




Natural language signatures of psilocybin microdosing

Camila Sanz¹ · Federico Cavanna^{1,2} · Stephanie Muller¹ · Laura de la Fuente^{1,3} · Federico Zamberlan^{1,4} · Matías Palmucci¹ · Lucie Janeckova⁵ · Martin Kuchar^{5,6} · Facundo Carrillo⁷ · Adolfo M. García^{8,9,10,11,12} · Carla Pallavicini^{1,2} · Enzo Tagliazucchi^{1,13} 

Received: 28 February 2022 / Accepted: 26 May 2022 / Published online: 9 June 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Rationale Serotonergic psychedelics are being studied as novel treatments for mental health disorders and as facilitators of improved well-being, mental function, and creativity. Recent studies have found mixed results concerning the effects of low doses of psychedelics (“microdosing”) on these domains. However, microdosing is generally investigated using instruments designed to assess larger doses of psychedelics, which might lack sensitivity and specificity for this purpose.

Objectives Determine whether unconstrained speech contains signatures capable of identifying the acute effects of psilocybin microdoses.

Methods Natural speech under psilocybin microdoses (0.5 g of psilocybin mushrooms) was acquired from thirty-four healthy adult volunteers (11 females: 32.09 ± 3.53 years; 23 males: 30.87 ± 4.64 years) following a double-blind and placebo-controlled experimental design with two measurement weeks per participant. On Wednesdays and Fridays of each week, participants consumed either the active dose (psilocybin) or the placebo (edible mushrooms). Features of interest were defined based on variables known to be affected by higher doses: verbosity, semantic variability, and sentiment scores. Machine learning models were used to discriminate between conditions. Classifiers were trained and tested using stratified cross-validation to compute the AUC and *p*-values.

Results Except for semantic variability, these metrics presented significant differences between a typical active microdose and the inactive placebo condition. Machine learning classifiers were capable of distinguishing between conditions with high accuracy (AUC ≈ 0.8).

Conclusions These results constitute first evidence that low doses of serotonergic psychedelics can be identified from unconstrained natural speech, with potential for widely applicable, affordable, and ecologically valid monitoring of microdosing schedules.

Keywords Psychedelics · Psilocybin · Microdosing · Language · Machine learning

Introduction

Psychedelic microdosing consists in consuming relatively small amounts of serotonergic compounds, approximating the perceptual threshold—typically, 10–20% of a full dose (Fadiman and Korb 2019; Kuypers et al. 2019; Ona and Bouso 2020; Polito and Stevenson 2019). Unlike higher psychedelic doses, microdosing is expected to produce minimal

acute effects with sustained effects that can last 1 or 2 days. Accordingly, this practice involves interspersing resting days with dosing days (two to four times per week) (Kuypers et al 2019). Microdosing has been gaining popularity over recent years, with lysergic acid diethylamide (LSD) and psilocybin (in the form of psychoactive mushrooms) being the compounds most frequently consumed for this purpose (Hutten et al. 2019a; Lea et al. 2020a, b, c; Polito and Stevenson 2019; Szigeti et al. 2021). Despite the illegal status of psychedelics in most countries, several websites contain discussions and suggestions related to microdosing and its effects, with users claiming that this practice can improve mood, stimulate productivity, and improve cognitive functions as well as mental concentration (Anderson et al. 2019a,

✉ Camila Sanz
camilasanz@gmail.com

✉ Enzo Tagliazucchi
tagliazucchi.enzo@googlemail.com

Extended author information available on the last page of the article

b; Lea et al. 2020a, b, c). Microdosing is also used for the self-treatment of mental health disorders such as depression or anxiety (Hutten et al. 2019b; Lea et al. 2020a, b, c; Lea, et al., 2020a, b, c), and it has been suggested as a model for the clinical use of psychedelics (Kuypers 2020).

The positive effects of microdosing are supported by multiple observational, survey-based, and open-label studies (Anderson et al., 2019a, b; Hutten et al. 2019a; Johnstad 2018; Lea et al. 2020a, b, c; Polito and Stevenson 2019; Prochazkova et al. 2018; Rootman et al. 2021). However, these generally involve self-selected samples and lack adequate control conditions; thus, given the subtle effects of microdosing, expectations and pre-existing traits may play a fundamental role in the perceived effects (Kaertner et al. 2021; Olson et al. 2020). In contrast, studies following double-blind and placebo-controlled experimental designs have found less support for positive outcomes of microdosing (Bershad et al. 2019; Family et al. 2020; Hutten et al. 2020; Szigeti et al. 2021; van Elk et al. 2021; Yanakieva et al. 2019). Yet, these studies generally include tasks and questionnaires validated in the context of high doses of psychedelics, which might lack the specificity and sensitivity necessary to capture the subtler effects induced by microdosing. In turn, these limitations could hinder the development of translational approaches and the validation of therapeutic models, among other potential limitations, thus raising the need for novel methods to study low doses of psychedelic substances.

We adopted an alternative approach based on natural language processing (NLP). As a first step, this approach consisted in extracting linguistic features from unconstrained speech produced by the participants during the acute effects, using them as input to machine learning algorithms. NLP is characterized for being an objective, non-invasive, cost-effective, and scalable tool to investigate ecologically valid data (Sanz et al. 2022; Tagliazucchi 2022). Unlike standard questionnaires, which constrain reports to a possibly sub-optimal pre-selected set of questions, this approach is capable of automatically identifying and capturing informative semantic and grammatical features of speech, allowing to distinguish between different experimental conditions (Agurto et al. 2020; Bedi et al. 2014; Corcoran et al. 2018; Norel et al. 2020; Sanz et al. 2022; Sanz et al. 2021). Importantly, besides informing the contents of the drug-elicited experience, NLP allows to investigate the modulation of language production itself, which can be informative of drug action beyond what is reported during the subjective acute effects (Tagliazucchi 2022). While NLP has been applied to investigate the effects of different compounds, including psychedelics (Bedi et al. 2014; Cox et al. 2021; Cox and Johnson, 2021; Hase et al. 2022; Sanz et al. 2021; Sanz et al. 2018; Zamberlan et al., 2018), there are no works to date applying this tool to the study of microdosing.

We analyzed the effects of microdoses of *Psilocybe cubensis* mushrooms (0.5 g dried material, a typical micro-dose) (Polito and Stevenson 2019; Szigeti et al. 2021; van Elk et al. 2021) on verbal reports. Data was obtained following a double-blind placebo-controlled experimental design with two different measurement weeks per participant (Cavanna et al. 2022). During each week, participants received either the active dose or the placebo and were interviewed about their feelings, expectation, perception, mood, creativity, and alertness. We obtained and analyzed two metrics based on our previous results of NLP applied to speech under the effects of high doses of psychedelics: verbosity (total length of the speech samples, in words) and semantic variability (variability of the time series obtained by computing the similarity between the meaning of consecutive words) (Sanz et al. 2021). We also investigated the mean sentiment score (use of terms linked to positive/negative sentiment) to account for the purported effects microdosing on mood (Cameron et al. 2020; Hutten et al. 2020; Lea et al. 2020a, b, c; Lea et al. 2020a, b, c; Polito and Stevenson 2019). Based on these previous reports, we hypothesized that psychedelic microdoses would increase verbosity and mean sentiment scores, and decrease the semantic coherence of speech. Moreover, we evaluated whether these metrics presented differences between participants who correctly identified the experimental condition and those who did not, which we labeled as *unblinded* and *blinded*, respectively. Finally, we implemented statistical tests and machine learning models to classify the experimental condition (active dose of psilocybin vs. placebo) and the condition identification as given by the participants (blinded vs. unblinded) based on these NLP features.

Methods

Participants

Thirty-four native Spanish-speaking volunteers (11 females: 32.09 ± 3.53 years; 23 males: 30.87 ± 4.64 years) were enrolled, via social media or word of mouth. There were two measurement weeks per participant, separated by at least 1 week. All participants enrolled or obtained a higher education degree, had normal or corrected-to-normal vision, and successfully completed all stages of the experiment.

Participants reported 11 ± 14.9 past experiences with serotonergic psychedelics of which 1.5 ± 2.3 were considered challenging, and 6 participants reported past experience with microdosing. Some participants also reported occasional use of cannabis (20), yerba mate (17), coffee (15), tea (7), tobacco (7), MDMA (1), and cocaine (1). The following exclusion criteria were assessed after an initial interview to present a general overview of the experiment and sign the

informed consent form: past diagnosis of psychotic disorders, bipolar disorders (type 1 or 2), depressive disorders, obsessive–compulsive disorder, anxiety disorder, dysthymia, panic disorder, and/or neurological disorders. Those presenting bulimia, anorexia, and/or substance abuse/dependence over the last 5 years (excluding nicotine) were also excluded. Participants were asked to stop their habitual use of psychoactive drugs (including alcohol and caffeine) during the duration of the experiment.

This study was conducted in accordance with the Helsinki declaration and approved by the Committee for Research Ethics at the Universidad Abierta Interamericana (Buenos Aires, Argentina), protocol number 0–1054. Speech measurement and analysis was part of a larger protocol pre-registered in www.clinicaltrials.gov, identifier number: NCT05160220. All participants provided their written informed consent to participate in the study. Participants did not receive financial compensation for their participation.

Experimental protocol

This experiment followed a double-blind placebo-controlled design, where two conditions (0.5 g of dried *Psilocybe cubensis* and the same weight of an edible mushroom) were randomized and assigned to the 2 weeks of the experiment. All participants underwent both experimental sessions. This assignment was done by a third party and blinded to both participants and researchers.

During Wednesday and Friday of each week of the experiment, volunteers attended the laboratory facilities and consumed either capsules with the active dose or the placebo. On Fridays, approximately 2.30 h after dosing, participants were interviewed by a mental health professional member of the research team (FC). Given that participants completed several tasks on Wednesday (see Cavanna et al. 2022), the interviews were conducted only on Fridays to avoid

confounds due to participant exhaustion. The following questions were asked: “How do you feel right now?”; “Do you feel according to how you expected?”; “Are you feeling changes in your perception?”; “Are you feeling changes in your mood?”; “Are you feeling changes in your level of imagination or creativity?”; “Are you feeling changes in your level of attention, alertness or energy?”. In the following, we refer to these questions as “feeling”, “expectation”, “perception”, “mood”, “creativity”, and “alertness”, respectively. After the interviews conducted on Friday, participants were asked to identify the condition corresponding to that week (either active dose or placebo). Participants were not asked to justify their answer, and the actual condition was not disclosed to the participants until the end of the study (for a schematic representation of the experimental design, see Fig. 1).

Chemical characterization of the mushroom samples

Psilocybe cubensis mushrooms from three independent sources were consumed by the participants of this experiment. The active dose was composed of 0.5 g of ground dried material. Samples of 150 mg from each source of material were isolated and sent for analysis (performed by LJ and MK) to the Laboratory of Forensic Analysis of Biologically Active Substances, University of Chemistry and Technology Prague, Czech Republic, which resulted in the concentration of alkaloids psilocybin, psilocin, baeocystin, and norbaeocystin, averaged across samples.

Analysis of mushroom samples revealed the average concentration of the following alkaloids: psilocybin (640.2 µg/g), psilocin (950.7 µg/g), baeocystin (50.4 µg/g), and norbaeocystin (12.5 µg/g). Therefore, the active dose contained 0.32 mg of psilocybin, 0.48 mg of psilocin, 0.025 mg of baeocystin, and 0.0063 mg of norbaeocystin.

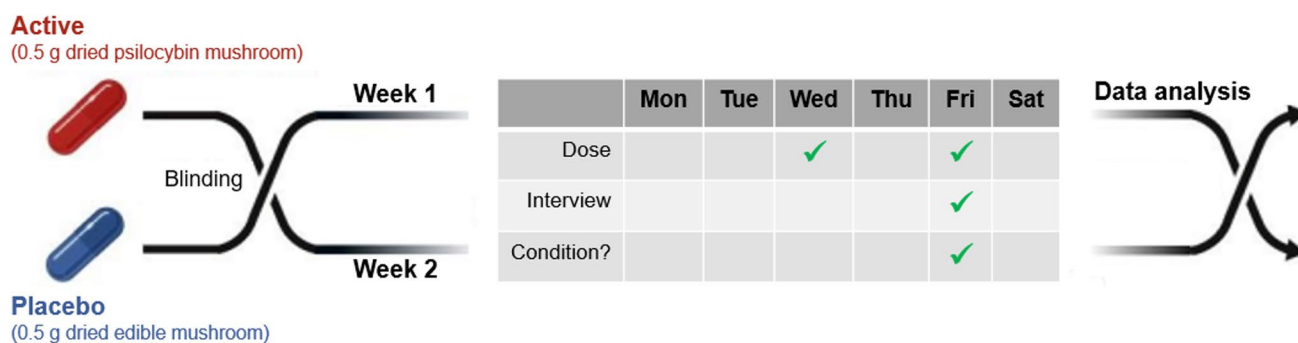


Fig. 1 Double-blind placebo-controlled experimental design. The first week, participants were either randomly assigned the active dose or the placebo (in the example shown in the figure, the active condition was assigned to the second week). The remaining experimental condition corresponded to a following week. Researchers ignored

the content of the capsules until the analysis stage. The following schedule was repeated on both weeks: on Wednesday and Friday, participants consumed the dose, and only on Friday subjects were interviewed and asked to identify the experimental condition

Since at least 1 month passed between the end of the experiment and the chemical analysis, despite conserving the samples under optimal conditions, psychoactive material (i.e., psilocybin and psilocin) may have been lost during this time period (Gotvaldova et al. 2021).

Speech analysis

Transcript preprocessing and labeling

Interviews were recorded and manually transcribed by a technician with background in linguistics who was blind to the experimental condition, and afterwards reviewed by a member of the research team as a quality check (both were native Spanish speakers). The parts corresponding to the interviewer were removed from the transcripts, leaving only the parts corresponding to the participants. Interview questions were analyzed separately to determine which specific topics were affected by microdosing. Each transcript was labeled as active dose or placebo, depending on the experimental condition, and as blinded or unblinded (depending on whether the participant correctly identified the condition). This resulted in four possible combinations (blinded active dose, blinded placebo, unblinded active dose, and unblinded placebo).

Verbosity

We computed verbosity scores by counting the number of words (including repetitions and stopwords) produced by the participants when answering the questions of the interview.

Semantic variability

We computed a metric of semantic variability based on the distance between consecutive words spoken by the subjects. Using Python's TreeTagger library with AnCorra Spanish corpus (<http://clic.ub.edu/corpus/es/ancora>), we applied a series of pre-processing steps including splitting the transcripts into its individual words, removal of punctuation marks and other symbols, conversion to lowercase, removal of stopwords, and lemmatization (i.e., conversion of each word to its base form). Next, we used the FastText model (<https://fasttext.cc>) pre-trained with the Common Crawl and the Wikipedia corpus in Spanish (comprising around 2,000,000 unique words) to obtain the word embedding. The embedding assigns each word a vector, so that terms with similar meanings are mapped into vectors that are close in the embedding space. Therefore, the semantic distance between words can be obtained as the cosine angle between their corresponding vectors (Corcoran et al. 2018). Afterwards, each term in the pre-processed transcripts was identified by its vector and the semantic distance between

consecutive terms was computed; thus, the transcripts were represented as time series containing the cosine distances between consecutive terms. Finally, the semantic variability was computed by taking the variance of these time series (Sanz et al. 2021). We used word instead of sentence embeddings since the answers provided by the participants consisted of comparatively few sentences (3.6 ± 3.3 sentences), which was considered insufficient to adequately estimate the variance. Previous work investigating the acute effects of medium/high doses of serotonergic psychedelics showed that semantic variability is increased during the acute effects relative to the placebo (Sanz et al. 2021; Wießner et al. 2021).

Sentiment analysis

Sentiment analysis is an NLP technique used to determine whether text data conveys positive, negative, or neutral sentiment, with applications to quantify the content of experiences elicited by psychedelics (Qiu and Minda 2021). Sentiment analysis of the speech transcripts was implemented in Python using a sentence-level model pre-trained for Spanish (github.com/ayllote/senti-py). Briefly, the model pipeline included standard text pre-processing steps such as conversion to lowercase, accent removal, standardization of linguistic expressions typical of regional Spanish, conversion of verbs to their infinitive forms, introduction of bigrams consisting of adjectives and their negation, and removal of custom words considered useless for the quantification of sentiment. These steps were followed by a univariate feature selector and a multinomial naive Bayes classifier trained using data obtained from websites in Spanish that offer services and allow users to comment and rate their experience ("TripAdvisor", "Pedidos Ya", "MercadoLibre", among others), comprising approximately 1 million training samples. The parameters and hyper-parameters of this pipeline were found by the use of a grid search combined with K-fold cross validation ($K=10$). Based on its words, the model assigns to each sentence a score ranging from 0 (negative) to 1 (positive), with a score of ≈ 0.5 assigned to neutral sentences.

The transcript of each response was first split into individual sentences, which were used as input for the sentiment analysis model. Finally, the average score across all sentences was computed, resulting in a mean sentiment score (MSS) for the answers to each question in the interview. The following three sentences extracted from the transcripts illustrate cases with positive, neutral, and negative sentiment, according to the model:

"I feel tired, I've been tired all day, I don't believe that... I keep thinking that I took two placebos". Negative sentiment (MSS = 0.066).

“No, I think it was quite normal eh... like overall I feel that I was calmer as when the stuff that normally happens presented itself, I could sort of see it as one step further back you know like more relaxed, but aside from that it was quite normal, I didn't see any big changes”. Neutral sentiment (MSS = 0.49). “Now complete peace of mind”. Positive sentiment (MSS = 0.83).

Statistical analyses

Statistical analyses were conducted using non-parametric tests to evaluate differences in the median verbosity, semantic variability, and MSS. Wilcoxon signed-rank tests were applied for active dose vs. placebo and the comparisons between the corresponding subgroups (active dose vs. placebo restricted to blinded and unblinded subgroups), which are within-subject comparisons. Mann–Whitney U tests were used for blinded vs. unblinded and the comparisons between the corresponding subgroups (blinded vs. unblinded restricted to active dose and placebo). For both tests, the alpha level (0.05) of statistical significance was corrected via the Bonferroni criterion with $N=6$ (number of questions in the interview).

We obtained summary statistics (means and variances) from two previous studies investigating similar NLP metrics in psychiatric patients and in healthy subjects under the acute effects of psychedelics (Elvevag et al. 2007; Sanz et al. 2021). To determine the relationship between sample size and statistical power, we used the formula $n = \frac{(\sigma_1^2 + \sigma_2^2)(z_{1-\alpha/2} + z_{1-\beta})^2}{\Delta^2}$ (Rosner 2015), where n is the sample size, $\alpha = 0.05$ is the type 1 error probability, β is the type 2 error probability (one minus the statistical power), Δ is the absolute difference between means, $\sigma_{1,2}^2$ are the variances, and z is the critical Z value corresponding to the chosen α and β . Using these parameters, we confirmed that $n = 34$ resulted in statistical power greater than 0.8 (equivalently, $\beta < 0.2$).

Boxplot representations for each answer and group were used to graphically summarize the results. Each boxplot comprises from the 25th percentile (lower quartile) to the 75th percentile (upper quartile); whiskers extend from these percentiles up to 1.5 times the interquartile range with a line indicating median values. Participants' responses were represented with single points above the boxplot, except for outliers (below/above the lower/upper quartile ± 1.5 times the interquartile range) which were removed.

Statistical analyses were conducted using Python's SciPy library (<https://scipy.org>).

Machine learning analysis

Machine learning models were trained using the scikit-learn (scikit-learn.org) library. In total, each sample consisted of 12 features, given by the verbosity scores and the MSS for each question in the interview (the semantic variability was excluded since it was not significantly different between conditions, see results and supplementary material). Random forest classifiers (1000 estimators with default parameters, scikit-learn.org) were used to classify these samples according to the experimental condition and the condition identification given by participants. This classifier was chosen due to its good performance without need for fine parameter tuning, allowing to avoid the need for a separate testing dataset. In total, six models (the same hyper-parameters and training–testing procedure) were implemented, two trained to distinguish between experimental conditions (psilocybin vs. placebo) and blinded vs. unblinded, and the remaining four trained to perform the same classification, but restricted to subgroups following the four possible combinations: blinded active dose, blinded placebo, unblinded active dose, and unblinded placebo.

For each classifier, samples were divided into 3 folds via stratified cross-validation to preserve the proportion of labels. Then, each of the folds was used at least once for testing while the remaining were used to train the model; the average of the area under the ROC curve (ROC AUC) across folds was used to determine the model performance. This procedure was repeated 100 times with and without shuffling of the sample labels. This shuffling was used to break the relationship between the features and the labels, thus resulting in a null-model used to estimate chance level performance. Next, the number of times the AUC of the model trained with shuffled labels exceeded the value obtained without label shuffling divided by the total number of iterations (100) resulted in a p -value representing the probability of obtaining the given ROC AUC value assuming chance distribution of the sample labels between classes.

Results

Statistical analysis and machine learning classification

Active dose vs. placebo

We did not find statistically significant differences in the semantic variability when comparing the psilocybin vs. placebo conditions (for further details, see the supplementary material).

Participants under the active dose exhibited higher verbosity than under placebo condition in their answers to all

questions in the interview (Fig. 2A). Significant differences were found for the answers related to the questions “perception” ($T=67$, $p=0.0034$, $r=0.67$), “mood” ($T=51.5$, $p=0.0016$, $r=0.75$), and “alertness” ($T=79$, $p=0.0047$, $r=0.61$). The remaining answers raised p -values larger than 0.026 and did not survive Bonferroni corrections. For the group under the active dose, median values and interquartile ranges were $M=63.5$, $IQR=82.25$ (“perception”); $M=50$, $IQR=85.5$ (“mood”); and $M=65$, $IQR=60.5$ (“alertness”), and for the group under the placebo condition $M=10.5$, $IQR=39.75$ (“perception”); $M=10$, $IQR=19.25$ (“mood”); $M=33$, $IQR=63.5$ (“alertness”).

Median values of MSS for participants under the active dose were around 0.4, suggesting an overall neutral state (Fig. 2B). Participants under the microdosing condition scored higher than under the placebo in all their answers, indicating more positive sentiment. Significant differences were found for the answer related to “perception” ($T=72$, $p=0.0039$, $r=0.65$); the other answers resulted in p -values larger than 0.031 and did not survive Bonferroni corrections. The median value and interquartile range for the group under the active dose was $M=0.37$, $IQR=0.11$ and for the group under the placebo condition $M=0.22$, $IQR=0.19$.

These results suggest that the active dose generated subtle effects leading to longer answers and increased positivity

in the speech of the participants when reporting changes to perception, mood, and alertness. Answers related to general feelings, expectations, and creativity remained the same regardless of the experimental condition.

Classification for the instance without label shuffling yielded an AUC value of 0.79 ± 0.04 , which was significantly higher ($p < 0.01$) than the one found after label shuffling (0.52 ± 0.10), indicating classifier accuracy above chance level. Histograms representing the AUC values across all iterations are shown in Fig. 2C.

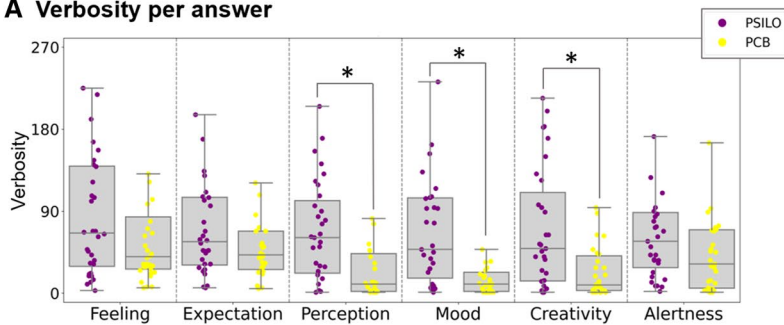
Blinded vs. unblinded

Pairwise comparisons between blinded and unblinded participants did not result in significant differences for any of the answers in the interview, yielding p -values larger than 0.013 and 0.1 for the verbosity and MSS, respectively. Participants in the blinded group showed higher median verbosity values (Fig. 3A) and lower MSS (Fig. 3B) than participants in the unblinded group, but neither of these changes were significant at $p=0.05$ with Bonferroni correction for multiple comparisons.

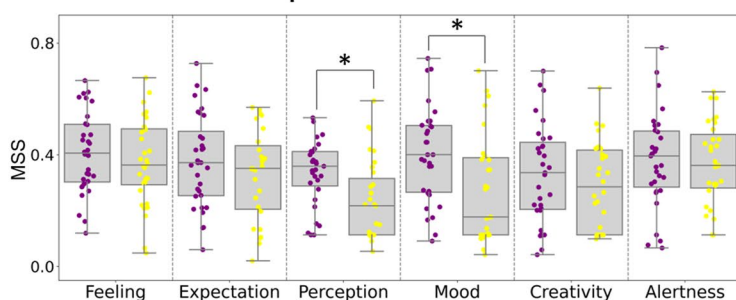
The results of the machine learning analysis (Fig. 3C) were consistent with the lack of statistically significant differences between groups, yielding an AUC value of

PSILO vs PCB (ALL)

A Verbosity per answer



B Mean sentiment score per answer



C Machine learning classifier

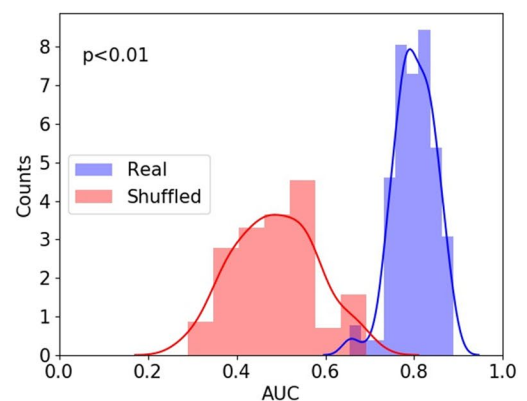
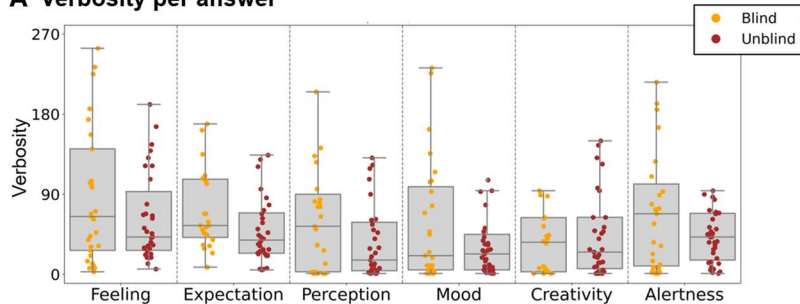


Fig. 2 The active dose (PSILO) increased verbosity and positive sentiment relative to the placebo (PCB). **(A)** Boxplots of the verbosity for each question and condition. Significant pairwise differences are indicated with an asterisk ($*p < 0.05$, Bonferroni corrected for mul-

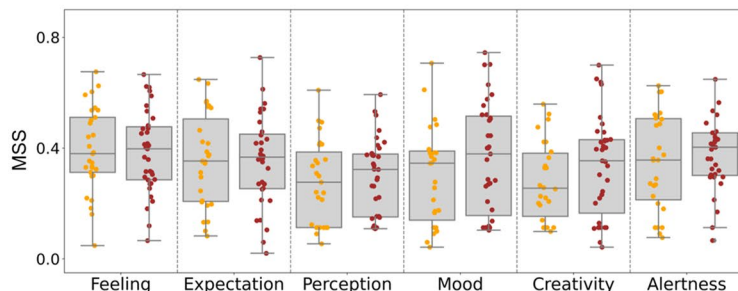
iple comparisons, $N=6$; Wilcoxon signed-rank test). **(B)** The same information as in panel A but for the MSS. **(C)** Histograms of AUC values obtained from the classification with (red) and without (blue) label shuffling across 100 iterations

Blind vs Unblind (ALL)

A Verbosity per answer



B Mean sentiment score per answer



C Machine learning classifier

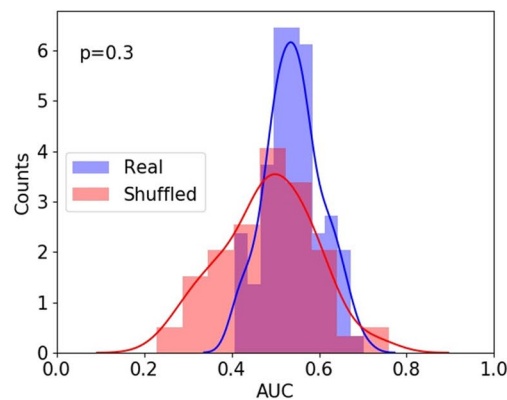


Fig. 3 The analysis of verbosity and MSS did not reveal significant differences between blinded and unblinded participants. **(A)** Boxplots of the verbosity for each question and condition. Significant pairwise differences are indicated with an asterisk ($*p < 0.05$, Bonferroni cor-

rected for multiple comparisons, $N=6$; Mann–Whitney U test). **(B)** The same information as in panel **A** but for the MSS. **(C)** Histograms of AUC values obtained from the classification with (red) and without (blue) label shuffling, across 100 iterations

0.53 ± 0.06 , which did not differ ($p = 0.37$) from the one obtained after label shuffling (0.49 ± 0.11).

Statistical analyses and classifiers restricted to subgroups of participants/conditions

Statistical and machine learning analyses were repeated for all possible combinations of subgroups: blinded active dose, blinded placebo, unblinded active dose, and unblinded placebo. Wilcoxon signed-rank and Mann–Whitney U tests were used to compare verbosity, semantic variability, and sentiment scores for the subgroups. The performance of the classifiers is summarized in Fig. 4. Boxplots corresponding to these results, together with statistical analyses, are included in the supplementary material.

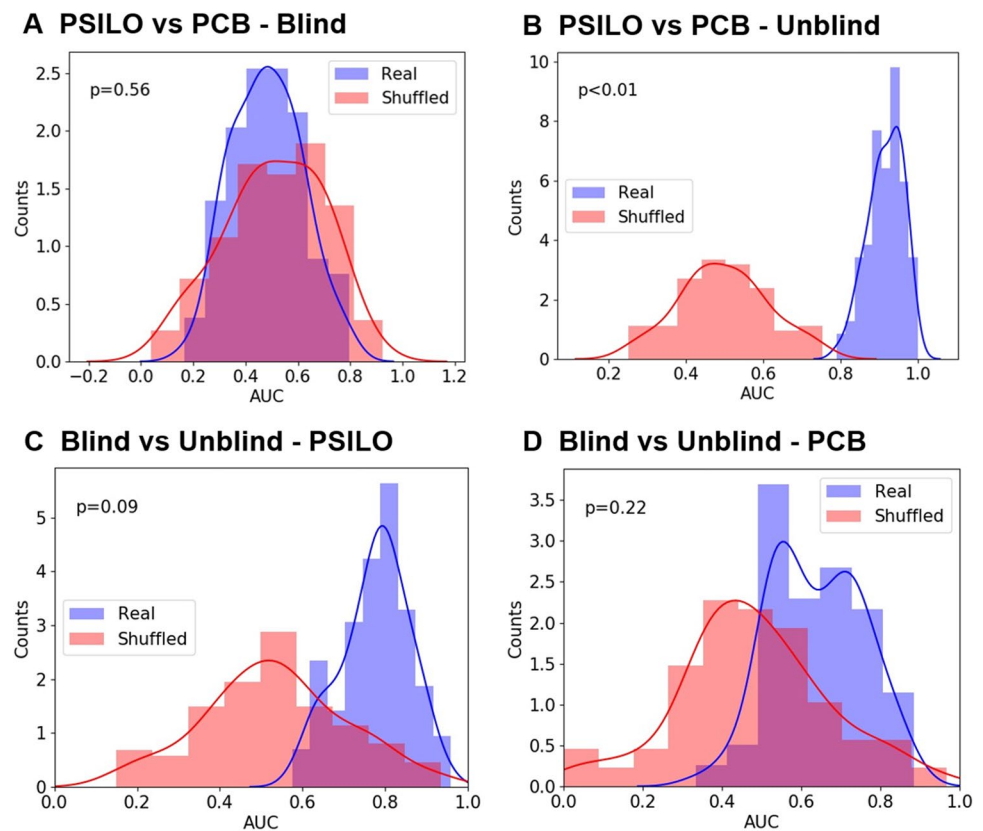
The classifier trained to distinguish between the active dose and the placebo condition, but restricted to blinded participants only (Fig. 4A), yielded an AUC value of 0.48 ± 0.13 . This did not differ significantly ($p = 0.56$) from the value obtained by shuffling labels (0.52 ± 0.18). In contrast, when restricting the samples to unblinded participants (Fig. 4B), both conditions were robustly classified as compared to the classifier with label shuffling ($p < 0.01$). AUC values were 0.51 ± 0.17 and 0.92 ± 0.05 for the instances with and without shuffling labels, respectively.

Machine learning results for blinded vs. unblinded, but restricted to the active dose and the placebo conditions only (Fig. 4C and D), did not present statistical differences as compared with the classifiers with label shuffling ($p = 0.09$ and $p = 0.22$, respectively). Nevertheless, classification of blinded vs. unblinded participants limited to the active dose condition presented a p -value close to the alpha level of significance. Under this restriction, AUC values were 0.50 ± 0.18 and 0.76 ± 0.08 for the instances with and without label shuffling. When restricted to the placebo condition, AUC values were 0.50 ± 0.18 and 0.63 ± 0.17 , with and without shuffling labels, respectively.

Discussion

We explored the effects of psychedelic microdosing on natural speech via measures of semantic variability, verbosity, and sentiment scores. The last two measures discriminated between the active dose and the placebo condition. Furthermore, random forest classifiers successfully distinguished between groups based on this information. In contrast, no significant differences in these measures were found between participants who correctly identified the experimental condition and those who did not; consistently, machine learning

Fig. 4 Only the classifier between the active dose and the placebo condition restricted to unblinded participants yielded a significant accuracy ($p < 0.05$), as compared with the instance with label shuffling. Histograms of the AUC values obtained from classification with (red) and without (blue) label shuffling for all the following subgroups: **(A)** active dose vs. placebo restricted to blinded participants, **(B)** active dose vs. placebo restricted to unblinded participants, **(C)** blinded vs. unblinded participants, restricted to the active dose, **(D)** blinded vs. unblinded participants, restricted to the placebo condition



classification yielded an AUC value close to chance level. Finally, statistically significant differences and a robust classification between active dose and placebo were achieved for participants who correctly identified the experimental condition.

Previous work has shown that higher doses of LSD increase verbosity as well as semantic variability (i.e., reduces speech coherence) (Carhart-Harris et al. 2014; Sanz et al. 2021; Wießner et al. 2021), which is consistent with the general hypothesis of more disordered or entropic brain activity elicited by psychedelics (Carhart-Harris et al. 2014). Insofar as semantic variability indirectly reflects the organization of the stream of thoughts, these previous results suggest that psychedelics result in a hyperassociative state, which in turn might facilitate creativity (Girn et al. 2020). Our results show that semantic variability is not affected by low doses of psilocybin, questioning whether microdosing is capable of enhancing specific aspects of cognition related to creativity through a scrambling effect such as the one postulated for higher doses.

Concerning the results of sentiment analysis, microdosing users generally report an improvement in their mood (Anderson et al. 2019a, b; Hutten et al. 2019a; Johnstad 2018; Lea et al. 2020a, b, c; Polito and Stevenson, 2019; Prochazkova et al. 2018; Rootman et al. 2021). Accordingly, increased MSS of natural speech could reflect the

positive effect of psilocybin on mood and subjective well-being. There is evidence showing links between positive sentiment (assessed using automated NLP tools) and subjective well-being. This research shows that sentiment analysis can be used to develop message-level predictive models for multiple components of well-being (Schwartz et al. 2016). Moreover, sentiment analysis can be used to infer variations in subjective well-being and mood attributed to external factors, such as weather (Hannak et al. 2012) and urban environments (Plunz et al. 2019). Interestingly, increased MSS was not only observed in the answer to the question about mood included in the interview, but also in the answer to other questions. This suggests that microdosing could be capable of inducing a state of positive mood, which generally affects verbal expression, and might be indicative of improved mental health (Babu and Kanaga, 2022). The same applies to increased verbosity, which could reflect more enthusiasm, motivation, and energy during the acute effects of the microdose (Sanz et al. 2021). Considered separately, these results are compatible with alternative scenarios such as increased verbosity due to nervousness, or increased MSS due to the inclusion of positive terms in the speech that are not directly implicated with the mood of the participants. However, when considered together, they support a synergistic interpretation favoring the induction of a state of

positive mood, even though additional experiments should be conducted to exclude alternative interpretations.

Labeling participants as blinded or unblinded revealed that (approximately) half of them correctly identified the condition, a percentage consistent with the chance level. Nevertheless, we note that speech-based features allowed to classify between placebo and psilocybin when the analysis was restricted to participants who correctly identified the condition (Fig. 4B), which did not happen for those who failed in the identification (Fig. 4A). This suggests a possible effect of the subjective effects and the corresponding changes in speech in the correct unblinding of the experimental condition; however, more data is necessary to formally establish this association.

Comparisons between blinded and unblinded participants suggest that expectations could play an important role in the perceived effects of microdosing. This could explain the lack of significant differences and poor classifier performance in the comparison between placebo and active dose, restricted to participants who did not correctly identify the experimental condition. However, expectation effects were not apparent in the comparison between blinded and unblinded participants restricted to the active dose, suggesting that microdosing could generate effects that cannot be fully explained by the identification of the experimental condition. The issue of unblinding is pervasive to all studies of psychedelic microdosing and, more generally, to the study of compounds capable of eliciting profound alterations in the state of consciousness (Kuypers et al. 2019; Muthukumaraswamy et al. 2021; Schenberg 2021; Szigeti et al. 2021; van Elk et al. 2021). Future studies should explore more adequate control conditions and experimental paradigms capable of alleviating these concerns. Also, the participants were asked to identify the condition before the interview, and neither confidence levels nor justifications were requested. Changing the schedule or asking subjects to justify their answer could be useful to further understand the relationship between changes in speech production and unblinding.

Natural language processing tools allowed us to reveal significant differences and to obtain a robust classification of the conditions without resorting to questionnaires that were not formulated with the specific objective of studying low doses of psychedelics. Other advantages of these methods include scalability, low implementation cost, and capacity to process large volumes of unstructured data produced under conditions of ecological validity (Tagliazucchi 2022). Because of these advantages, NLP tools could be useful to provide long-distance guidance for individuals following future therapeutic protocols with serotonergic psychedelics (Kuypers 2020). On the other hand, the process of conducting interviews could be lengthy and time consuming, representing a limitation unless novel tools to automate the process are developed.

Our study presents some limitations stemming from its experimental design. We considered microdosing effects over periods of 1 week; thus, long-term outcomes associated with cumulative effects cannot be captured by our speech analysis. However, since NLP measures can be extracted remotely and automatically, this issue could be mitigated by asking participants to self-record short speech samples and then to submit them for analysis. This observation also highlights the potential applicability of these findings as a tool to monitor the effects of microdosing based on samples of ecological validity. While we only considered semantic variability, verbosity (which are increased by higher doses of psychedelics) (Sanz et al. 2021), and the MSS (due to reports of mood enhancements induced by microdosing), future studies should explore other specific measures that could more adequately capture the effects of microdosing on cognition and mental function. Also, the acoustic analysis of speech samples (e.g., prosody) could yield valuable information, but remains largely unexplored in the context of psychedelics, both for low and high doses (Tagliazucchi 2022).

In conclusion, we characterized natural language produced under effects of low doses of psilocybin, extracting markers from unstructured and unconstrained speech that are compatible with improved mood of the participants, and which might be difficult to capture using more traditional methods. These results highlight the value of recording brief samples of natural language before, during, and after the acute effects of psychedelic compounds.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00213-022-06170-0>.

Acknowledgements We thank Sol Pérez Vázquez for her assistance with the logistics of this study.

Funding This study was funded by grant PICT-2019-02294 awarded by Agencia Nacional de Promoción Científica y Tecnológica (Argentina). Partial funding was received from the Ministry of Health of the Czech Republic (grant no. NU21-04-00307) and the Grant Agency of the Czech Republic (grant no. 20-25349S). Adolfo García is an Atlantic Fellow at the Global Brain Health Institute (GBHI) and is supported with funding from GBHI, Alzheimer's Association, and Alzheimer's Society (Alzheimer's Association GBHI ALZ UK-22-865742); ANID, FONDECYT Regular [1210176]; and Programa Interdisciplinario de Investigación Experimental en Comunicación y Cognición (PIIECC), Facultad de Humanidades, USACH.

Declarations

Conflict of interest The authors declare no competing interests.

References


Agurto C, Cecchi GA, Norel R, Ostrand R, Kirkpatrick M, Baggett MJ, Wardle MC, Wit H, Bedi G (2020) Detection of acute

- 3,4-methylenedioxyamphetamine (MDMA) effects across protocols using automated natural language processing. *Neuropsychopharmacology* 45(5):823–832. <https://doi.org/10.1038/s41386-020-0620-4>
- Anderson T, Petranker R, Christopher A, Rosenbaum D, Weissman C, Dinh-Williams LA, Hui K, Hapke E (2019a) Psychedelic microdosing benefits and challenges: an empirical codebook. *Harm Reduct J* 16(1):43. <https://doi.org/10.1186/s12954-019-0308-4>
- Anderson T, Petranker R, Rosenbaum D, Weissman CR, Dinh-Williams LA, Hui K, Hapke E, Farb NAS (2019b) Microdosing psychedelics: personality, mental health, and creativity differences in microdosers. *Psychopharmacology* 236(2):731–740. <https://doi.org/10.1007/s00213-018-5106-2>
- Babu NV, Kanaga EGM (2022) Sentiment analysis in social media data for depression detection using artificial intelligence: a review. *SN Comput Sci* 3(1):74. <https://doi.org/10.1007/s42979-021-00958-1>
- Bedi G, Cecchi GA, Slezak DF, Carrillo F, Sigman M, de Wit H (2014) A window into the intoxicated mind? Speech as an index of psychoactive drug effects. *Neuropsychopharmacology* 39(10):2340–2348. <https://doi.org/10.1038/npp.2014.80>
- Bershad AK, Schepers ST, Bremmer MP, Lee R, de Wit H (2019) Acute subjective and behavioral effects of microdoses of lysergic acid diethylamide in healthy human volunteers. *Biol Psychiatry* 86(10):792–800. <https://doi.org/10.1016/j.biopsych.2019.05.019>
- Cameron LP, Nazarian A, Olson DE (2020) Psychedelic microdosing: prevalence and subjective effects. *J Psychoactive Drugs* 52(2):113–122. <https://doi.org/10.1080/02791072.2020.1718250>
- Carhart-Harris RL, Leech R, Hellyer PJ, Shanahan M, Feilding A, Tagliazucchi E, Chialvo DR, Nutt D (2014) The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front Hum Neurosci* 8:20. <https://doi.org/10.3389/fnhum.2014.00020>
- Cavanna F, Muller S, de la Fuente LA, Zamberlan F, Palmucci M, Janeckova L, Kuchar M, Pallavicini C, Tagliazucchi ER (2022) Microevidence for microdosing with psilocybin mushrooms: a double-blind placebo-controlled study of subjective effects, behavior, creativity, perception, cognition, and brain activity. *Biorxiv*
- Corcoran CM, Carrillo F, Fernandez-Slezak D, Bedi G, Klim C, Javitt DC, Bearden CE, Cecchi GA (2018) Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry* 17(1):67–75. <https://doi.org/10.1002/wps.20491>
- Cox DJ, Garcia-Romeu A, Johnson MW (2021) Predicting changes in substance use following psychedelic experiences: natural language processing of psychedelic session narratives. *Am J Drug Alcohol Abuse* 47(4):444–454. <https://doi.org/10.1080/00952990.2021.1910830>
- Cox DJ, Johnson MW (2021) Verbal behavior related to drug reinforcement in polysubstance cannabis users: comparison across drugs. *Exp Clin Psychopharmacol*. <https://doi.org/10.1037/pha0000404>
- Elvevag B, Foltz PW, Weinberger DR, Goldberg TE (2007) Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophr Res* 93(1–3):304–316. <https://doi.org/10.1016/j.schres.2007.03.001>
- Fadiman J, Korb S (2019) Might microdosing psychedelics be safe and beneficial? An Initial Exploration. *J Psychoactive Drugs* 51(2):118–122. <https://doi.org/10.1080/02791072.2019.1593561>
- Family N, Maillet EL, Williams LTJ, Krediet E, Carhart-Harris RL, Williams TM, Nichols CD, Goble DJ, Raz S (2020) Safety, tolerability, pharmacokinetics, and pharmacodynamics of low dose lysergic acid diethylamide (LSD) in healthy older volunteers. *Psychopharmacology* 237(3):841–853. <https://doi.org/10.1007/s00213-019-05417-7>
- Girn M, Mills C, Roseman L, Carhart-Harris RL, Christoff K (2020) Updating the dynamic framework of thought: creativity and psychedelics. *Neuroimage* 213:116726. <https://doi.org/10.1016/j.neuroimage.2020.116726>
- Gotvaldova K, Hajkova K, Borovicka J, Jurok R, Cihlarova P, Kuchar M (2021) Stability of psilocybin and its four analogs in the biomass of the psychotropic mushroom *Psilocybe cubensis*. *Drug Test Anal* 13(2):439–446. <https://doi.org/10.1002/dta.2950>
- Hannak A, Anderson E, Barrett LF, Lehmann S, Mislove A, Riedewald M (2012) Tweetin' in the rain: Exploring societal-scale effects of weather on mood. In *Proceedings of the International AAAI Conference on Web and Social Media* 6(1):479–482
- Hase A, Erdmann M, Limbach V, Hasler G (2022) Analysis of recreational psychedelic substance use experiences classified by substance. *Psychopharmacology* 239:643–659
- Hutten N, Mason NL, Dolder PC, Kuypers KPC (2019a) Motives and side-effects of microdosing with psychedelics among users. *Int J Neuropsychopharmacol* 22(7):426–434. <https://doi.org/10.1093/ijnp/pyz029>
- Hutten N, Mason NL, Dolder PC, Kuypers KPC (2019b) Self-rated effectiveness of microdosing with psychedelics for mental and physical health problems among microdosers. *Front Psychiatry* 10:672. <https://doi.org/10.3389/fpsy.2019.00672>
- Hutten N, Mason NL, Dolder PC, Theunissen EL, Holze F, Liechi ME, Feilding A, Ramaekers JG, Kuypers KPC (2020) Mood and cognition after administration of low LSD doses in healthy volunteers: a placebo controlled dose-effect finding study. *Eur Neuropsychopharmacol* 41:81–91. <https://doi.org/10.1016/j.euronuro.2020.10.002>
- Johnstad PG (2018) Powerful substances in tiny amounts: an interview study of psychedelic microdosing. *Nordisk Alkohol Nark* 35(1):39–51. <https://doi.org/10.1177/1455072517753339>
- Kaertner LS, Steinborn MB, Kettner H, Spriggs MJ, Roseman L, Buchborn T, Balaet M, Timmermann C, Erritzoe D, Carhart-Harris RL (2021) Positive expectations predict improved mental-health outcomes linked to psychedelic microdosing. *Sci Rep* 11(1):1941. <https://doi.org/10.1038/s41598-021-81446-7>
- Kuypers KP, Ng L, Erritzoe D, Knudsen GM, Nichols CD, Nichols DE, Pani L, Soula A, Nutt D (2019) Microdosing psychedelics: more questions than answers? An overview and suggestions for future research. *J Psychopharmacol* 33(9):1039–1057. <https://doi.org/10.1177/0269881119857204>
- Kuypers KPC (2020) The therapeutic potential of microdosing psychedelics in depression. *Ther Adv Psychopharmacol* 10:2045125320950567. <https://doi.org/10.1177/2045125320950567>
- Lea T, Amada N, Jungaberle H (2020a) Psychedelic microdosing: a subreddit analysis. *J Psychoactive Drugs* 52(2):101–112. <https://doi.org/10.1080/02791072.2019.1683260>
- Lea T, Amada N, Jungaberle H, Schecke H, Klein M (2020b) Microdosing psychedelics: motivations, subjective effects and harm reduction. *Int J Drug Policy* 75:102600. <https://doi.org/10.1016/j.drugpo.2019.11.008>
- Lea T, Amada N, Jungaberle H, Schecke H, Scherbaum N, Klein M (2020c) Perceived outcomes of psychedelic microdosing as self-managed therapies for mental and substance use disorders. *Psychopharmacology* 237(5):1521–1532. <https://doi.org/10.1007/s00213-020-05477-0>
- Muthukumaraswamy SD, Forsyth A, Lumley T (2021) Blinding and expectancy confounds in psychedelic randomized controlled trials. *Expert Rev Clin Pharmacol* 14(9):1133–1152. <https://doi.org/10.1080/17512433.2021.1933434>
- Norel R, Agurto C, Heisig S, Rice JJ, Zhang H, Ostrand R, Wacnik PW, Ho BK, Ramos VL, Cecchi GA (2020) Speech-based characterization of dopamine replacement therapy in people with Parkinson's disease. *NPJ Parkinsons Dis* 6:12. <https://doi.org/10.1038/s41531-020-0113-5>

- Olson JA, Suissa-Rocheleau L, Lifshitz M, Raz A, Veissiere SPL (2020) Tripping on nothing: placebo psychedelics and contextual factors. *Psychopharmacology* 237(5):1371–1382. <https://doi.org/10.1007/s00213-020-05464-5>
- Ona G, Bousso JC (2020) Potential safety, benefits, and influence of the placebo effect in microdosing psychedelic drugs: a systematic review. *Neurosci Biobehav Rev* 119:194–203. <https://doi.org/10.1016/j.neubiorev.2020.09.035>
- Polito V, Stevenson RJ (2019) A systematic study of microdosing psychedelics. *PLoS One* 14(2):e0211023. <https://doi.org/10.1371/journal.pone.0211023>
- Plunz RA et al (2019) Twitter sentiment in New York City parks as measure of well being. *Landsc Urban Plan* 189:235–246
- Prochazkova L, Lippelt DP, Colzato LS, Kuchar M, Sjoerds Z, Hommel B (2018) Exploring the effect of microdosing psychedelics on creativity in an open-label natural setting. *Psychopharmacology* 235(12):3401–3413. <https://doi.org/10.1007/s00213-018-5049-7>
- Qiu TT, Minda JP (2021) Recreational psychedelic users frequently encounter complete mystical experiences: trip content and implications for wellbeing. *psyarxiv*
- Rootman JM, Kryskow P, Harvey K, Stamets P, Santos-Brault E, Kuypers KPC, Polito V, Bourzat F, Walsh Z (2021) Adults who microdose psychedelics report health related motivations and lower levels of anxiety and depression compared to non-microdosers. *Sci Rep* 11(1):22479. <https://doi.org/10.1038/s41598-021-01811-4>
- Rosner B (2016) *Fundamentals of biostatistics*. Cengage Learning, Boston, MA
- Sanz C, Carrillo F, Slachevsky A, Forno G, Gorno Tempini ML, Villagra R, Ibañez A, Tagliazucchi E, García AM (2022) Automated text-level semantic markers of Alzheimer's disease. *Alzheimer Dement Diagn Assess Dis Monit* 14:e12276
- Sanz C, Pallavicini C, Carrillo F, Zamberlan F, Sigman M, Mota N, Copelli M, Ribeiro S, Nutt D, Carhart-Harris R, Tagliazucchi E (2021) The entropic tongue: disorganization of natural language under LSD. *Conscious Cogn* 87:103070. <https://doi.org/10.1016/j.concog.2020.103070>
- Sanz C, Zamberlan F, Erowid E, Erowid F, Tagliazucchi E (2018) The experience elicited by hallucinogens presents the highest similarity to dreaming within a large database of psychoactive substance reports. *Front Neurosci* 12:7. <https://doi.org/10.3389/fnins.2018.00007>
- Schenberg EE (2021) Who is blind in psychedelic research? Letter to the editor regarding: blinding and expectancy confounds in psychedelic randomized controlled trials. *Expert Rev Clin Pharmacol* 14(10):1317–1319. <https://doi.org/10.1080/17512433.2021.1951473>
- Schwartz HA, Sap M, Kern ML, Eichstaedt JC, Kapelner A, Agrawal M, Blanco E, Dziurzynski L, Park G, Stillwell D, Kosinski M, Seligman ME, Ungar LH (2016) Predicting individual well-being through the language of social media. *Pac Symp Biocomput* 21:516–527. <https://www.ncbi.nlm.nih.gov/pubmed/26776214>
- Szigeti B, Kartner L, Blemings A, Rosas F, Feilding A, Nutt DJ, Carhart-Harris RL, Erritzoe D (2021) Self-blinding citizen science to explore psychedelic microdosing. *Elife*. <https://doi.org/10.7554/eLife.62878>
- Tagliazucchi E (2022) Natural language as a window into the subjective effects and neurochemistry of psychedelic drugs. *Frontiers in Pharmacology* 900
- van Elk M, Fejer G, Lempe P, Prochazkova L, Kuchar M, Hajkova K, Marschall J (2021) Effects of psilocybin microdosing on awe and aesthetic experiences: a preregistered field and lab-based study. *Psychopharmacology*. <https://doi.org/10.1007/s00213-021-05857-0>
- Wießner I, Falchi M, Palhano-Fontes F, Oliveira Maia L, Feilding A, Ribeiro S, Mota N, Barros de Araujo D, Tófoli LF (2021) Low-dose LSD and the stream of thought: increased discontinuity of mind, deep thoughts and abstract flow. *Psychopharmacology* 39:1721–1733
- Yanakieva S, Polychroni N, Family N, Williams LTJ, Luke DP, Terhune DB (2019) The effects of microdose LSD on time perception: a randomised, double-blind, placebo-controlled trial. *Psychopharmacology* 236(4):1159–1170. <https://doi.org/10.1007/s00213-018-5119-x>
- Zamberlan F, Sanz C, Martinez Vivot R, Pallavicini C, Erowid F, Erowid E, Tagliazucchi E (2018) The varieties of the psychedelic experience: a preliminary study of the association between the reported subjective effects and the binding affinity profiles of substituted phenethylamines and tryptamines. *Front Integr Neurosci* 12:54. <https://doi.org/10.3389/fnint.2018.00054>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Camila Sanz¹ · Federico Cavanna^{1,2} · Stephanie Muller¹ · Laura de la Fuente^{1,3} · Federico Zamberlan^{1,4} · Matías Palmucci¹ · Lucie Janeckova⁵ · Martin Kuchar^{5,6} · Facundo Carrillo⁷ · Adolfo M. García^{8,9,10,11,12} · Carla Pallavicini^{1,2} · Enzo Tagliazucchi^{1,13} 

¹ Departamento de Física, Universidad de Buenos Aires and Instituto de Física de Buenos Aires (IFIBA – CONICET), Pabellón I, Ciudad Universitaria (1428), CABA, Buenos Aires, Argentina

² Fundación Para La Lucha Contra Las Enfermedades Neurológicas de La Infancia (FLENI), Montañeses 2325, C1428 CABA Buenos Aires, Argentina

³ Institute of Cognitive and Translational Neuroscience, INECO Foundation, Favaloro University, Buenos Aires, Argentina

⁴ Department of Cognitive Science and Artificial Intelligence, Tilburg University, Tilburg, the Netherlands

⁵ Forensic Laboratory of Biologically Active Substances, Department of Chemistry of Natural Compounds, University of Chemistry and Technology Prague, Prague, Czech Republic

⁶ Department of Experimental Neurobiology, National Institute of Mental Health, Klecany, Czech Republic

⁷ Applied Artificial Intelligence Lab (ICC-CONICET), Pabellón I, Ciudad Universitaria (1428), CABA, Buenos Aires, Argentina

⁸ Cognitive Neuroscience Center (1644), Universidad de San Andrés, Buenos Aires, Argentina

- ⁹ National Scientific and Technical Research Council (1425), Buenos Aires, Argentina
- ¹⁰ Global Brain Health Institute (94143), University of California-San Francisco, San Francisco, CA, USA
- ¹¹ Global Brain Health Institute (94143), Trinity College Dublin (D02), Dublin, Ireland
- ¹² Departamento de Lingüística Y Literatura, Facultad de Humanidades (9160000), Universidad de Santiago de Chile, Santiago, Chile
- ¹³ Latin American Brain Health Institute (BrainLat), Universidad Adolfo Ibañez, Santiago, Chile