design was used, which is susceptible to placebo effects and learning effects. In the current study, we used a more rigorous placebo-controlled design which is less prone to Type I error. Our results conceptually align with previous research assessing sub-acute effects of moderately-large doses of psychedelic truffles on creativity. Mason et al. (2019) showed that psilocybin containing truffles sub-acute were increased divergent creativity (originality and flexibility). Similarly, Kuypers et al. (2016) investigated the acute effect of large doses of Ayahuasca on both convergent and divergent thinking using the PCT task. The results showed that participants provided more original answers (controlled for fluency) in the PCT and their performance worsened on convergent thinking. With regards to microdosing, the current results fit well with a recent online study carried out by Anderson et al. (2019). The authors assessed creativity using the Unusual Uses Task (UUT) in microdosing participants in the field and compared their performance with non-microdosing controls. The result indicated that the microdosing group performed significantly higher on three creativity indexes: the degree of clever, more uncommon, and more remote responses.

In sum, our results suggest that microdosing does not improve all facets of creativity, yet positive effects of active microdosing were found for originality. Furthermore, the active condition produced a higher percentage of original solutions as compared to the placebo condition, which suggests a trade-off between quality and quantity of divergent responses. These effects were further moderated by prior psychedelic experience.

While these results are promising, several limitations must be noted. First, based on our power calculations the current study was slightly underpowered and as a consequence false-positive (Type I) and false-negative (Type II) errors cannot be ruled out. Secondly, microdoses were self-administered by participants in a naturalistic setting and therefore experimenters did not have a full control over the dosage. While participants followed a dosing schedule and were requested to report any changes to their dosing times, unintentional or willful distortions regarding their dosing could not be excluded. In light
of these limitations, we decided to run an additional set of experiments with the aim to replicate and further build on the current findings.

4. EXPERIMENT 2

4.1. Rationale
The goal of the second experiment was to replicate and build upon findings from Experiment 1. Experiment 2 overlapped closely with the procedure in Experiment 1, yet several alterations were added. Firstly, the psychedelic dose was increased to evaluate possible dose-dependent effects in microdosing. The role of 5-HT2AR signalling in cognition is complex, and, thus, several possibilities of dose-response outcome could be anticipated. First, if we assume a positive linear relationship between psychedelic dose and cognitive flexibility, we would anticipate more pronounced effects on divergent thinking while convergent creativity would suffer. However, several recent papers examining the effects of microdosing on cognition suggested a non-linear relationship between dose size, showing more pronounced effects of lower doses on cognition and BDNF blood levels related to neuro-plasticity (Hutten et al., 2020, 2021; Yanakieva et al., 2019). It is thus possible that serotonergic functions follow an inverted-U-shape function, similar to the dopamine function in cognitive control (Cools & D'Esposito, 2011; Vijayraghavan et al., 2007). If this is the case, it could be hypothesized that lower doses of psychedelic (as administered in Experiment 1) produce more pronounced positive effects on creativity in general, while large microdoses may render cognition too diverted to reach an optimal balance in cognitive control. Furthermore, compared to Experiment 1, in Experiment 2 the duration of the microdosing period was extended to four weeks. A prolonged microdosing period was recommended by previous researchers as optimal to achieve more salient microdosing benefits (Fadiman & Kob, 2017). This effect could hypothetically exist through build-up neuro-chemistry and gradual alteration in baseline brain connectivity (Carhart-Harris et al., 2016; Mason et al., 2021). In that case, we would anticipate the most pronounced effects at the end of the microdosing period. However, a negative relationship between the dosing period and its effects was also plausible considering that 5-HT-2A signaling may down-regulate 5-HT-2 receptors functioning imperative to for psychedelic action (Carhart-Harris & Nutt, 2017).

4.2. Method

4.2.1. Study design overview
In line with Experiment 1, the second study was carried out as a randomized, double-blind, placebo-controlled and between-participant design including three assessment points (baseline, acute 1 and acute 2). The testing consisted of the same creativity tasks (AUT, PCT) and subjective measures as in Experiment 1 (see Table 1). Participants were again tested in the labs of the Faculty of Social Sciences at Leiden University, yet this time the microdosing workshop was organized by MI & PSN in Amsterdam and promoted by a MI & PSN online platforms. As compared to the previous study the
duration of the trial was extended to approximately four weeks and consisted of 10 dosing days (see Fig. 1b). Participants were again tested at baseline and after the 6th dose to mirror the procedure in Experiment 1. Yet, this time the 6th dose referred to the Acute 1 session while in Experiment 1 the Acute 1 session referred to the 2nd dose. Furthermore, as compared to Experiment 1, the second follow-up session (Acute 2) was examined after the 10th dose. This Acute 2 session allowed us to examine the effects of microdosing during a prolonged microdosing period. As compared to Experiment 1, the sub-acute effects were not measured considering the lack of significant results in Experiment 1. Furthermore, as compared to Experiment, the dosage was increased to ~1 gram of fresh psychedelic truffles and the placebo doses (packed again in opaque capsules) contained non-psychadelic truffles mixed with rice to produce identical sound and weight effects to the active microdoses.

4.2.2. Participants

Due to the limited capacity of the workshop, only 100 healthy participants who passed the PSN screening were invited to take part in the microdosing event. From these 100 participants, 96 participants were interested to take part in the associated research. In total 83, healthy participants filled in additional screening and attended the baseline session at Leiden University, 71 participants attended the Acute 1 session (after 6th dose) and 66 participants attended the Acute 2 session (after 10th dose). Furthermore, from the sample of 66 Acute 2 participants, another 5 participants had to be excluded as they missed the Acute 1 session or took other psychoactive drugs during the trial. The final AUT sample consisted of 61 healthy participants who completed both follow-up sessions in a good order (30 in placebo and 31 in the experimental group), with a mean age of 27.75 (SD = 6.32). PCT data were collected in the good order for 63 participants, but 8 participants had to be excluded, due to incorrect interpretations of the task instructions. This is not surprising considering that most participants in the current trial had no prior experience with psychological paradigms. The final PCT sample thus consisted of 55 participants (26 in a placebo and 29 in the experimental group) with an average age of 28.1 (SD = 6.1). Notably, due to a software error 8 participants in the AUT task and 6 participants in the PCT task had missing data at baseline. Nevertheless, considering that within-participant effects were not of primary research interest and baseline data served mainly as a control measure to examine possible pre-existing creativity differences among groups, the data for the missing participants were mean imputed (mean per condition) after the control analyses were carried out to preserve experimental power. All participants provided informed consent before the study onset, approved by the local ethical committee (Leiden University, Institute for Psychological Research).

4.2.3. Instruments

Alternate Uses Task (AUT). The AUT measure mirrored the procedure in Experiment 1, requiring participants to think of creative uses of ordinary items. Participants had 5 minutes per item (2 items presented at each session) and as standard, the AUT was rated by two different raters according to five
different variables (for scoring see Experiment 1). Yet, in contrast to previous experiments, the AUT items were presented in different order. For reliability analyses and the list of AUT items used in each experiment, see Supplementary materials (Table S2, Experiment 2).

**Picture Concept Task (PCT).** Compared to Experiment 1, the PCT measure in Experiment 2 consisted of 15 items presented at baseline and 21 trials presented in the follow-up session in order to increase power of measurement in the experimental phase. Before the analyses, the PCT scores were Z-scored to account for the unequal number of items across the two sessions.

**Subjective effects.** Participants were again asked to reflect on their subjective microdosing effects during every experimental session. Following procedure in Experiment 1, subjective effects in terms of microdosing strength and expectations over own group allocation were measured at every follow-up session.

4.2.4. Main Analyses

The Acute 1 session measures were taken after the 6th microdose (previously acute 2) and the Acute 2 session measures were taken after the 10th dose. PCT was tested only twice at baseline and at Acute 1 session, corresponding to testing points in Experiment 1. The analysis strategy followed the procedure at Experiment 1 and data were analyzed by Mixed-design rmANOvas with session entered as the within-participant factor and group (placebo vs active microdose) as the between-participant factor. The differences in subjective beliefs concerning own group allocation were analyzed with χ2 tests at every follow-up session. The group differences in the perceived psychoactive microdosing strength were analyzed by rmANOvas with the session (baseline, Acute 1, Acute 2) entered as the within-participant factor and group (placebo vs active microdose) as the between-participant factor. However, compared to Experiment 1, the exploratory analyses assessing moderating effects of previous psychedelic experience were not carried out because the majority of participants in Experiment 2, had prior psychedelic experience and thus a comparison could not be carried out.

4.3. Results

4.3.1. Sample

Groups did not differ in terms of age, weight, gender, BMI, time of dosing and previous psychedelic experience (please see Table 2: Experiment 2). Notably, when comparing the demographic information with Experiment 1, the population in the second study was significantly older ($M = 27.76$, $SD = 6.32$) than in the first experiment ($M = 23.96$, $SD = 5.1$); $t (118) = 3.85$, $p < 0.001$) and significantly more cases of participants with prior psychedelic experience. Specifically, 88.9% had psychedelic experience in the second experiment and only 60.4% in the first experiment, $\chi^2(1, N = 107) = 11.52$, $p = 0.001$, Cramer’s $V = 0.328$). The result of the power calculation remained the same as in Experiment 1, which indicated that, to achieve a medium effect size, at least 78 participants were required for the AUT and 82 participants were required for the PCT (see Supplementary materials).
4.3.2. Subjective effects

Interestingly, even after increasing the active dose's size, participants in the active psilocybin condition did not correctly estimate their group allocation. The chi-square test was not significant at the first, \( \chi^2 (2, N = 61) = 0.435, p = 0.805, \text{Cramer’s } V = 0.084 \), nor the second follow-up sessions \( \chi^2 (2, N = 61) = 0.863, p = 0.650, \text{Cramer’s } V = 0.119 \). Thus, even after four weeks of dosing participants performed poorly in estimating their group allocation. The result suggests that the blinding manipulation was successful and placebo effects were present even after the prolonged microdosing period see Table 3. (Experiment 2). The group differences in the perceived psychoactive microdosing strength showed a main effect of session, \( F (2,118) = 19.510, p < 0.001, \eta^{2}p = 0.249 \), indicating that the strength of the dosing subjectively changed across the three measurements. As indicated at Table 4. (Experiment 2) the subjective dose effects again diminished across the three measurements. As indicated at Table 4. (Experiment 2) the subjective dose effects again diminished between the workshop \( M = 37.13, SD = 31.88 \) the first follow-up session \( M = 20.96, SD = 26 \) and the last experimental session \( M = 16.85, SD = 20.6 \) ; \( t \leq 5.107, p < 0.001 \), regardless of experimental condition. However, the rmANOVA analyses showed a non-significant main effect for condition, \( F (1,59) = 0.034, p = 0.854, \eta^{2}p < 0.001 \) and an interaction between group and session, \( F (2,118) = 0.133, p = 0.876, \eta^{2}p = 0.002 \). This result indicates that participants did not break blind regarding their experimental condition, since participants perceived comparable psychedelic strength across the active and placebo condition.

4.3.3. Divergent thinking

The effect of microdosing on the five measures of the AUT was tested by five sets of 2 x 3 rmANOVAs with significance level adjusted to alpha level \( p < 0.01 \). The main effects of session were significant for all measures \( Fs \geq 17.8, Ps \leq 0.001, \eta^{2}p \geq 0.23 \) when controlling for the effect of condition. Similarly to Experiment 1 this effect was not of experimental interest and can be interpreted as a by-product of item difficulty and task repetition. More importantly, the main effects of condition were not significant for the divergent scores when controlling for the effect of time, \( Fs \leq 37.85, Ps \geq 0.001, \eta^{2}p \leq 0.391 \). Crucially, the interactions between session and condition were not significant. Specifically, the two-way interaction for fluency score \( F (1.6, 99) = 0.115, p = 0.857, \eta^{2}p = 0.002 \), flexibility score \( F (2,118) = 0.643, p =0.527, \eta^{2}p = 0.011 \), elaboration score, \( F (2,118) =0.290, p =0.749, \eta^{2}p = 0.005 \), originality score, \( F (2,118) = 1.315, p =0.272, \eta^{2}p = 0.022 \) and originality ratio \( F (2, 118) = 1.674, p = 0.192, \eta^{2}p = 0.028 \) were statistically insignificant. Notably the \( Fs \) were larger than 1 for originality and originality ratio and the direction of the effect was similar to the result in Experiment 1 (see Figure 4, Experiment 2). However, considering the lack of significant findings the post-hoc analyses were not carried out.

4.3.4. Interactions with subjective effects.

The exploratory analyses indicated that the expectations regarding own group allocation and subjective drug strength did not interact with the null findings in Experiment 2 when expectation was added as a between-participant factor. For the analyses and the results see Supplementary materials (Section S4).
4.3.5. Convergent thinking

An independent samples t-test showed that the experimental and control groups were similar in PCT performance at the baseline assessment, \( t(49) = 0.037, p = 0.848 \). The mixed-design rmANOVA indicated no main effect of session \( (F(1,53) < 0.001, p = 0.995, \eta^2 p < 0.001) \), when controlling for the effect of condition. The main effect of condition, \( (F(1,53) = 0.280, p = 0.599, \eta^2 p = 0.005) \) was also insignificant when controlling for the effect of session. Crucially, the interaction effect between time and condition was not significant, \( (F(1,53) = 0.021, p = 0.913, \eta^2 p < 0.001) \), which suggests that active microdosing did not exert effect on convergent thinking as compared to placebo (see Figure 4, Experiment 2). Furthermore, the exploratory analyses indicated that the expectations over own group allocation did not significantly interact with these null effects (see Supplementary materials, Section S4).

Figure 4. Results of Experiment 1 and Experiment 2, show mean Z-scores for the five divergent scores measured by AUT and one convergent score measured by PCT as a function of group (Placebo vs. Active psilocybin) across the three testing session. Vertical capped lines indicate standard error of the mean. Asterisks indicate significant differences between active and placebo sessions with Bonforni adjusted \( \alpha \) level (*\( p < .005 \), **\( p < .001 \)).

4.4. Discussion

The result of Experiment 2 did not replicate the findings from Experiment 1. While some results showed a descriptively similar pattern in the divergent scores in the active microdosing condition (e.g., originality ratio), the difference between the two conditions was not significant. The PCT analyses indicated that microdosing did not improve convergent thinking, which is a result consistent with the
first experiment. We aimed to build on the previous experiment by increasing the microdosing duration and dose size. Interestingly, even with a higher dose and trial duration, participants were not breaking blind in this experiment. Participants reported diminishing drug effects regardless of their experimental condition, and the strength of subjective drug effect was rated comparable across the two conditions (placebo vs active microdose). These results are consistent with Experiment 1 in supporting the prevalence of placebo effects. The lack of a significant effect on divergent thinking indicates that the effects observed in Experiment 1 are not robust. The instability of observed effects could be related to methodological or sampling differences across the two trials. Firstly, it is possible that the dose-effect relationship in microdosing is non-linear and thus more pronounced effects may be found for specific functions in lower ranges of a microdose, which is was an observation corroborating with recent dose-dependent microdosing research (Hutten et al., 2020, 2021; Yanakieva et al., 2019).

Secondly, participants in Experiment 2, were approximately four years older then in Experiment 1. Considering previous research showing reduction in regional serotonin receptors and serotonin transporters in normal aging (Banerjee & Podda, 2016; Meltzer et al., 1998), interaction effects between serotonin induced drug-effects and age are plausible. As shown in Supplementary materials (section S4.1) participants in Experiment 1 showed better baseline performance in all creativity measures which could be either interpreted as by-product of aging (Palmiero et al., 2014, 2017) or could be accounted for by differences in previous psychological training across the two samples. Specifically, in Experiment 2 participants consisted of the general public while the subjects in the first experiment were mostly recruited among the Leiden University students with previous experience with psychological testing. Previous experience with psychological paradigms could be interpreted as advantage in detecting subtle drug-effects as psychologically trained participants are likely to vary less in terms of task comprehension and practice effects as compare to participants carrying out psychological assessment for the first time. Furthermore, varying duration of the two trials (6 doses vs. 10 doses) as well as differences in prevalence of previous psychedelic experience across the two experiments may have differentially affected participants motivation as well as expectation biases (e.g. accumulation and adjustment of expectation with every dose).

Finally, the second experiment was potentially under-powered as according to the power calculation 72 participants were required to detect medium size and only 61 participants were included. To overcome affirmed power issues and shed light on inconsistencies across findings, we carried out Experiment 3, implementing a within-participant cross-over design.

5. EXPERIMENT 3

5.1. Rationale
Experiment 3 aimed to conceptually replicate and build on previous findings while improving statistical power by means of utilizing a within-participant cross-over design. Specifically, previous studies have found that adaptation of the PCT task including a divergent component, was sensitive to sub-acute and acute effects of large doses of psychedelics (Kuypers et al., 2016; Mason et al., 2019). In order to investigate whether microdosing effects on creativity generalize to other divergent paradigms an adaptation of the PCT task was used in Experiment 3. The dose was again increased to 1.5 gram of fresh truffles to account for possible dose-dependent effects. The hypothesis and analysis procedure was pre-registered on the Open Science Framework (https://osf.io/cn8z4/). In the preregistration we hypothesized to find a positive effect of microdosing psychedelics on the divergent scores (fluency, originality, and originality ratio). We did not have specific hypotheses as to whether psilocybin should increase or reduce the convergent scores, due to the inconsistencies in previous findings. Finally, we predicted to find interaction between block order and condition, as we suspected that microdosing effects may diminish over time with build-up tolerance.

5.2. Methods

5.2.1. Study design overview

The initial screening mirrored procedures in the first two experiments and excluded participants with a past medical history or clinical illness. Participants who passed the screening process attended the PSN workshop, where they prepared psychedelic microdoses. As compared to previous Experiment 1 and 2, Experiment 3 implemented a within-participant cross-over design. The testing took place at the labs of University of Amsterdam, and the Amsterdam ethical committee approved the experimental protocol (project no. 2019-SP-10060; see: lab.uva.nl). The creativity tasks were part of a larger test battery administered by researchers with various study interests. The full package of the studies and the hypotheses can be found online (https://osf.io/cn8z4/). The procedure of Experiment 3 was similar to Experiments 1 & 2, except that all participants self-administered 14 microdoses in total in cross-over fashion. The order of active psychedelic doses vs. placebo was switched after the 7th dose. The microdosing period consisted of approximately three weeks of active dosing in one block, two weeks of a wash-out period, and three weeks of placebo self-administered in the other block. The active dosing block and placebo dosing block were contra-balanced across participants (Figure 1:C). At the end of the wash-out period, participants could pick-up the other seven doses in the local Smartshop to resume dosing. This way, we could prevent the direct comparison of the packaged doses, while participants kept them at home. Participants did not have to visit university for baseline measurement considering the cross-over nature of the trial (for list of DVs see Fig 4: Experiment 3). Participants were aware that they would either receive active microdoses or inactive placebos at one of the two experimental blocks. PSN yet again concealed participants' group allocation from the researchers until the end of the analyses.
Fig 4. The four dependent measures (PCT-d) and control measures (Control) used in Experiment 3. Right panel shows DVs used in the final Mega-analysis (PCT-d, AUT).

<table>
<thead>
<tr>
<th>Measures</th>
<th>Experiment 3</th>
<th>Mega-analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Active</td>
<td>Placebo</td>
</tr>
<tr>
<td>Outcome measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluency</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Originality</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Ratio</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Convergent</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Control measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Microdosing strength</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Condition expectation</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

5.2.2. Participants
Due to limited capacity in the MI & PSN workshop, approximately 100 healthy participants who passed PSN screening were invited to join the microdosing workshop. From the 100 participants, only 75 started out with our study and filled in the initial research screening information. Only participants who complied with behavioral guidelines and completed both testing blocks in good order were included in the final analyses (e.g., active microdosing block and placebo block). Twenty-five participants dropped out after the first block and additional 16 participants dropped out during the second block of testing (i.e., approximately eight weeks after the first dose). In total 34 participants completed both required sessions, but 7 participants were further excluded after screening (5 took other psychoactive drugs and 2 participants self-administered microdoses longer than 2.5 hours before testing). This yield a final sample of 27 healthy participants (13 females) between the age of 20 to 48 with mean age 31.3 years ($SD = 10.00$) from which 14 participants started with the active doses.

5.2.3. Instruments
**Picture concept task – divergent (PCT-d).** Instead of AUT task, we used a divergent version of PCT task, here referred to as the PCT-d. Similarly to previous research (Kuypers et al. 2016; Mason et al. 2019) the PCT-d accounts for both convergent and divergent scores simultaneously. Specifically, each participant completed 18 trials, where each stimulus contained between 9 and 12 color pictures shown in a matrix of three or four rows. However, each trial was this time presented for 2 minutes, instead of 30 seconds as in the previous studies. As in Experiment 1 and 2, participants had a time limit of 30 seconds to find a single convergent solution by connecting one picture in all the picture rows presented. After completing the convergent task, participants were asked to report as many alternative solutions (e.g., the divergent responses) within the remaining time. Following previous research and our preregistration the dependent variables consisted of one convergent and three divergent parameters (see
Fig. 4, Experiment 3). The scoring of the convergent and divergent thinking followed the same scoring strategy as described in Experiment 1 and 2, and was administered by two independent raters. The PCT-d was administered twice across two testing blocks (i.e., during active psychedelic block and passive placebo block). For more details about the PCT-d and pre-processing please see the Supplementary material (Section S5).

**Subjective effects.** Participants were again asked to reflect on the psycho-active microdosing effects. Following procedure in Experiment 1, subjective effects in terms of microdosing strength and expectation of own group allocation were measured at every testing block (i.e., block 1 and block 2). However, as compared to previous experiments we did not have any quantifiable data regarding past psychedelic experience, due to data collection error. Yet, informal observations indicated that the majority of participants had prior experience with psychedelics.

### 5.2.4. Main Analyses

The differences in subjective beliefs concerning own group allocation were analyzed with χ² tests at each of the two experimental blocks and microdosing strength across the two drug conditions was analyzed with independent sample t-tests at each of the two testing blocks. Following the statistical pre-registered analyses plan, the primary analyses were carried out with the PCT-d task, including one convergent score and three divergent scores. Specifically, a 2 x 4 rmANOVA was conducted with the type of creativity score (i.e., convergent score, fluency, originality, and originality ratio) and the type of treatment (placebo vs active microdose), both entered as the within-participant factors. Next, the analyses were re-run with 2 x 2 x 4 Mixed-design rmANOVAs, this time including block order as a between-participant factor (e.g., placebo given in block one vs block two) to control for order effect. Post-hoc comparisons were corrected for multiple comparisons, using Bonferroni correction, by adjusting the critical p-value (p < 0.013) according to the four comparisons of interests.

### 5.2.5. Exploratory analyses of subjective effects.

In addition to the pre-registered analyses, exploratory analyses (following the strategy of Experiment 1 and 2) were run to assess the effect of participants’ expectations over own group allocation (e.g., active, placebo, not-sure) on the change in the outcome measures. The data were analyzed by 2 x 4 x 3 rmANOVAs with the type of treatment (placebo vs active microdose) and creativity scores (i.e. convergent score, fluency, originality, and originality ratio) as the within-participant factors and expectation regarding group allocation in block 1 and 2 (e.g., active, placebo, not sure) as the between-participant factor. In addition, correlations were run between the ratings of subjective drug strength at each block and the change in the outcome measures.

### 5.3. Results

#### 5.3.1. Sample analyses

Independent samples t-tests and chi-square tests were conducted to ensure that randomization was successful and that groups were comparable before the start of the intervention. As indicated in the
Table 3. (Experiment 3), groups were comparable across all demographic factors. The population in the third study was significantly older ($M = 31.25, SD = 9.79$; $t (99) = 4.49, p < 0.001$) than in the first ($M = 23.96, SD = 5.1$) and second experiment ($M = 27.75, SD = 6.32$; $t (111) = 2.45, p = 0.016$). All participants had prior psychedelic experience. The reliability scores (Cronbach's Alpha) ranged from excellent (i.e., fluency) to good (i.e., originality) confirming satisfactory data quality. For the reliability analyses as well as more detailed description of the PCT-d task and its pre-processing steps please refer to the Supplementary materials (Experiment 3). A power calculation for the PCT-d scores (administered 2 times) in cross-over within-participant design indicated that a total of 24 participants should be included within the analyses in order to detect differences between conditions with a moderate effect size ($\eta^2_p = 0.06$), given alpha = 0.05 and power = 0.80.

5.3.2. Subjective effects
The differences in subjective beliefs concerning own group allocation were analyzed with $\chi^2$ tests at each of the two experimental blocks. The chi-square analysis indicated that the placebo manipulation was successful, during the first block of testing (after 7$^{th}$ dose), ($X^2 (2, N = 21) = 5.588, p = 0.061, \text{Cramer's } V = 0.516$), however at the following block (14$^{th}$ dose), ($X^2 (2, N = 20) = 11.209, p = 0.004, \text{Cramer's } V = 0.749$) participants were breaking blind. Next, the doses’ psychoactive strength across the two conditions were analyzed. The independent samples t-tests indicated non-significant difference between conditions at the first block (7$^{th}$ dose) ($t(32) = 1.40, p = 0.173, \text{Cohen’s } d = 0.541$), yet in the second block (14$^{th}$ dose) participants in active condition rated the microdosing effects significantly higher ($M = 33.23, SD = 30.09$) than in the placebo condition ($M = 4.28, SD = 8.01$; $t (32) = 3.23, p = 0.002, \text{Cohen’s } d = 1.33$). Furthermore, Table 5. (Experiment 3) indicates that subjective microdosing effects diminished over time regardless of condition between the workshop and first follow-up session ($t (34) = 2.74, p = 0.011, \text{Cohen’s } d = 0.52$), which is a finding consistent with the previous two experiments.

5.3.3. Divergent thinking
The $2 \times 4 \text{rmANOVA}$ indicated no main effect of condition ($F (1,26) = 1.544, p = 0.225, \eta^2_p = 0.056$), when controlling for the type of creativity score, suggesting that condition has not significantly affected overall creativity. More importantly, the interaction between creativity and condition reached significance ($F (3,78) = 2.803, p = 0.045, \eta^2_p = 0.097$), suggesting that at least one of the creativity scores significantly differed between the two conditions. Next, four post-hoc paired-sample t-tests were performed to assess which of the creativity measures show significant differences between the conditions (Bonferroni corrected). Post-hoc t-tests on divergent scores revealed that microdosing elicited a significant higher score on the originality ratio in the active microdosing condition ($M = 0.316, SD = 0.78$) when compared to placebo ($M = 0.1, SD = 0.95$), $t(26) = 2.68, p = 0.012, \text{Cohen’s } d = 0.518$). The effect remains significant even after Bonferroni correction with adjusted alpha ($p = 0.0125$). However, the effect on fluency score ($t(26) = 1.32, p = 0.773, \text{Cohen’s } d = 0.056$) and unweighted originality score ($t(26) = 1.348, p = 0.189, \text{Cohen’s } d = 0.259$), were not significant. For the result, see
In the pre-registration, we expected to find significant effects of microdosing on all three divergent scores.

5.3.4. Convergent Thinking

The post-hoc paired-sample t-test on convergent scores revealed that the difference between the active microdosing condition \((M = 0.05, SD = 0.922)\) and placebo condition \((M = 0.103, SD = 0.99)\), was not significant \(t(26) = 0.29, p = 0.773, \text{Cohen's } d = 0.056\). The lack of effect on convergent thinking corroborates with the previous two experiments. We did not have specific pre-registered predictions regarding convergent thinking.

5.3.5. Order effect

Next, as pre-registered, the analysis was re-run with the session block (active at block one vs active at block two) added as the between-participant factor to control for order effects. The \(2 \times 2 \times 4\) \(rmANOVA\) showed no main effect of a block \((F(1,25) = 0.052, p = 0.821, \eta^2_p = 0.002)\) on creativity when controlling for condition and type of creativity. The two-way interaction between condition and block, \((F(1,25) = 0.194, p = 0.663, \eta^2_p = 0.008)\), and three-way interaction between condition, block, and creativity \((F(3,25) = 0.066, p = 0.978, \eta^2_p = 0.003)\), indicated that block order did not significantly interact with the other factors. However, the key interaction between condition and creativity was insignificant \((F(3,78) = 2.679, p = 0.053, \eta^2_p = 0.097)\). This effect did not corroborate with our pre-registration, where we hypothesized to find a significant interaction between blocks, yet considering the relatively low sample size and the additional between-participant factor, the power of the analyses may have been too low.

5.3.6. Interactions with subjective effects.

The exploratory analyses indicated no main effects and interaction effects with expectations at the first, \((Fs \geq 1.361, p \leq 0.247)\) and second testing block, \((Fs \geq 1.149, p \leq 0.340)\). This suggests that participants' expectations did not significantly interact with the results. However, while controlling for the variance accounted by expectation the key interaction between condition and creativity was insignificant, and the key interaction between condition and creativity was also insignificant \((F < 0.99, p > 0.401)\). The lack of interaction is not surprising considering the fact that condition and expectation significantly overlapped in the second block since the participants broke blind. For the full exploratory analyses please see Supplementary materials (Section S5.3).

5.4. Discussion

The aim of Experiment 3 was to shed light on the inconsistencies between the findings from Experiment 1 and Experiment 2 and, if possible, to substantiate previous findings with sufficient statistical power. The G*Power analyses indicated that 26 participants were sufficient in detecting a medium-size effect for the current design, and thus, the sample size in this trial was considered satisfactory \((N = 27)\). Experiment 3, showed a similar pattern of results as Experiment 1. However, as
compared to Experiment 1, only one out of the two prior effects was replicated. Namely, a significant
difference between conditions was found for the originality ratio score. This effect suggests that
participants produce a lower quantity but a higher quality of creative responses after the active
microdose than after placebo. This is a compelling result, considering the effect's generalizability across
two distinct trials and two different divergent thinking paradigms. Also, the effect did not interact with
participants' expectations regarding their group allocation. However, as compared to Experiment 1,
participants were breaking blind during the second experimental block. Even when explicit ratings of
drug effects did not seem to influence the change in creativity score, implicit biases related to
expectancy- and demand-effects cannot be fully ruled out.

Overall, the current result showed some promise in terms of replicability of microdosing effect
on the originality ratio. While the effect on originality ratio was anticipated (pre-registered), we did not
replicate the effect on unweighted originality observed in Experiment 1, which is a more commonly
used creativity index. Several possibilities regarding the inconsistent finding on unweighted originality
score may be considered. First, the statistical power of experiment 3 was calculated for detecting a
medium effect size. Thus, if small effects for unweighted originality score exist, our sample was not
large enough to detect it (Type 2 error). Secondly, participants in Experiment 3 were, on average, seven
years older than in Experiment 1; they mostly consisted of working professionals, and almost all
participants had prior psychedelic experience. Such findings indicate, again, that the effect of
microdosing on creativity is more subtle than originally anticipated and further replications would be
desirable, to confirm these findings and assess possible interactions between microdosing and
demographic factors.

6. MEGA-ANALYSES

6.1. Rationale

Before drawing final conclusions and interpreting the current finding, we decided to re-examine the
data through exploratory mega-analyses, which allowed us to increase statistical power and reduce
heterogeneity across trials. Specifically, the creativity indices overlapping across three experiments and
collected at similar time-points were pooled together and re-analyzed with standard frequentist and
Bayesian methods to provide additional information about the results.

6.2. Method

6.2.1. Data overview and pre-processing

As shown in Figure 2 (mega-analyses), the data that closely corresponded in experimental time-point
were pooled-together across the three studies. Specifically, we drew data from the second acute
measurement in Experiment 1 (6th dose); the first acute measurement in Experiment 2 (6th dose); and
the first experimental block in Experiment 3 (7th dose). The data extraction from the first block at the
Experiment 3 allowed for between-participant comparison (placebo vs. active microdosing) similarly
to Experiment 1 and Experiment 2 (see Figure 1: mega-analyses). Only the indexes that conceptually overlapped across the three trials were selected for the mega-analyses. Specifically, the convergent score, fluency, originality, and originality ratio were selected from the AUT and PCT(-d). Before data integration, the dependent variables were standardized (Z-scored) per each measure within each experiment to overcome scaling differences across tasks and measures. By increasing statistical power we were able to more effectively control for a variety of demographic factors and expectation biases which could moderate observed effects (e.g., experiment number, age, gender, weight, subjective drug strength and expectation).

6.2.2. Participants
Since participants who failed to attend the last testing session at Experiment 2 and Experiment 3 did not have to be excluded from the mega-analyses, we maximized our sample by including 24 new, not-yet-analyzed participants. Specifically, 61 participants were pooled from Experiment 1, 71 participants from Experiment 2 and 43 participants were pooled from Experiment 3, however after the data-screening 4 participants were excluded (3 took other drugs and one skipped more than 2 doses). The mega-analyses for the divergent scores consisted of pooled data of 171 participants from which 86 participants were in the active psilocybin and 85 in the placebo condition. The sample consisted of 88 females, the mean age of the sample was 27.22 (SD = 7.4), mean weight 70.2 kg (SD = 7.45) and a mean BMI of 23.55 (SD = 10.35). The same sample of participants was used to examine the subjective drug effects. For the convergent scores, data of 138 participants (71 females) with a mean age of 27.3 (SD = 7.6) mean weight 60.68 (SD = 11.26) and mean BMI 23.56 (SD = 11.54) were analyzed.

6.2.3. Main Analyses
First, the subjective microdosing effects were analyzed with $\chi^2$ test and independent sample t-tests. Since psychedelic effects are thought to be highly dose-dependent, we examined to what extent increments in dose across trials interacted with subjective microdosing effects. The subjective strength of microdosing in the active condition was examined across the 3 trials with a one-way ANOVA. The experimental trial was entered as the between-participant factor and subjective dose strength as the dependent variable. In addition, regression analyses were run between the relative dose size (0g, 0.65g, 1g, 1.5g) divided by the weight of the participants and the ratings of subjective microdosing strength to examine dose-dependent effects. The main goal of the mega-analyses was to assess possible differences in the four creativity scores (e.g., a convergent thinking, fluency, originality, and the originality ratio) between the two conditions (e.g., placebo vs active microdose) at a single time point (after the 6th/7th dose) (See Figure 1). The data was analysed using four independent sample t-tests corrected for multiple comparisons (Bonferroni correction $p \leq 0.013$). In addition, a regression analysis was run between the relative dose size (dose/weight) and each of the four dependent measures to examine if the relative dose-size of microdosing had significant effects on creativity. Finally, in addition to standard statistical methods, we calculated Bayesian (posterior) probabilities associated with the occurrence of the null
[p(H0|D)] and alternative [p(H1|D)] hypotheses for each analysis. To estimate Bayesian probabilities, we implemented the procedure previously implemented by Mason (2011).

6.2.4. Exploratory control analyses

Finally, to determine to what degree subjective and demographic factors interact with the creativity scores, we performed a series of two-stage hierarchical regressions. The data was entered in a linear hierarchical regression with each of the four creativity scores entered separately as the modeled variable. In the first step, these nuisance variables were entered in order to control for their variance (i.e. Creativity index= $b_0 + (b_1 \times age) + (b_2 \times gender) + (b_3 \times weight) + (b_4 \times experimental\ trial\ number)$ + $(b_5 \times drug\ strength)$ + $(b_6 \times expectation$). In the second step the condition (active vs. placebo) was included as a regressor (Creativity index= $b_0 + (b_1 \times age) + (b_2 \times gender) + (b_3 \times weight) + (b_4 \times experimental\ trial\ number)$ + $(b_5 \times drug\ strength)$ + $(b_6 \times group\ expectation) + (b_7 \times condition$).

Next, the relative dose-dependent effects of microdosing on creativity were examined by dose size corrected for participants’ weight (dose/weight) used as the main regressor for creativity. As such, the same two-step hierarchical analysis was repeated but this time the relative dose size (dose/participant’s weight) entered in the second step (i.e. e.g. Creativity index= $b_0 + (b_1 \times age) + (b_2 \times gender) + (b_3 \times drug\ strength)$ + $(b_4 \times group\ expectation) + (b_5 \times dose/weight$). A power calculation for hierarchical linear regression indicated that a total of 82 participants should be included within the analyses in order to detect differences between conditions with a moderate effect size ($Partial\ R^2 = .09$), given alpha =0.05 and power = 0.80.

6.3. Results

6.3.1. Subjective effects

First, the differences in subjective drug effects were examined. Independent samples t-tests showed that the active condition ($M = 22.29, SD = 25.64$) and placebo condition ($M = 20.9, SD = 25.32$), $t(169) = 0.355, p = 0.723$, had comparable ratings of microdosing strength at this point of experimental procedures. The chi-square analysis of expectation regarding group allocation was also not significant ($X^2 (3, N = 158) = 4.302, p = 0.231$). The null result suggests that the blinding procedure was successful at this time point ($6^{th}/7^{th}$ dose) even after increasing statistical power. Next, the subjective strength of microdosing was examined across the three experiments. As could be anticipated the one-way ANOVA was significant ($F(2,83) = 1.034, p =0.029$) indicating that subjective drug effects increased with higher dose. The Pearson’s correlation between-participants weight and subjective microdosing effects was negatively correlated ($r(153) = 0.181, p = 0.025$) which suggests mediating effects of weight on the psychoactive microdosing effects (see Supplementary materials for full analyses, Section S6.1).

6.3.2. Divergent thinking

Independent-samples t-tests showed non-significant differences between the active ($M = -0.074, SD = 1.01$) and placebo ($M = 0.056, SD = 0.97$) condition for the fluency score, $t(168) = 0.864, p = 0.389$, with
Bayes factor indicating moderate evidence towards the null effect, \( BF_{01} = 4.722 \). The originality score in active condition (\( M = 0.15, SD = 1.07 \)) and placebo condition (\( M = -0.19, SD = 0.85 \)) was initially significant, \( t(169) = 2.28, p = 0.024, d = 0.349 \), but Bayes factor indicated only anecdotal evidence for the effect, \( BF_{10} = 1.25 \). Finally, the group difference for the originality ratio was significant even after correcting for multiple comparisons, \( t(166) = 0.723, p = 0.002, d = 0.483 \), with participants in the active microdosing condition scoring higher (\( M = 0.22, SD = 1.03 \)) than participants in the placebo condition (\( M = -0.245, SD = 0.89 \)) and Bayes factor showed relatively strong evidence for the effect, \( BF_{10} = 9.902 \).

### 6.3.3. Divergent thinking, relative dose size

The regression analyses replicated previous analyses and showed that relative dose size (dose/participant’s weight) predicted the originality ratio (\( F(1, 159) = 9.24, p = 0.003, R^2 = 0.055 \)) with Bayes factor indicating strong evidence for the alternative hypothesis, \( BF_{10} = 11.24 \). The result for the Fluency score was again insignificant (\( F(1, 159) = 0.07, p = 0.78, R^2 < 0.001 \)), with moderate evidence for the null hypothesis, \( BF_{01} = 5.69 \). Yet, as compared to the previous analyses the result of the linear regression analysis showed that relative dose size (dose/participant’s weight) significantly predicted the unweighted originality score, (\( F(1, 159) = 7.39, p = 0.007, R^2 = 0.04 \)), with Bayes factor indicating moderate evidence for the effect, \( BF_{10} = 4.91 \). This result suggests that a higher microdose is required for heavier participants to show effects for unweighted originality (see Supplementary materials for full analyses output, Section S6.2).

### 6.3.4. Convergent thinking

The result for converged scores showed that there was not a significant difference between the active microdosing condition (\( M = 0.024, SD = 0.852 \)) and placebo (\( M = 0.056, SD = 1.05 \)), \( t(136) = -0.197, p = 0.844, d = 0.034 \). Bayes factor indicated moderate evidence towards the null effect, \( BF_{01} = 5.46 \). Similarly, the regression analysis between the relative dose (dose/participant’s weight) and convergent score was not significant (\( F(1, 138) = 0.01, p = 0.78, R^2 < 0.001 \)) with moderate evidence for null effect, \( BF_{01} = 5.39 \).
Figure 5. Results of Experiment 3 and Mega-analyses, show mean Z-scores for the three divergent scores and one convergent score measured by PCT-d and AUT as the function of group (Placebo vs. Active psilocybin). Vertical capped lines indicate standard error of the mean. Asterisks indicate significant differences between active and placebo sessions with Bonfornni adjusted α level (*p < .0125, **p < .001).

6.3.5. Exploratory analyses of possible mediating factors

Next, we ran a hierarchical regression for each of the four creativity scores separately with demographic information and subjective drug effects as predictors. The data were screened for assumptions of linearity, normality, homoscedasticity, and multicollinearity. All assumptions were sufficiently met, allowing us to proceed with the hierarchical regression. For the full model summaries, assumption checks and coefficient tables, please see the Supplementary materials, mega-analyses (Section S.6).

The multiple linear regressions indicated that demographic and drug effects did not significantly predict any of the four creativity scores in the first step of the analyses, $R^2$s < 0.049, $P$s > 0.062. These results corroborate with previous ANOVA analyses and indicate that drug effects and demographic variables did not significantly predict creativity scores alone. A different outcome was found for the second block analyses. When the drug condition was added as a predictor, the full model became statistically significant for unweighted originality, ($R^2$ =0.117, $p = 0.015$) and originality ratio ($R^2$ =0.178, $p < 0.001$) but did not improve the model for convergent score and fluency score, $R^2$s < 0.095, $P$s > 0.593. A similar effect was found when the relative dose ratio (dose/participant’s weight) was added in the second step instead of the experimental condition. The overall model significantly predicted originality, ($R^2$ =0.085, $p = 0.022$) and originality ratio ($R^2$ =0.102, $p = 0.007$) but non-significant effects were found for convergent score and fluency score, $R^2$s < 0.039, $P$s > 0.315. These
results overall corroborate previous t-test and regression analyses, see Supplementary materials for the complete output of the mega-analyses Supplementary material (Section S.6).

7. Discussion

This study aimed to investigate the longitudinal effects of microdosing psilocybin on convergent and divergent creativity while controlling for expectation/placebo effects. We were specifically interested in determining whether psychedelic microdosing induces cognitive flexibility and, thus, promotes divergent thinking through 5-HT-2A signaling (Freksa et al., 2012; Girn et al., 2020; Carhart-Harris & Nutt, 2017).

Data from 171 attendees collected across three longitudinal placebo-controlled experiments at two experimental sites indicated that microdosing psilocybin positively predicts the overall originality of creative responses (originality ratio). In other words, participants in the active microdosing condition produced a higher quality of divergent responses rather than a higher quantity of divergent answers. No effects of microdosing were found for convergent thinking or the fluency score. The effect on unweighted originality was initially unstable and rendered insignificant after the multiple comparisons correction. Yet, once we controlled for the relative dose size (dose/weight of participants), microdosing significantly predicted the unweighted originality score. These effects were present even after controlling for expectation biases and demographic factors, suggesting that subjective effects did not account for these results alone.

Effect of microdosing on divergent creativity.

The effect of microdosing on originality ratio seems especially relevant, with regards to hypothesis suggesting increased cognitive flexibility in psychedelic states (Carhart-Harris & Nutt, 2017). When people engage in divergent thinking, most common answers are usually reported first and original ideas are generated later on in the divergent process (Beaty & Silvia, 2012; Christensen et al., 1957; Johns et al., 2001; Phillips and Torrance, 1977). This effect is known as a serial-order effect in creativity and could be understood as an increased spread of activation within semantic networks (Kenett, 2018; Mekern et al., 2019). Specifically, the most common associations are readily available within the semantic hierarchy (i.e., heuristics) because they have been previously witnessed/experienced and thus can be easily drawn from long-term memory. The original solutions require a novel, and often a symbolic, combination of pieces of prior knowledge. To find original associations within the semantic hierarchy, participants must rely more heavily on executive functions to search for remote and indirect associations (Gilhooly et al., 2007; Vartanian et al., 2020). Computational models in creativity propose that original solutions are derived through an object’s de-construction into its sub-components, known as object feature maps (Mekern et al., 2019). For example, the overlap between apples, balls, and
balloons is "round", which can be exploited to search for novel, creative solutions. Such representations allow for overlap between features of distinct objects allowing for novel connection between seemingly unrelated concepts. As such, retrieval from long-term memory plays a vital role in generating divergent but relatively common ideas (e.g., using a brick as a chair, using a brick as a weapon) which are represented by the fluency and flexibility scores in AUT. On the other hand, generating original ideas requires higher spread of activation within the semantic networks during which dormant connections can be found through overlap in object representations (e.g., Use a crushed brick powder as a sunblock; use a brick as a prisoner's anchor, use a brick to prevent growth) (Beaty & Silvia, 2012; Mekern et al., 2019; Vartanian et al., 2020). This is relevant as original concepts can be interpreted as indications of high cognitive flexibility, in which wider net of feature representations can be conceived in parallel (Hommel, 2015).

**Previous studies.**

Current findings are congruent with previous studies indicating positive yet limited effects of high doses of psychedelics on divergent thinking (Mason et al., 2021; Mason et al., 2019;). For instance, Zegans and colleagues (1967) conducted a placebo-controlled study with a low dose of LSD and observed that participants in the LSD condition showed more original responses as compared to participants in the placebo condition. Similarly, Frecska et al., (2012) compared the performance on the Torrance task (divergent thinking task) between ayahuasca users and participants without recent psychedelic use and reported that Ayahuasca increased the number of highly original solutions, while other aspects of creativity (fluency and flexibility) were unaffected. The underlying mechanisms by which psychedelics interact with domain-specific creativity remain relatively unexplored. Existing neuroscientific models of creativity predominantly focused on the effect of dopaminergic pathways in convergent and divergent creativity (Akbari et al., 2010; Nijstad et al., 2010; Boot et al., 2017, 2019; Hommel, 2015) and as such microdosing may introduce a novel tool to study role of serotonin in human creativity.

**Mechanistic interpretation.**

While the neuro-cognitive mechanisms in microdosing remain hypothetical, several interpretations can be drawn from existing research conducted with regular psychedelic doses. Specifically, the contemporary theoretical (REBUS) framework proposed by Carhart-Harris and Friston (2019) offers a mechanistic account of psychedelic action, which may account for enhanced creativity. The model proposes that perception is combined through a predictive coding process that combines bottom-up sensory signals and top-down prior beliefs. Predictive processing facilitates more efficient information processing; however, prior beliefs ultimately biases our perception in favour of our prior expectations and as such may lead to confirmation biases and restricted creativity (Carhart-Harris & Friston, 2019; Dietrich, 2004). Psychedelics were proposed to relax such high-level priors and liberate bottom-up signals which can in turn access conscious experience more easily. By reducing the weight of prior
beliefs, psychedelics were suggested to abolish pervasive thoughts and biases, inducing more unconstrained, disinhibited cognition which was previously associated with creativity (Hommel, 2015; Girn et al., 2020). Reduction in top-down priors was further argued to increase sensitivity to bottom-up signals and enhance the level-of-detail of sensory input (Carhart-Harris & Friston, 2019). An increase in bottom-up signals may allow for a broader exploration of the feature-space (e.g., shapes, edges, textures) through which new creative connections can be discovered (Mekern et al., 2019).

In terms of the neural mechanisms, psychedelic were observed to desynchronize neural populations within the prefrontal and associative regions and enhance global functional connectivity (Muthukumaraswamy et al., 2013). Specifically, the internally focused Default Mode Network (DMN) and externally focused Task-Positive Networks (TPN), which are typically negatively correlated (Cole et al., 2012), become less functionally distinct under psychedelics. Relevantly to current findings, previous research suggests that increased switching between functional networks predicts the quality of divergent thinking (Jung et al., 2013). The breakdown of the neuro-cognitive hierarchy in the psychedelic state was hypothesized to enhance the repertoire of possible states and representations that the brain can occupy (Carhart-Harris et al., 2016; Carhart-Harris & Friston, 2019). Such an increase in the brain's global flexibility may, in turn, increase the richness of subjective experience and the volume of sensory and mnemonic information, which can facilitate novel insights. This effect was operationalized by the so-called 'entropic brain hypothesis' (Carhart-Harris, 2018), which asserts that brain function in a psychedelic state is becoming more malleable, unstable and thus more sensitive to its environmental factors. Preliminary evidence for this hypothesis was recently put forward by Mason et al. (2021), who showed that divergent thinking increased seven days after a large dose of psilocybin and was predicted by changes within and between-network connectivity of the default mode network.

In sum, we hypothesize that while large doses of psychedelics acutely induce hyper-flexible, unconstrained cognition, microdosing facilitates only a slight increase in flexibility (i.e., state of low inhibition) and in turn facilitates increase spread of divergent associations (Hommel, 2015).

Subjunctive effects.

Regarding the results on subjective effects we found that repeated microdosing was overall well-tolerated. While several participants requested dose reduction no adverse psychological or physical effects were reported. Subjective drug effects were shown to be dose dependent and diminished over dosing period, independently of experimental condition. Prominent placebo effects were found in all three trials and placebo effects were especially pronounced for those with previous drug experience (in Experiment 1). These effects could be interpreted as a by-product of state-conditioning account of placebo effects (for review please see Schwarz et al., 2016). Effective drugs can be considered a stimulus that creates a particular internal state that researchers consider responsible for the impact of the drug on personal experience and behavior (Büchel et al., 2014). Through accumulating experience with a drug and its effects, the intake of the drug will be accompanied by particular expectations of
these effects, which through response conditioning may induce same or similar psychopharmacological state to the one that intake of the actual drug would create. In other words, expectations need not be interpreted as creating artifacts, but they may come as potent as the real drug effects.

**Limitations**

While results of current study are promising the current study is not without limitations and as such current findings can only be considered preliminary. The first limitation was the current trial’s semi-naturalistic setting, which limited our control over the drug administration. Participants reported the time of dosing at every session but, we could not guarantee the exact time of administration beyond self-reports. Furthermore, as compared to clinical trials utilizing chemically synthesized psilocybin, we could not control for the exact amount of psychedelic alkaloids at every dose. While we tried to account for the relative dose size and post-hoc analyzed drug content in the lab, we were not able to fully control the drug content. Another question of selection-bias arises, because participants who choose to participate in psychedelic trials may be implicitly or explicitly biased. Specifically, even if we controlled for explicit expectation biases our sample mostly consisted of participants with prior psychedelic experience which may not be representative of the general population. For instance, participants with extensive psychedelic experience may have already altered their brain connectivity with psychedelics prior to the study and thus may show a differential effect from the naïve participants. Alternatively, the expectation bias in those with psychedelic experience may induce psychopharmacological reactions mimicking psychedelic effects. Future studies should consider the recruitment of only naive participants and attempt to replicate these findings with more systematic and controlled lab-based drug-administration. Overall, in order to draw more firm conclusions high-powered independent replication would be necessary to establish the robustness of these effects.

### 7.1. Conclusions

In sum, the current study suggests that microdosing does not improve all facets of creativity, yet positive effects of active microdosing were found for the number of original answers. Furthermore, the active condition produced a higher ratio of original solutions as compared to the placebo condition, which suggests a trade-off between quality and quantity of divergent responses. While this study provided some support for enhancing properties of microdosing on divergent quality the effects were more subtle than initially anticipated. Regarding the qualitative effects, prominent placebo effects were found in the ratings of subjective effects in all three trials and were especially pronounced for those with prior psychedelic experience (Experiment 1). Subjective drug effects were shown to be dose dependent and diminished over time, independently of experimental condition. Microdosing may thus present a promising avenue of studying the interaction between psychedelics, cognition, and placebo effects in a more controlled manner. Future well-powered clinical trials are necessary to fully understand and appreciate the impact of microdosing on human creativity and its underlying brain mechanisms.
8. References


