**Annotation guidelines for PsyNamic**

V8: July 05, 2024 (minor adjustments to grammar and some clarifications to items in red)

**Authors**: Vera Lara Bernhard, David Brüschweiler, Pia Härvelid, Julia Bugajska, Bernard Hild, Helena Aicher, Milan Scheidegger, Benjamin V. Ineichen

# **1) Inclusion criteria for papers**

First step before starting the annotation: check whether the paper meets our eligibility criteria. All these points need to be fulfilled in order to be included for annotation:

We will include all clinical studies which report on the use of psychedelic substances (see below) in human populations, this includes healthy volunteers (including recreational drug users) as well as individuals with medical conditions/disorders.

We will **exclude**:

-Clinical studies not using classical psychedelic substances including THC, CBD, cocaine, heroin, and similar substances.

-Animal studies and non-systematic reviews (including comments, editorials or similar article types). Of note, studies including animals and humans will be included.

-Corrigenda.

-Exclude studies which only focus on purely operative pain and/or anesthesia aspects of ketamine, e.g., studies finding optimal ketamin doses for hip replacement surgery (except the study also assesses postoperative depression).

Exclude eligible abstract based on these criteria (no annotation required).

# **2) General rules on how to annotate:**

1) When annotating tokens:

A) **Always consider the context**: What is the population (P), intervention (I), control (C), and outcomes (O) of the text of interest. The models we train DO “learn”/”understand” context.

B) **Only annotate what is relevant for the study of interest**, e.g., when annotating species tokens: “In this study, we used LSD. To test the robustness of the findings, these experiments should be repeated using psilocybin in future studies.” Only LSD should be annotated (psilocybin is not part of this study).

2) There are 2 types of annotation tasks:

-**Text annotation**: The entire text is classified according to pre-specified labels.

-**Token annotation**: the relevant word or word span is annotated. Annotate all occurrences.

General rule: NA is mostly reserved for systematic reviews and unknown for non-systematic review study types for which the information is unknown

# **3) Additional rules on how to annotate:**

3) Include incorrect spelling/grammar if pertinent.

4) If possible, the labelled word string should not be a combination of terms with and without brackets. E.g. "oral appliance (OA) device" should result in two labelled words "oral appliance" and "OA". Another example: “Bone marrow stem cells (BMSC) transplantation”, no annotation of transplantation.

5) For token annotation: No annotation for cases in which an incorrect token was to be annotated (e.g., the Prodigy interface not allowing a separation), e.g., “we studied this therapy in multiple sclerosis(MS)”: Multiple sclerosis cannot be annotated without including the bracket (due to the Prodigy interface), i.e., do NOT annotate this. For sentence annotation: Annotate the whole sentence even if you have to include the first word of the following sentence due to the Prodigy interface, e.g., “We randomly assigned rats to willed movement training (WM):5 rats per …” Annotate everything including the 5 (since Prodigy will not be able to exclude the “5” from the annotation).

6) Do NOT annotate punctuation at the end of the sentence (see above).

# **4) Data annotation**

**Text level annotation**

|  |  |
| --- | --- |
| **Parameter** | **Definition and comments** |
| **Study characteristics** | |
| Study type | **Def**: The methodological classification of research based on its design, purpose, and approach.  **Selection**: select only one (the best fitting)  Levels:  •**Randomized controlled trial (RCT):** A study where individuals are randomly assigned to one of two or more groups, with one receiving a psychedelic substance and the others receiving a placebo or standard treatment, to compare outcomes directly. I.e., the study needs to check to attributes: 1) being randomized and 2) being controlled (e.g., with a placebo or an active control).  •**Cohort study:** An observational study where a defined group of people (the cohort) is followed over time to study outcomes related to an exposure to psychedelic substance(s), comparing results within the group or with a different group not exposed to the substance. In most cases, it is explicitly mentioned that this is a cohort study.  •**Real-world study:** Examines the relationship between variables (e.g., the use of a psychedelic substance and mental health outcomes) at a single point across a real-world population and setting. This includes real-world studies on recreational drug use, survey-based studies e.g., on MDMA use among healthy adults or studies assessing emergency department visits of people ingesting MDMA. Studies recruiting recreational drug users for additional interventions or experimental procedures such as taking blood or conducting EEG do NOT belong to this category.  •**Study protocol:** A protocol of a clinical trial outlining methodology of a study still to be conducted. Annotate as if the study has been conducted.  •**Systematic review and/or meta-analysis:** A comprehensive summary of all the available primary research on a specific psychedelic-related research question. Should mention the term systematic review and/or meta-analysis in the title/abstract. Meta-analyses only refer to published studies, i.e., pooling of different RCTs would be “Other”.  •**Qualitative study:** Explores subjective experiences, perceptions, and the impact of psychedelic substances through interviews, focus groups, or content analysis. Also includes studies focusing on religion/spirituality related to psychedelic substances. **IGNORE LABEL**  •**Case report:** A detailed report of the symptoms, signs, diagnosis, treatment, and follow-up of one individual patient/healthy participant.  •**Case series:** A collection of case reports involving patients who were given similar interventions.  •**Other**: A study not belonging to one of the above categories or not classifiable based on the provided information. A study pooling different RCTs together belongs to this category. |
| Study purpose | **Def: T**he overall purpose of the study.  **Selection:** select zero or more, i.e., a study can have more than 1 purpose such as assessing efficacy and safety.  Levels:  •Efficacy endpoints: This item should be checked if the study measures efficacy of a psychedelic substance to treat a disease (or, in rare circumstances, to improve mental health in a healthy population), e.g., whsurveyether a psychedelic substance improves depression or anxiety.  •Safety endpoints: This item should be checked if the study discusses potential adverse events/side effects OF A PSYCHEDELIC SUBSTANCE (Do NOT annotate adverse events not necessarily linked to the psychedelic, e.g., suicide in MDMA polydrug users should not be annotated as adverse events). It is not relevant whether adverse events are reported or not, e.g., a study reporting “the adverse events were mild and transient” or “no adverse events were noted” should both be checked. This also includes neurotoxicity studies like studies looking at memory impairment after long-term MDMA use. Also includes studies assessing emergency department visits of individuals consuming e.g., MDMA, or AEs related to physiological functions (e.g., increase in body temperature or blood pressure).  •Pharmacokinetics: Examines how the body absorbs, distributes, metabolizes, and excretes the psychedelic substance. This item should be checked if body substance levels are measured on at least one time point.  •Mechanism of action: The study assesses potential mechanisms of action of the psychedelic substance (mostly on a molecular/mechanistic level). This includes neuroimaging studies (e.g., fMRI or PET scans to see how psychedelics affect brain activity and connectivity), neuropharmacology studies (e.g., how psychedelics interact with specific neurotransmitter systems, such as serotonin 5-HT2A receptors), genetic expression studies (e.g., assessing how psychedelics affect gene expression within the brain), and electrophysiological studies (e.g., measuring changes in the electrical activity of neurons in response to psychedelic substances).  •Disease model study: The study uses the psychedelic substance to induce a model psychosis (e.g., ketamine is commonly used for that).  •Demographic study (“epidemiological study”): e.g., a study assessing MDMA use among a real-world population. This also includes studies about adverse events of psychedelic substances in a recreational/real-world setting.  Additional information:  •Generous study purpose checking (e.g., study about preventing ecstasy in general population also check safety)  •Note, rarely, a study has no such study purpose, e.g., a study which asks psychologists about their experience with using psychedelics to treat their patients.  •Reduced verbal fluency in ecstasy users would be safety outcome. |
| Study control | **Def**: The control type of the study.  **Selection**: select one or more.  Levels:  •Placebo  •Active control (e.g., midazolam as a control for ketamine)  •Pre-post: A study design where measurements are taken on the same subjects at two points in time: before (pre) and after (post) a psychedelic intervention. **IGNORE LABEL**  •Other control (e.g., non-users, pre-post, cross-over)  •Non-controlled  •Not applicable (e.g., for systematic reviews)  Additional information:  •E.g., Midazolam is an active control  •Control: refers to the control condition (e.g., if ketamine is tested in anxious and non-anxious patients without control non-controlled) |
| Data type | **Def**: Whether the data are collected longitudinally or cross-sectionally.  **Selection**: select only one  Levels:  • Longitudinal short: Outcomes are measures during at least two different time points > 24 hours apart but the total follow-up time is < 3 months.  • Longitudinal long: Outcomes are measures during at least two different time points > 24 hours apart and the total follow-up time is ≥ 3 months.  • Cross-sectional: Outcomes are measured during one single time point (or more than one time point within ≤ 24 hours).  • Unknown: Data type is unknown  • Not applicable (e.g., for systematic reviews or for studies pooling a variety of other studies together without providing information about data type)  Additional information:  •If no information is provided about study length (or only very indirectly like cross-over study) select unknown |
| Data collection | **Def**: The approach how the data are collected.  **Selection**: select only one  Levels:  • Retrospective: Data has been collected retrospectively (case series and reports are always retrospective)  • Prospective: Data has been gathered prospectively (always select prospective for study protocols)  • Unknown: Data collection unknown.  • Not applicable (e.g., for systematic reviews)  Additional information:  •Important is when the actual OUTCOME was measured, i.e., an abstract describing a patient receiving ketamine over 4 weeks, but the outcome was only measured twice within 2 hours = cross-sectional.  •Retrospectively or prospectively needs to be explicitly stated. Otherwise unknown.  •Always select “prospective” for study protocols. |
| Number of participants | **Def**: The number of participants studied. Include all study participants (including the ones not receiving psychedelics).  **Selection**: select only one.  Levels:  • 1-20  • 21-40  • 41-60  • 61-80  • 81-100  • 100-199  • 200-499  • 500-999  • ≥1000  • Unknown  • Not applicable (e.g., for most systematic reviews)  Additional information:  •In case a systematic review reports the number of individuals in the abstract, select the according choice.  •If pregnant women are the study participants, select double the number of participant (mother and child)  •If a study protocol mentions a range of participants likely being included, select NA. |
| Sex of participants | **Def**: The sex of participants studied.  **Selection**: select one or more.  Levels:  • Male  • Female  • Both sexes (We assume both sexes if nothing reported and ≥ 5 included subjects)  • Unknown (check this if neither the number of participants nor their sex is reported)  • Not applicable (e.g., for most systematic reviews)  Additional information:  •If pregnant women are the study participants, select both sexes if >5 pregnant women.  •In case a systematic review specifies the sex of included subjects and/or the number of individuals as > 5, select the respective checkbox. |
| Age of participants | **Def**: The age of participants studied.  **Selection**: select only one.  Levels:  • Pediatric (< 18 years old)  • Adult (≥18 years) (We assume adults if nothing reported)  • Unknown  • Not applicable (e.g., for most systematic reviews)  Additional information:  •In case a systematic review specifies the age of included subjects, select the respective checkbox.  •In case a study only mentions “adolescents” without data on actual age check pediatric |
| **Substance(s)** | |
| Substances | **Def**: The psychedelic substance(s) studied.  **Selection**: select one or more.  Levels (substances:  • Ketamine (not further specified)  • S-Ketamine  • R-Ketamine  • MDMA (Ecstasy)  • LSD  • Psilocybin (or psilocin)  • Psychedelic mushrooms (as organic compound, including psychedelic truffles).  • Ayahuasca (DMT plus MAO-inhibitor, also 5-MeO-DMT plus MAO-inhibitor)  • DMT / N,N-DMT (sometimes referred to as “tryptamines”)  • 5-MeO-DMT (sometimes referred to as “tryptamines”)  • Mescaline  • Ibogaine / iboga (includes noribogaine)  • Salvinorin A (sometimes referred to as Salvia or Salvia Divinorum)  • Combination therapy: One or more of these substances combined with any other non-psychedelic substance or any other type of therapy (e.g., electroconvulsive therapy or psychotherapy) and in a therapeutic context (i.e., do NOT annotate combination therapy if an MDMA polydrug abuser is also smoking cannabis)  • Analogue (check this and the respective substance if it is not the exact substance listed above but closely related [claimed by the study authors], e.g, MDE is an analogue of MDMA check “MDMA” and “analogue”)  • Unknown (e.g., when just mentioning psychedelics without further specification)  Additional information:  •For example: “In addition to standard treatment” check “combination therapy” (if we assume that standard treatment includes psychotherapy)  •If examples are provided for psychedelics, e.g., “such as psilocybin and lsd”, but the study includes other non-reported substances, check all mentioned substances plus “unknown”.  •Racemic ketamine select ketamine (except abstract particularly mentions R- and/or S-Ketamin) |
| Application form | **Def:** The form of substance application.  **Selection:** select one or more.  Levels:  • Oral  • Nasal  • Intravenous  • Smoking (includes inhalation)  • Subcutaneous  • Other (does include sublingual)  • Unknown  • Not applicable (e.g., for systematic reviews) |
| Regimen | **Def:** What therapeutic regimen did the study use?  **Selection:** select zero or more (a study could have both single and multiple applications potentially)  Levels:  • Microdosing (as defined by the study authors)  • Application at multiple times points (e.g., LSD given on at least 2 time points or ketamine weekly for 3 months; or regular mdma-users for recreational drug use)  • Single dose (substance is only applied once, also counts as single dose if substance is applied more than once over up to 1 hour)  • Unknown (e.g., in a cross-sectional study on psychedelic use among the general population and it is not clear how many times the participants consumed psychedelic substances)  Additional information:  •This item only applies to individuals actually receiving the substance, i.e., if healthy controls are mentioned in an abstract not receiving a substance (in addition to some other treatment group), do NOT select unknown but only select the regimen for the treatment group receiving the substance.  •If you select microdosing, do NOT select “multiple applications” in addition.  • Could potentially be annotated for systematic reviews.  • PER substance (i.e. once DMT and once ketamine in the same study is single dose).  • For systematic reviews: select “Regimen: unknown”, except a systematic review assess a very specific regimen (e.g., microdosing)  • Very low dose of the psychedelic substance used as placebo does not count as one application. |
| Setting | **Def**: The setting in which the psychedelic substance is applied.  **Selection**: select one or more.  Levels:  •Clinical: the substance is applied in a clinical setting, e.g., a room of a clinic (this will be the large majority).  •Naturalistic: the substance is applied in a naturalistic setting, e.g., within a traditional ayahuasca ceremony in nature.  •Party setting  •Other setting  •Unknown: Select only if it is unclear from the context. E.g., select this for studies which look at memory functions from MDMA users in which the setting is not described.  •Not applicable (e.g., for most systematic reviews, except e.g., systematic reviews which look at naturalistic settings of psychedelic substances)  Additional information  •Select clinical if no particular setting is mentioned and the study has likely been conducted in a clinical setting. Select unknown only if it is unclear from the context in which setting the substance was applied. |
| Substance naivety | **Def:** Whether the study participants are naïve to psychedelic substances. A study could involve both naïve and experienced users.  **Selection:** select one or more.  Levels:  • Substance-naïve participants (study participants have not used psychedelic substances in the past)  • Substance-non-naïve participants (study participants did use psychedelic substances in the past)  • Unknown (this should be checked for studies where it is not clear whether study participants have consumed such substances, e.g., in surveys on drug use. However, if healthy controls are part of the study cohort for which we do not know whether they have consumed substances, do NOT check this box)  • Not applicable (e.g., for systematic reviews)  Additional information  •The substance naivety is only rarely reported on abstract level. Make a judgment call based on the abstract content to check the option(s), e.g., a study which assess the effect of ketamine on depressive patients, check unknown  •Substance-naivety includes controls  •Non-naïve relates to any psychedelic substance |
| **Clinical measures** | |
| Condition | **Def**: The overarching conditions of the study participants under investigation.  **Selection**: select zero or more.  Levels:  • Psychiatric condition (e.g., PTSD, depression, anxiety disorder, addiction such as smoking addiction. Check this and the corresponding disorder below e.g, in case of depression: check both “psychiatric condition” and “depression”).  • Depression  • Anxiety  • Post-traumatic stress disorder (PTSD)  • Alcoholism  • Other addictions (includes addictions like smoking and others)  • Anorexia  • Alzheimer’s disease  • Non-Alzheimer dementia  • Substance abuse (e.g., when the study participants are polydrug users/abusers of certain substances and the effect of one or more psychedelics is studied on an outcome DIFFERENT to treating the addiction to drugs)  • (Chronic) pain (e.g., cluster headache)  • Palliative setting (end of life)  • Recreational drug use  • Healthy participants (will also include larger population-based studies for e.g., recreational drug use where we can assume that most participants do not have major medical conditions. Is commonly checked together with recreational drug use)  Additional information:  •Of note, this only applies to the condition being treated, i.e., if a study assesses whether chronic recreational MDMA use is causing depression, do NOT annotate depression as condition.  •If no information about participants is provided, we can reasonably assume that they are healthy |
| Outcomes | **Def:** Does the study measure/use/report certain specified outcomes?  **Selection**: select zero or more.  Levels:  •The ability to drive a car  •Suicide (incl. suicidal ideation or related questionnaires)  •Functional MRI (fMRI)  •MRI (e.g., arterial spin labelling. Note: check this and the fMRI box if fMRI has been used)  •PET, SPECT (or similar imaging approaches)  •Physiological functions (e.g., heart rate, blood pressure, body temperature, pupillary diameter, BMI)  •Electroencephalography (EEG/MEG), Of note: includes MEG  •Mental functions (mental functions, e.g., ability to recognize faces, form language, mindfulness, memory, measures of mood or depression, impulsivity)  •Psychedelic experience, e.g., mystical experience questionnaire or visual scale for psychedelic experience.  •Soluble biomarker (e.g., proteins in blood or CSF or also in non-liquid body compartments such as hair. This also includes the substance itself measured in some body compartment, e.g., in pharmacokinetic studies. Includes genotyping [requires blood sampling])  •Surveys (e.g., questionnaires about how commonly psychedelic substances are consumed in a recreational setting). This does NOT include acquisition of patient data in the clinic. It mostly refers online surveys.  •Interviews (person-to-person)  •Neurotoxicity (measuring biological neurotoxicity more or less directly, e.g., MR-spectroscopy to measure toxicity metabolites or cell assays of patients. This does NOT include indirect measures of neurotoxicity such as adverse events of MDMA use)  •Emergency department visits (e.g., studies assessing reasons for emergency department visits of users of psychedelic substances). |
| Clinical trial phase | **Def:** stages in which scientists conduct experiments with a therapeutic intervention to obtain sufficient evidence for a process considered effective as a medical treatment (phase 0-4), i.e., considering efficacy and/or safety of the therapeutic intervention. NOTE: Medical devices and behavioural therapies (like meditation) do not have phases.  **Selection:** select zero or more (could be a phase 1/2 for example).  Levels:  •Phase 1  •Phase 2  •Phase 3  •Phase 4  •Unknown: fulfilling above definition but phase not explicitly (or implicitly) mentioned.  •Not applicable: Not fulfilling the above definition, e.g., an fMRI study assessing the changes upon lsd or a survey among MDMA users or a systematic review.  Additional information:  •Of note: the exact phase is rarely being reported, i.e., most studies will either be unknown or not applicable. |
| Study conclusions | **Def:** The overarching finding of the study. Only applies to studies testing a therapeutic impact of psychedelic substances on medical conditions (e.g., psychiatric conditions, headache).  **Selection:** select only one.  Levels:  • Positive, having a beneficial therapeutic impact (e.g., if a study reports: “lsd might decrease suicidal ideation in depression”)  • Negative (e.g., if a study reports: “ketamine worsened neuropathic pain in patients”)  • Neutral (e.g., if a study reports: “we did not observe an effect of MDMA on avoidance behaviour”)  • Mixed: both negative and positive conclusion  • Unknown: no conclusion  • Not applicable (e.g., for a study assessing fMRI changes of lsd or for a study protocol)  Additional information:  •Information is mostly found in the last 2 – 3 sentences of the abstract. |

**Token level annotation**

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Definition and comments** | **Examples** |
| Application area | **Def**: The condition or disease of interest studied under the impact of the psychedelic substance in the present paper.  Additional information:  •Our working definition for the **Application** is any "state labeled as diseases by virtue of consensus on prevalent sociocultural and medical values". It has to have "clearly identifiable diagnostic features and disease progression, and response to specific treatment."  •As a general guideline, annotated entities should be conditions that have an ICD-10/11 code.  •**Annotate only in title and/or the introductory part of the abstract, i.e., when the problem or objective of the study is discussed (including abbreviations)**. Do NOT annotate outcomes, such as “the Hamilton Depression Scale” should not be annotated. Similarly, “antisuicidal response” should not be annotated. Or: “among participants, depression was reduced”, do NOT annotate depression.  •Do NOT annotate antidepressant.  •Do not annotate over brackets: “…or bipolar disorder (BD) type 1 and type 2”, only bipolar disorder and BD should be annotated as two entities.  •Do NOT annotate very generic concepts like “psychiatric condition”  •Annotate sweeping statements like “…with major depressive and bipolar disorder” as one entity.  •Annotate the reasonably necessary string, e.g., “active suicidal ideation”, active suicidal ideation should be annotated. However, adjectives could be annotated, “Depressed patients”, annotate Depressed but not patients. “depressive symptoms” should both be annotated (also symptoms of depression). In “refractory anxiety disorder”, annotate the entire string.  •Further defining characteristics should be included: Acute/Chronic; Active/Inactive; Mild/Moderate/Severe; End Stage/Early Stage; Drug-resistant; Total/Partial; Intermittent/Relapsing and others. Similarly, "Post-stroke" should be annotated instead of only "stroke" because it refers to the phase after the acute stroke.  •Annotate abbreviations. Abbreviations should be annotated separately. For instance,“Huntington disease (HD)”should be separated into two annotations: “Huntington disease” and “HD.  •Of note, in “Managing dissociative symptoms following esketamine”, dissociative symptoms are NOT the treatment target but the target of the study and should NOT be annotated (context is relevant).  •Cigarette addiction is a disorder but not smoking cessation.  •Can be empty, i.e., no annotation per abstract.  •Include .^ in the title for highlighting |  |
| Dosage | **Def:** The applied dosage(s) of the psychedelic substance(s) in the present study.  •Only annotate for studies testing not more than 1 psychedelic substance (e.g., do not annotate in a study testing lsd and psilocybin, BUT annotate in a study testing psilocybin against midazolam)  •Do NOT annotate doses of other substances or of the psychedelic substance in other studies (e.g, mentioned by referencing other studies)  •Do not annotate concentrations of solutions but dosages, i.e., do NOT annotate 500mg/10 ml, but annotate 2mg/kg.  •Annotate the number and the corresponding unit.  •Annotate all reported doses for the respective psychedelic substance.  •Annotate only for single substance studies (i.e., do not annotate in abstracts which report the dosage of two or more psychedelic substances).  •Annotate the number including the unit.  •Annotate also the term “between” if pertinent, similarly “or” and “and” should be annotated. Do not annotate “per device”.  •Annotate also relative dosages (e.g., 1 mg/kg), do not annotate the terms “bw” or “body weight” or similar if pertinent  •Do NOT annotate doses measured in e.g., plasma (like 4 mg/ml)  •Can be empty, i.e, no annotation per abstract. | Lsd was used in this study (100-150 mg).  In our study, between 15 and 20 mg of psilocybin was used.  Here, 0.4 mg/kg bodyweight DMT was used.  The patient received 56 or 84 mg ketamime.  The patient received 40 mg or 120 mg MDMA |

Regular expressions for NCTS to identify linked registries.

**Interrater agreement Round 1**

A screenshot of a computer

Description automatically generated

|  |  |  |
| --- | --- | --- |
| **Measure** | **Priority** |  |
| Study type | 1 |  |
| Study purpose | 1 |  |
| Study control | 2 |  |
| Data type | 2 |  |
| Data collection | 2 |  |
| Number of participants | 1 |  |
| Sex of participants | 1 |  |
| Age of participants | 1 |  |
| Substances | 1 |  |
| Application form | 1 |  |
| Regimen | 2 |  |
| Setting | 2 |  |
| Substance Naivety | 2 |  |
| Condition | 1 |  |
| Outcomes | 1 |  |
| Clinical Trial phase | 3 |  |
| Study conclusion | 3 |  |