

Czech Technical University in Prague

Faculty of Electrical Engineering

SAN Final Assignment

Timur Abragimovič, Jakub Benetin, Čestmír Vejmla



January 1, 2025

1 Introduction

Novel Psychoactive Substances (NPS) are emerging xenobiotics with often unknown pharmacological properties, toxicity, and interaction profiles. Despite tighter regulations, the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) notes a continuous rise in both the number of detected NPS and associated intoxications. As of 2024, more than 830 NPS have been monitored, many of which remain widely available in various drug markets due to ease of online distribution and incomplete legal control. These substances can range from potent synthetic opioids, such as carfentanil and etonitazene, to emerging cannabinoids (e.g., hexahydrocannabinol) and novel dissociatives. Their widespread misuse constitutes a global health threat—synthetic opioids alone contribute to more than 200 overdose deaths per day in the United States. Some NPS are so potent that they resemble chemical warfare agents in toxicity, raising concerns not only for users but also for public safety and healthcare providers.

Yet, compared to traditional illicit or prescription drugs, NPS often lack comprehensive preclinical studies that systematically map their acute and chronic effects on the central nervous system (CNS). Since many NPS are structurally and pharmacologically diverse, research is complicated by their heterogeneity: psychoactive effects, receptor-binding profiles, and time courses (pharmacokinetics) can vary drastically even within a single chemical class. The evolving nature of NPS further underscores the urgency to characterize these substances rigorously, as newly synthesized compounds may not fit neatly into known pharmacological categories.

Electroencephalography (EEG) is a powerful tool for assessing the neurophysiological effects of novel psychoactive substances (NPS) on the central nervous system (CNS). Research has demonstrated that different pharmacological classes of NPS elicit distinct patterns of spectral power alterations, reflecting their unique mechanisms of action. Psychedelics such as psilocybin, which primarily act on serotonin 5-HT_{2A} receptors, induce global decreases in EEG power, often resulting in desynchronization of brain activity and reduced connectivity across networks [Vejmola et al., 2021, Páleníček et al., 2013]. Dissociatives like ketamine can suppress alpha and theta activity while often amplifying certain higher-frequency components, reflecting NMDA receptor antagonism and associated cortical network dysregulation. Ketamine exhibits a dose-dependent effects, with lower doses enhancing delta and gamma power, while higher doses increase coherence across all frequency bands, particularly in delta, theta and high-beta [Páleníček et al., 2011, Shokry et al., 2019]. Stimulants, categorized by their dopamine transporter (DAT) or serotonin transporter (SERT) affinities, display a range of effects: DAT-affinitive stimulants such as amphetamine typically enhance middle-frequency power, especially in alpha and theta bands, while SERT-affinitive drugs like MDMA tend to reduce overall EEG activity [Zanettini et al., 2019], with time-dependent effects [Tao et al., 2015]. Opioids such as fentanyl significantly increase the amplitude of slow-frequency (delta- and theta-band) EEG activity [Greenwald and Roehrs, 2005]. Cannabinoids, particularly CB₁ agonists like THC, are generally associated with decreases in EEG power, with THC specifically affecting neural oscillations in the theta band [Skosnik et al., 2018, Buonamici et al., 1982]. These findings on classical drugs underscore the potential of EEG as a tool for identifying distinct neurophysiological signatures linked to different NPS, aiding in the development of biomarkers for both toxicity and therapeutic effects.

2 Dataset

2.1 Origin

The EEG recordings were obtained within the ongoing project of the Ministry of the Interior entitled "New psychoactive substances: forensic-toxicological research centre" (VK01010212).

2.2 Drug classes

All psychoactive substances were divided into eight different classes according to their pharmacology, i.e., according to the profile of the receptors to which they bind. Psychostimulants were further divided into three groups of substances, according to their DAT/SERT ratios, which significantly determine the quality of their effect, according to [Luethi and Liechti, 2020] and [Rudin et al., 2021]. Cannabinoids were then divided into two groups according to whether they activate CB₁ and CB₂ receptors (cannabinoids (agonist) group) or whether they inhibit their activity (cannabinoids (negative

allosteric modulator) group). The schematic of the drug classes and their respective representatives can be seen in the figure below (Figure 1).

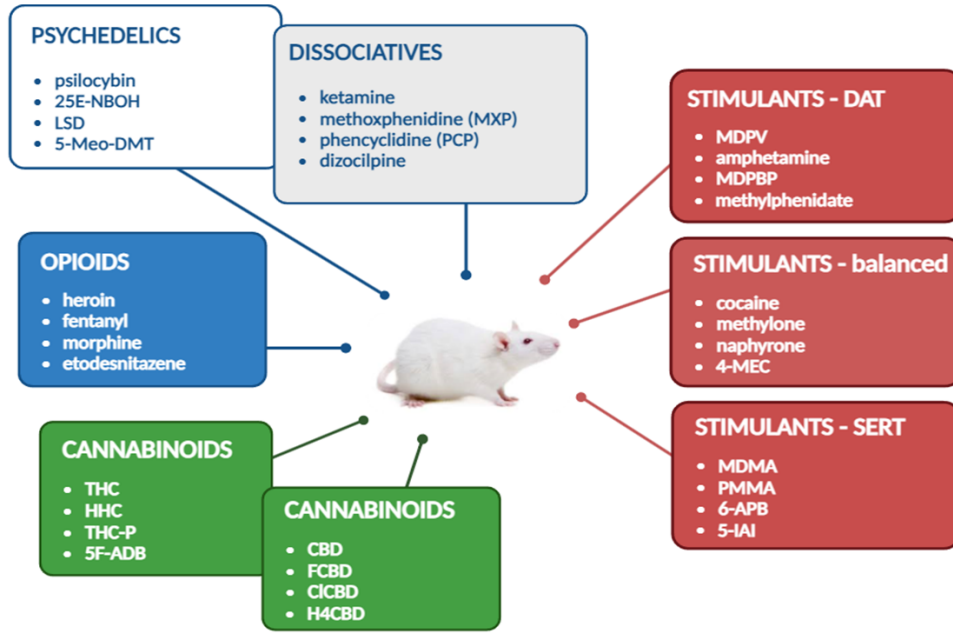


Figure 1: Drug classes and their respective representatives.

2.3 Study design

In addition to the eight different drugs, the rats were also injected with saline solution (VEH) and saline solution adulterated with ethanol (VEHET), which was necessary to dissolve some of the substances. In each run, one substance was selected for each group and injected in a cross-over design to \pm eighteen rats over ten weeks. Each rat was given each of the ten substances just once within ten weeks, in randomized order.

A total of four runs of experiments were recorded. In the first, third, and fourth runs, the same substances were applied but at three different doses. Doses were chosen based on the literature. In the second run, different substances were applied, dose-wisely corresponding to the dosing of drugs from the first run of the same category. A summary of the substances selected for the respective groups is shown in (Figure 2).

GROUP	RUN 1			GROUP	RUN 2		RUN 3 a 4		
	TREATMENT	dose mg/kg	file name		dose mg/kg	file name	TREATMENT	dose mg/kg	file name
VEH	SAL+deionized	x	x	SAL+deionized	x	x	SAL+deionized	x	x
VEHET	SAL+eth+Tween	x	x	SAL+eth+Tween	x	x	SAL+eth+Tween	x	x
psychedelics	psilocybin	5	psilocybin_H	25E-NBOH	5	25E_H	psilocybin	0,5/2	psilocybin_L/psilocybin_M
SERT	MDMA	5	MDMA_M	PMMA	20	PMMA_M	MDMA	2,5/10	MDMA_L/MDMA_H
BALANCED	cocaine	20	cocaine_H	MDMC	10	MDMC_M	cocaine	5/10	cocaine_L/cocaine_M
DAT	MDPV	2	MDPV_M	amphetamine	5	AMP_M	MDPV	1/4	MDPV_L/MDPV_H
dissociatives	MXP	20	MXP_M	ketamine	30	ketamine_M	MXP	10/40	MXP_L/MXP_H
opioids	heroin	0,25	heroin_M	fentanyl	20ug	fentanyl_M	heroin	0,05/2	heroin_L/heroin_H
CBR modulators	CBD	10	CBD_H	FCBD	10 H	FCBD_H	CBD	1/5	CBD_L/CBD_M
CBR agonists	HHC	10	HHC_H	THC	10 H	THC_H	HHC	1/5	HHC_L/HHC_M

Figure 2: Experimental design: the selected substances with their respective doses for 4 different runs of the experiment are shown. The abbreviations meanings are: L - low dose, M - medium dose, H - high dose. VEH - vehiculum = saline, VEHET - saline + ethanol

2.4 Data collection

Experiments were performed on adult male laboratory rats (Wistar strain). The animals were housed in pairs, the breeding cages were placed in a room with controlled temperature (22 ± 2 °C) and humidity (50-70 %), access to food pellets and water was ad libitum. Animals (n = 18 individuals per group) were recorded using an innovative EEG implant, which has a tetrahedral equidistant arrangement with 20 cortical electrodes (Figure 3). The electrodes were implanted during standard stereotactic surgery to optimally cover the entire brain.

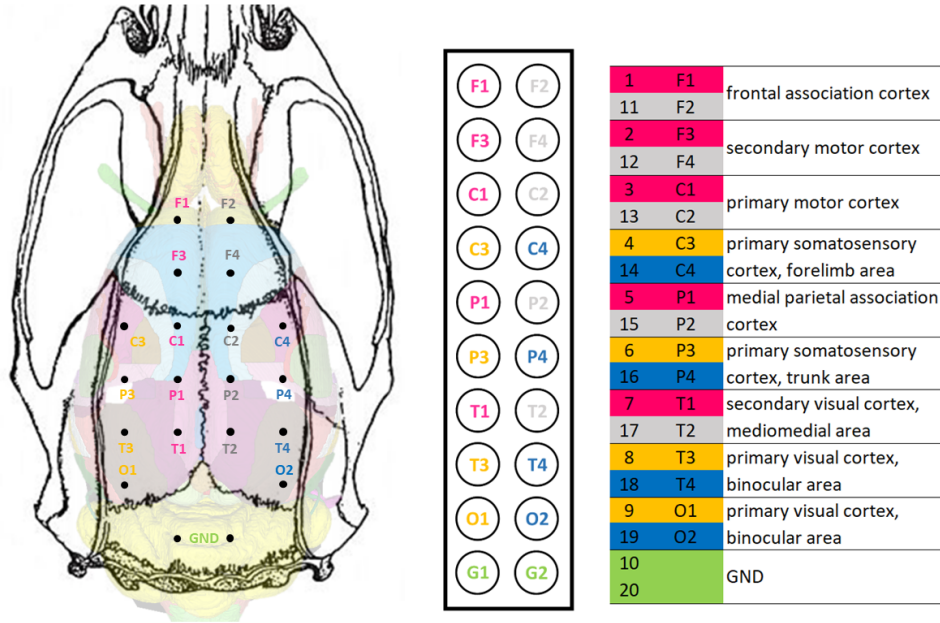


Figure 3: Electrode layout on the top of the rat skull (left). Electrode pinout - top view on a connector (middle). Electrode labels and numbering with the respective brain regions they overlay (right).

For EEG acquisition, rats were placed in a spacious cubicle the size of a standard open-field arena, i.e. 80x80 cm (Figure 4). The rats were allowed to move freely throughout the recording. After a 15-minute baseline recording, the rats were injected with one of the drugs and recording then continued for another 90 minutes.



Figure 4: Photos of the monitoring (upper) and recording (lower) room.

2.5 EEG data preprocessing and analysis

The preprocessing of EEG data was performed in BrainVision software by experts in a standard resting-state EEG pipeline: The EEG data were filtered (0.5-150 Hz), artifacts were rejected by semi-automatic approach, and if necessary, bad channels were interpolated using spherical splines. The continuous signal was then segmented into a baseline record (segment 0) and six 15-minute segments after treatment (segments 1-6). The average spectral power was then calculated for each 15-minute segment and frequency band: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-35 Hz) and gamma (35-100 Hz).

2.6 Data structure

Spectral outputs were exported as numpy matrices of the dimension: time segment (segment 0, segment 1, ..., segment 6) x EEG bands (delta, theta, ..., gamma) x channels (F1, F3, ..., O2), i.e. matrices of the dimension (7x5x18). The file names then reflect the rat identification number, treatment, dose, week of administration, and the arena in which the recording took place (arena 1 or 2). Thus, for

example, the file name "ID16_psilocybin_H_week7_ar2" refers to the recording of rat number sixteen, which received high-dose psilocybin in the seventh week of the experiment and was recorded in arena number 2.

3 Validation of Cross-over Design

3.1 Motivation and Hypotheses

The primary motivation of this analysis was to validate cross-over design per se by investigating whether baseline EEG spectral power in rats is influenced by external conditions and experimental design factors. Specifically, we aimed to assess (i) whether the baseline EEG power differs between two experimental arenas (Arena 1 vs. Arena 2), (ii) whether time (across ten consecutive weeks) plays a role in altering the EEG baseline, and (iii) whether there is a systematic effect of different experimental runs (Run 1, Run 2, Run 3, Run 4). Our initial hypotheses were that:

1. The baseline EEG power would not substantially differ between arenas under otherwise identical conditions.
2. There would be no large systematic differences across the four runs, assuming the protocols were consistently applied.
3. The baseline EEG might show minimal variation over ten weeks, reflecting stable physiological conditions in the animals.

3.2 Methods

3.2.1 Data Collection and Preprocessing

This report analyzed baseline segment records only. We constructed a single dataset containing the following columns:

- *rat_id*: unique identifier of each rat
- *arena*: categorical factor (Arena 1, Arena 2)
- *run*: categorical factor (Run 1, Run 2, Run 3, Run 4)
- *week*: numeric (1–10)
- *delta*, *theta*, *alpha*, *beta*, *gamma*: mean spectral power in each frequency band

We then used linear mixed-effects models (LMM) via the `lmer` function in `lmerTest`:

$$\text{Power}_{ijkl} = \beta_0 + \beta_{\text{arena}} \cdot \text{Arena}_i + \beta_{\text{run}} \cdot \text{Run}_j + \beta_{\text{week}} \cdot \text{Week}_k + (1|\text{rat_id}),$$

where *Power* (either delta, theta, alpha, beta, or gamma) was the response variable. Arena and run were treated as fixed categorical factors, week was treated as a numeric variable for a linear time trend, and *rat_id* was included as a random intercept to account for repeated measures within rats. We repeated this model for each of the five frequency bands. Post-hoc comparisons (Tukey) were performed using the `emmeans` package.

3.3 Summary of the Model

For each frequency band, the linear mixed model included the factors *arena*, *run*, and *week* plus a random intercept for individual rats. The overall structure provided:

- **Fixed effects:** Arena (1 or 2), Run (1–4), and Week (1–10)
- **Random effect:** (1|*rat_id*), allowing each rat to have its own baseline intercept
- **Outcome:** Speeds of spectral power changes or differences in baseline power attributed to each factor

3.4 Detailed Results by Frequency Band

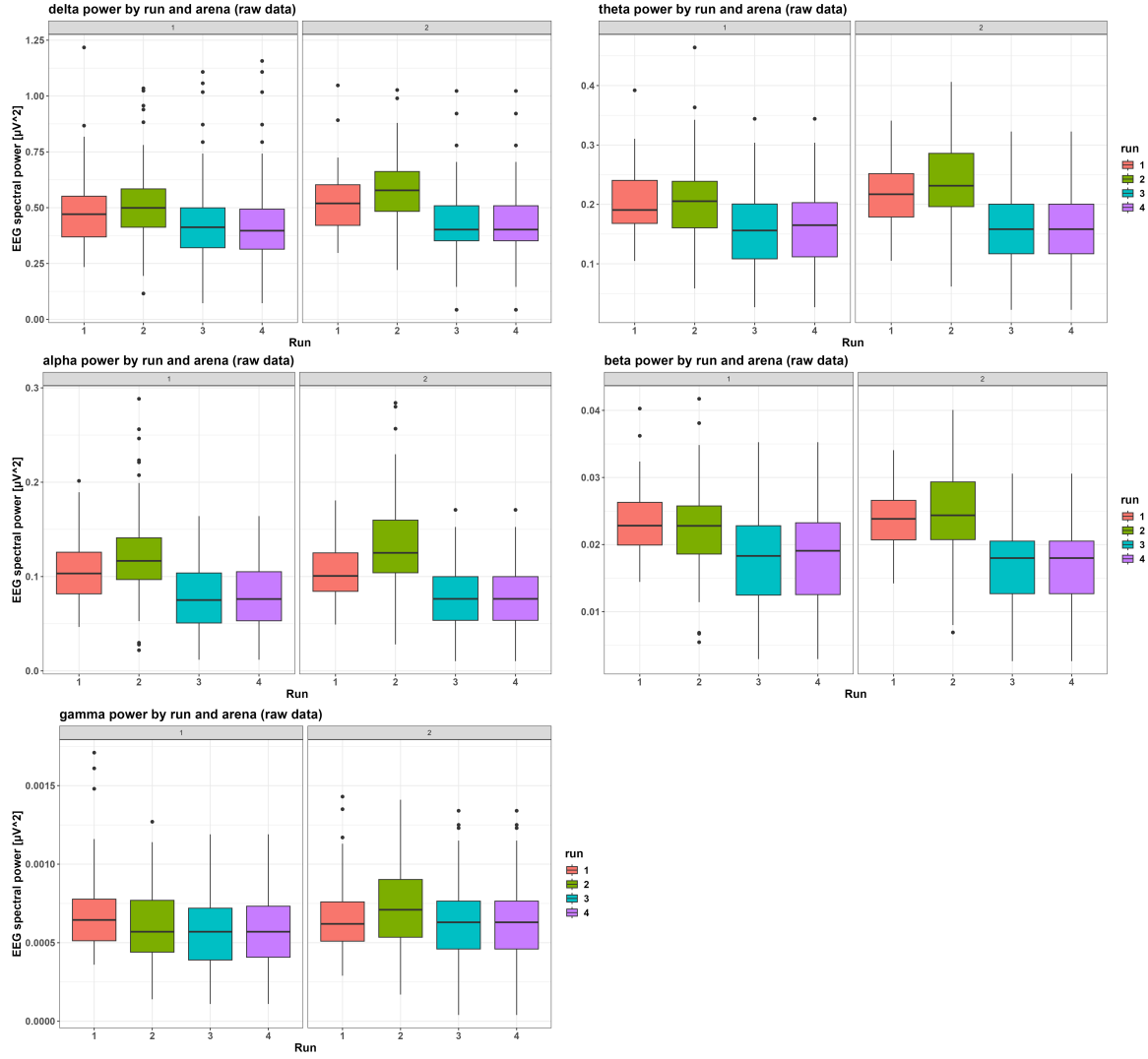


Figure 5: Boxplots of baseline EEG spectral power (μV^2) for each frequency band (δ , θ , α , β , γ) as a function of the experimental run. The x-axis shows the run (1–4), while the y-axis indicates the measured spectral power. Panels (facets) correspond to different arenas (Arena 1 or Arena 2). Each boxplot reflects the median, interquartile range, and potential outliers of the raw data for that run and arena, with fill colors distinguishing the run levels. This figure allows a direct comparison of how each frequency band’s baseline power distribution varies among the runs and between arenas.

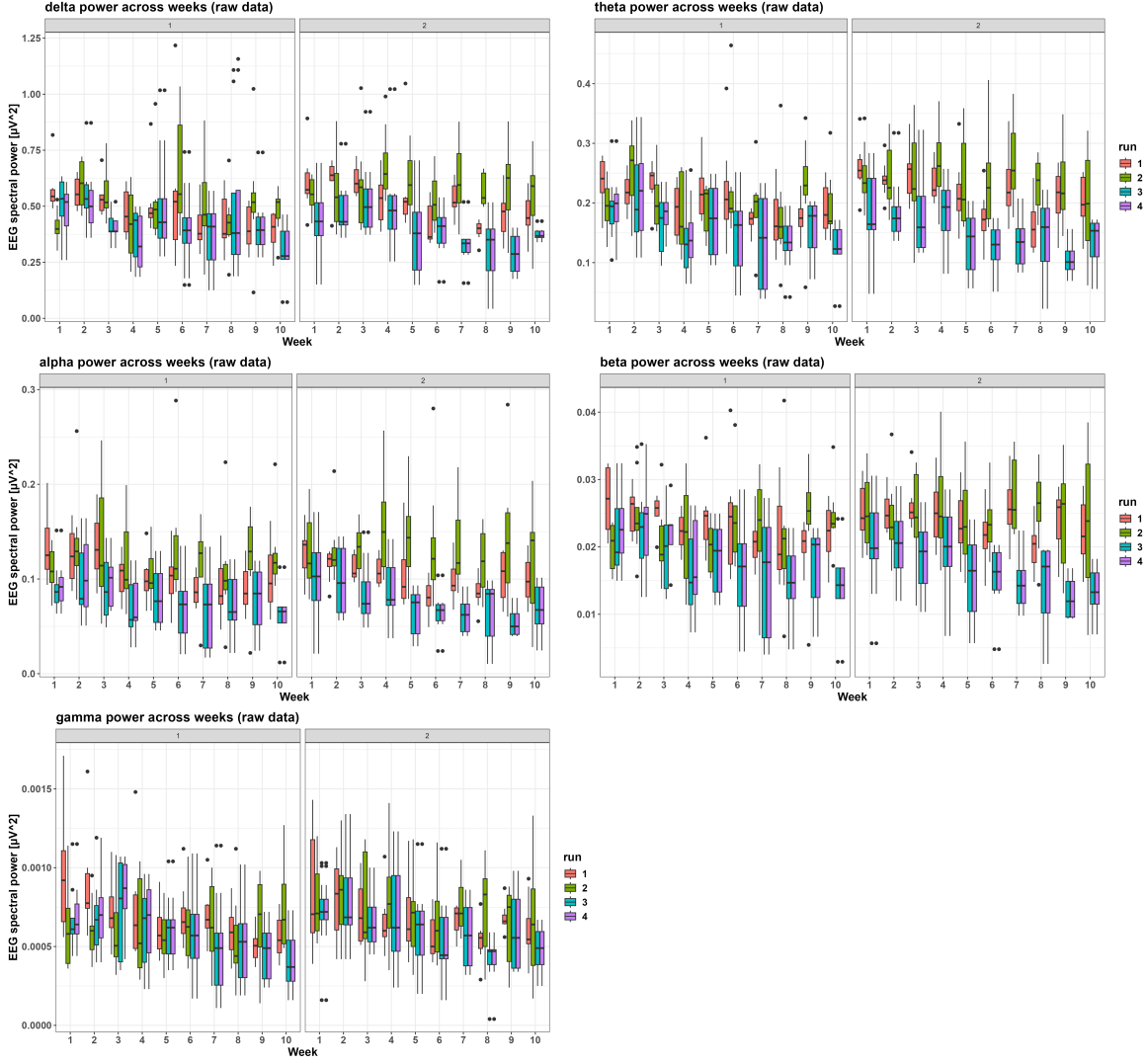


Figure 6: Boxplots of baseline EEG spectral power (μV^2) for each frequency band across the 10-week experiment. The x-axis shows the experimental week (treated as a factor), and the y-axis again represents spectral power in μV^2 . Each panel corresponds to a recording arena (Arena 1 or Arena 2), and each boxplot illustrates the distribution of raw baseline values for that specific week and run. Fill colors denote the run (1–4). By visualizing all weeks side by side, one can observe temporal trends or week-to-week variability in each frequency band’s baseline power, as well as differences between arenas and runs.

3.4.1 Delta Band

In the delta band, Arena 2 does not differ significantly from Arena 1 (Estimate = 0.0071, $p = 0.559$). Regarding the run factor, neither Run 2 (Estimate = 0.0360, $p = 0.282$) nor Arena 2 stand out, but Runs 3 (Estimate = -0.0870 , $p = 0.00768$) and 4 (Estimate = -0.0925 , $p = 0.00556$) both show significantly lower baseline values compared to Run 1. A highly significant negative effect of week is present (Estimate = -0.0147 , $p = 1.53e-12$), indicating a decreasing trend in delta power as the experiment progresses. Overall, delta power steadily declines over time, with Runs 3 and 4 showing lower levels than Run 1, while Run 2 is not distinctly different, and Arena 2 does not substantially deviate from Arena 1.

3.4.2 Theta Band

For the theta band, the effect of Arena 2 approaches significance (Estimate = 0.0067, $p = 0.0818$), but does not decisively differ from Arena 1. Run 2 (Estimate = 0.0104, $p = 0.5242$) also does not diverge substantially from Run 1, yet Runs 3 (Estimate = -0.0628 , $p = 0.000101$) and 4 (Estimate = -0.0606 , $p = 0.000231$) are clearly lower than Run 1. A strong negative week trend persists here as well (Estimate = -0.00716 , $p < 2e-16$), implying a reduction in theta power over time. In short, theta displays a consistent decrease across the weeks, with Runs 3 and 4 notably below Run 1, and Arena 2 only marginally higher than Arena 1.

3.4.3 Alpha Band

In the alpha band, Arena 2 (Estimate = 0.0016, $p = 0.469$) shows no meaningful departure from Arena 1. Run 2 is slightly above Run 1 (Estimate = 0.0201, $p = 0.0430$), whereas Runs 3 (Estimate = -0.0325 , $p = 0.000715$) and 4 (Estimate = -0.0320 , $p = 0.00110$) are significantly lower. Week exhibits a strongly negative pattern (Estimate = -0.00392 , $p < 2e-16$), consistent with an ongoing decline in alpha power over time. Thus, alpha diminishes each week, while Run 2 sits modestly higher, and Runs 3 and 4 tend to be lower.

3.4.4 Beta Band

Within the beta band, neither Arena 2 (Estimate = 0.000106, $p = 0.763$) nor Run 2 (Estimate = -0.00036 , $p = 0.837$) differ notably from their respective references, but Run 3 (Estimate = -0.00735 , $p = 2.75e-05$) and Run 4 (Estimate = -0.00713 , $p = 6.42e-05$) are clearly reduced compared to Run 1. The negative week effect (Estimate = -0.000653 , $p < 2e-16$) remains significant, reflecting a steady decrease in beta power over time. In summary, baseline beta levels also drop across the weeks, with Runs 3 and 4 clearly below Run 1, and Run 2 or Arena 2 showing no special deviation.

3.4.5 Gamma Band

Lastly, in the gamma band, Arena 2 appears slightly elevated (Estimate = $2.99e-05$, $p = 0.0441$) relative to Arena 1. Run 2 (Estimate = $-4.22e-05$, $p = 0.5442$) does not differ greatly, but Run 3 (Estimate = $-1.34e-04$, $p = 0.0431$) and borderline Run 4 (Estimate = $-1.31e-04$, $p = 0.0542$) both exhibit a modest decrease from Run 1. Once again, the strong negative week coefficient (Estimate = $-3.225e-05$, $p < 2e-16$) underscores a declining trend in gamma power over the course of the experiment. In essence, gamma follows the common downward trajectory, with a mild elevation in Arena 2 and slight dips in Runs 3 and 4 compared to Run 1.

3.5 Overall Conclusions

A significant *negative trend* across weeks emerged in all frequency bands ($p < 2e-12$ up to $p < 2e-16$). Runs 3 and 4 frequently exhibit lower baseline spectral power compared to Runs 1 and 2. In the alpha band, Run 2 still shows a slight increase, whereas in other bands it remains non-significant. For the most part, **arena** is not a strong factor. However, in gamma ($p = 0.044$) and to some extent theta ($p \approx 0.082$), Arena 2 appears marginally higher. Overall, the difference in baselines between Arena 1 and Arena 2 remains small. **Residual Diagnostics:** Residual plots indicate no major heteroskedasticity. The QQ-plot is fairly acceptable, although there is an occasional outlier (e.g., in the delta band, max scaled residual of about 5.27).

3.6 Discussion

For most bands, **arena** remains a minor factor. In certain cases (gamma, and borderline for theta), Arena 2 yields marginally higher baseline power than Arena 1, but the numeric difference is modest. The subtle differences identified between arenas are likely to be due to hardware, cables used, or possible interference specific to a particular arena.

All models show a strong time dependence (**week**), suggesting that baseline EEG consistently diminishes across the whole frequency spectrum from roughly Week 1 to Week 10. From a biological standpoint, these findings confirm that EEG recordings are not fully stable over many weeks, and

that experimental runs may differ in ways not immediately apparent. This effect likely reflects rat adaptation, minor implant changes, or aging. If this were an adaptation of the rat to the recording environment, reflected, for example, by the rat exploring less and tending to decrease in vigilance, we would expect discrete changes across frequency bands, such as a decrease in the theta band, reflecting reduced locomotion, and an increase in the delta band, reflecting a tendency to fall asleep [Hansen et al., 2019]. Given the consistent broadband decrease observed, we would hypothesize that the aging of the implant and the so-called host-probe interaction, a biological reaction of the tissue in which glial scar formation occurs and tends to isolate the electrode and thus likely increase the impedance [Polikov et al., 2005], are primarily responsible for the observed effect.

Regarding **run** differences, Runs 3 and 4 systematically produce lower spectral power than Runs 1 (and partially Run 2). This underscores that multi-run designs can introduce subtle variations—technical, environmental (different times of the year), or even rat cohort-related.

All of these findings suggest that in a cross-over design with subsequent comparisons of substances, it is necessary to normalize the EEG by the baseline recording to avoid possible bias.

4 Modeling Time-Dependent EEG Power Across Substance Doses

In this section, we model the time-dependent responses of various substances and compare them to the saline (VEH) recordings. We used all available data while disregarding the influence of arena and week when it was recorded. For each substance, dose, and frequency band, we used Linear Mixed Models (LMM). We allowed for random intercepts and slopes. This approach allowed us to test for differences in the intercept change (increase or decrease of EEG power) and slope (change in power dependent on time) relative to the intercept and slope of saline recordings.

Although linear models may not fully capture the nonlinear nature of EEG power changes (see Figure 7), we opted for this simpler, more interpretable approach due to the large dataset, accepting the tradeoff in model fit. These models will not generalize well for time segments outside our range (90 minutes total, when each rat was under the influence of a particular substance). However, if the trend is a straightforward linear decrease or increase in EEG power, it should capture it, since almost all substances used produce effects lasting up to several hours.

4.1 Data Overview and Preprocessing

In total, we had over 600 recordings, each containing 7 time segments (15-minute intervals), 5 frequency bands (delta to gamma), and 18 electrodes. The data was first averaged across the electrodes, reducing it to a 7x5 format per recording. Initial visualizations showed that the effects of different substances were largely global, affecting various brain regions similarly (e.g., increase in power, decrease, or no effect).

Time segment 0, which represents the baseline (pre-substance) recording, was subtracted from the subsequent time segments. This results in all values in time segment 0 being set to zero, and this segment was then discarded, leaving us with 6 time segments and 5 frequency bands for analysis.

The lowest number of recordings was for the medium dose of MDPV ($n = 11$), while the highest number was for the high dose of cocaine ($n = 35$). Reference saline recordings (VEH) were the most numerous ($n = 66$).

4.2 Fitted models

- **time**: numeric (1-6, time segment 1 is from 15-30 min, 2 from 30-45 min...)
- **dose**: Categorical factor (0, L, M, H)
- **rat id**: Unique identifier for each rat

We fitted a Linear Mixed-Effects Model (LMM) via the `mixedlm` function in `statsmodels`:

$$\text{Power}_{ijkl} = \beta_0 + \beta_{\text{time}} \cdot \text{time}_i + \beta_{\text{dose}} \cdot \text{dose}_j + \beta_{\text{time:dose}} \cdot (\text{time}_i \cdot \text{dose}_j) + (1 \mid \text{rat id}),$$

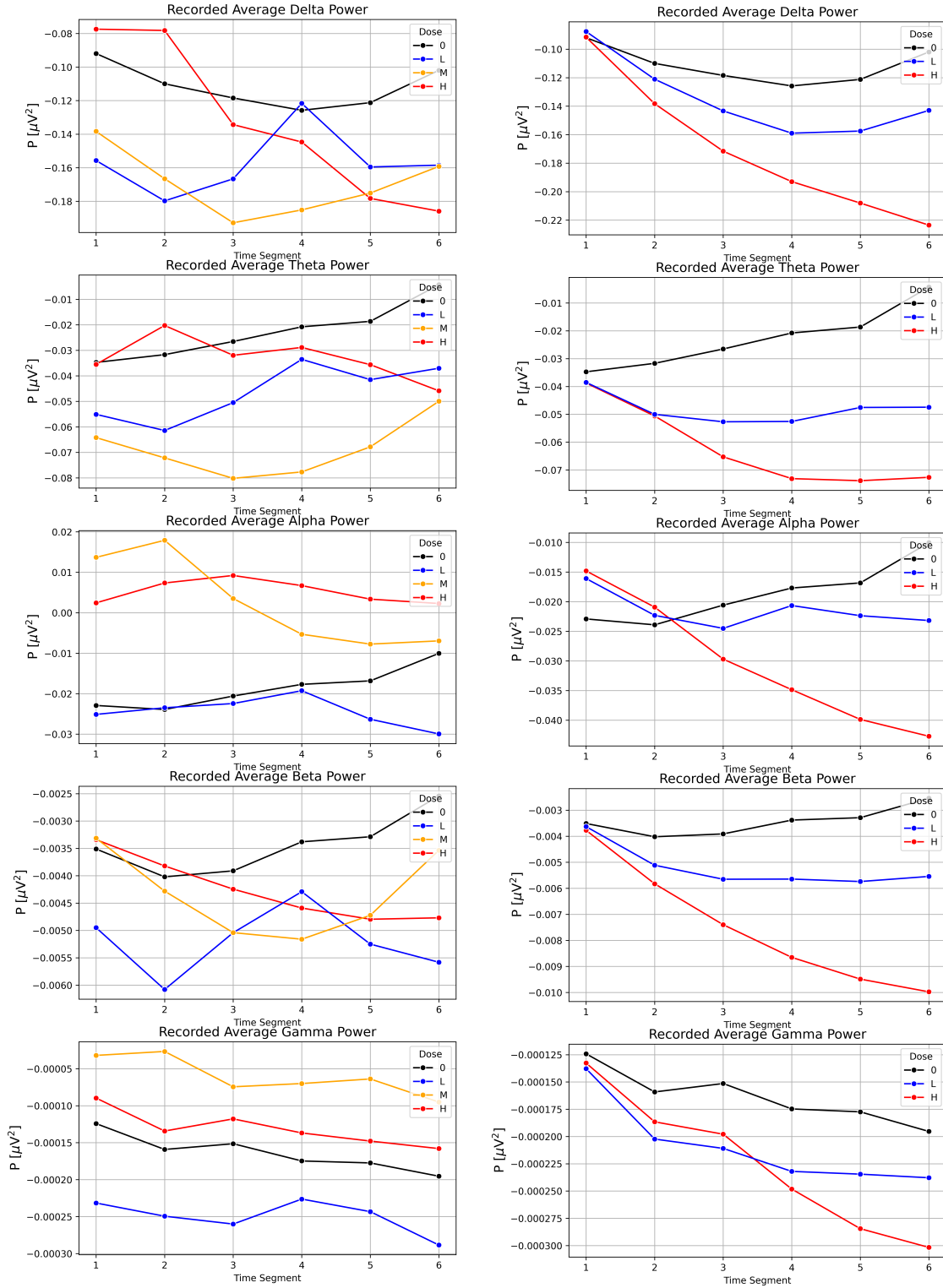


Figure 7: Left: Average power for different frequency bands after heroin administration. Right: Same, but for HHC. Dose 0 (black) represents average recorded power values after saline injection (VEH), averaged across all rats. Naturally, these values are identical in both plots. L (blue), M (orange), and H (red) correspond to low, medium, and high doses of the respective substance. This coloring scheme is used throughout this section. These plots illustrate the variability in the data and the non-linear trends observed across conditions, emphasizing the challenges of modeling EEG responses. Nevertheless, some frequency bands for certain dosages show a decent linear trend in time.

where **Power** was the response variable, **time** was treated as a continuous independent variable, **dose** as a categorical factor, and **rat id** was included as a random factor. The interaction term (**time** \times **dose**) was included to examine how the effect of time differs across doses. This was done for all frequency bands.

4.3 Results

Refer to Figure 9 as that contains all the results described in this section. Positive intercept values correspond with increased spectral power after administering the substance, negative intercepts to decrease. Similarly for slopes, positive values signify gradual increase of EEG power in time and negative ones a gradual decrease.

4.3.1 Saline Intercepts and Slopes

These are represented in Figure 9 as horizontal dashed lines. In all cases the intercept for saline recording is negative, meaning that after the initial 15 minute baseline recording, the power in each frequency band decreased slightly. Injecting saline should have no effect on spectral power. One possible explanation is that during the baseline recording, the rat was recently moved into a new environment for recording and thus might be more active at the start, busy with exploring the new enclosure. During such behavior, the brain would be more active. Slopes for saline recordings are close to zero (see Figure 8), as is to be expected.

4.3.2 Substance Intercepts and Slopes

For the intercepts of different substances at different dosages, CBD, cocaine, heroin, HHC, MDMA, FCBD and THC share similar negative values as the saline recordings across all frequency bands. Psilocybin and especially 25ENBOH usually have more negative intercepts across bands than saline recording. Among substances that cause an initial increase in spectral power are MDPV, AMP, ketamine and MDMC.

Most substances presented have quite unremarkable slopes that are similar to the saline recording for delta and theta bands. Notable exceptions are ketamine, MDMC and THC, which show a negative slope in for all bands. For ketamine and MDMC, this is expected, since these substances cause an initial increase in power and as the effects wear off, power gradually returns to normal. THC on the other hand starts with a negative intercept and from then on the EEG power across bands keeps decreasing throughout the 90 minutes.

4.4 Conclusion

This section provides an informative overview of the time-dependent effects of various substances and dosages on EEG power across frequency bands. For some substances, namely ketamine, MDMC, and THC the linear models fit adequately. However, substances with more transient or oscillatory effects, such as cocaine, MDMA and PMMA, might require alternative modeling approaches to fully capture their dynamics.

Naturally, each intercept and slope calculated carries a p-value, but this was not included, as it is dependent on the number of recordings were available to us for each dose. For an overview of how EEG power evolves in time based on substance and dose, the plots themselves are more informative and going into more depth for each individual substance is beyond the scope of this work.

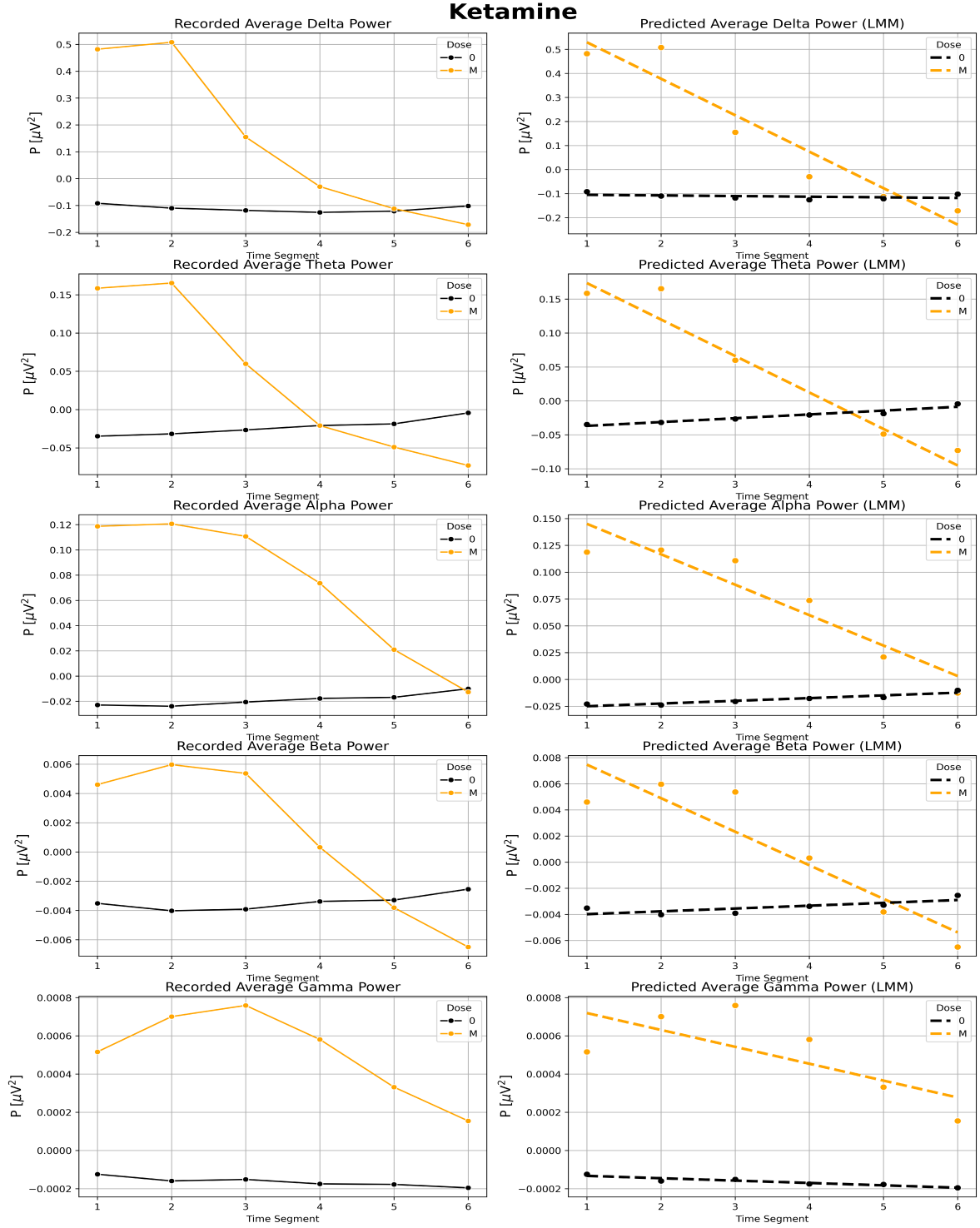


Figure 8: Showcase of ketamine, where a linear model might be appropriate. Left: Recorded average power values from all recordings at particular dosage for each frequency. Right: Fitted models. With respect to time, saline recordings are mostly constant, with only minimal slope. Ketamine on the other hand shows an initial increase in EEG power after administration and gradual decrease back to baseline with passing time across all frequency bands.

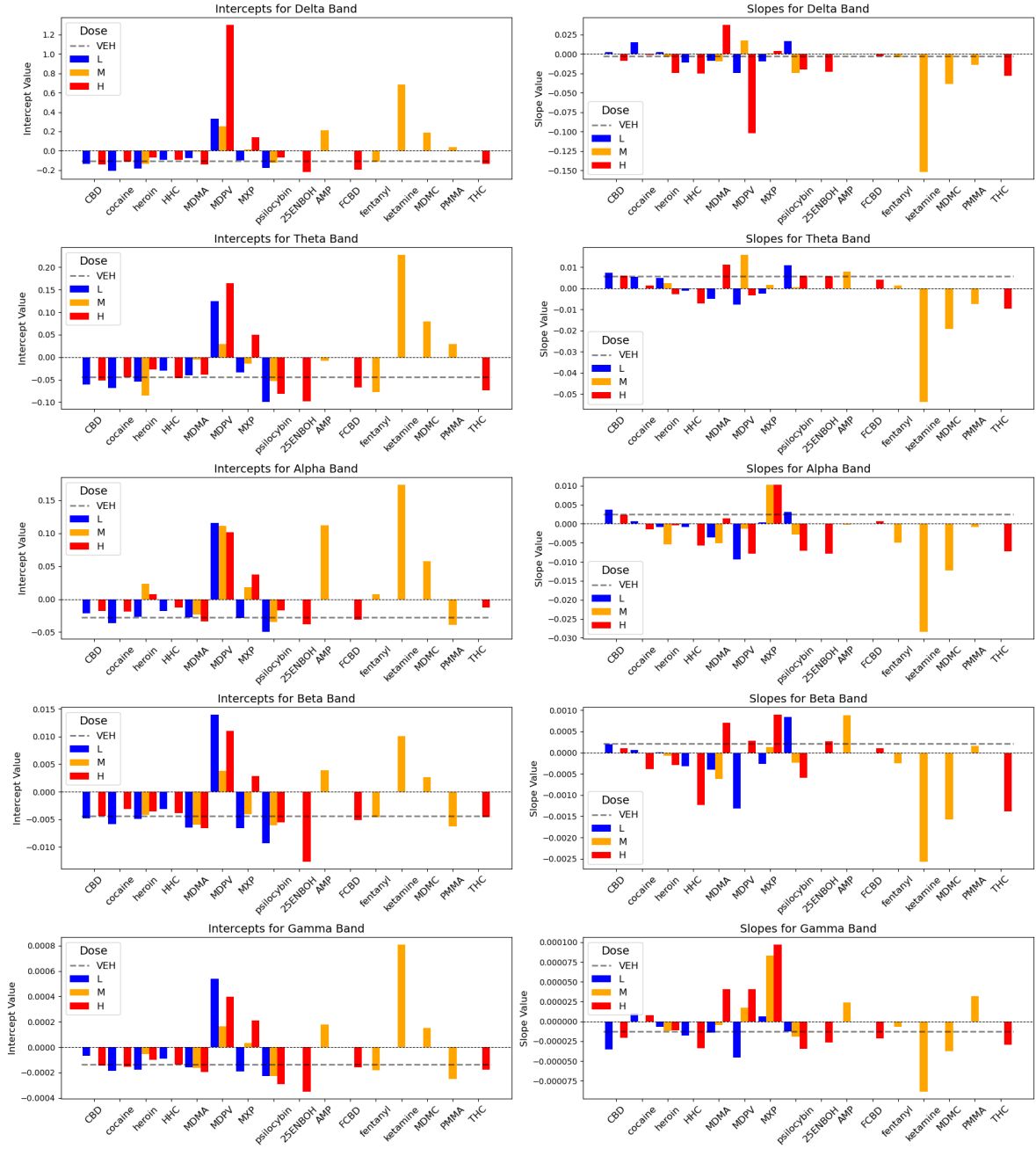


Figure 9: Intercepts and slopes for all different substances and dosages that were available. Horizontal lines in each plot correspond to the slope and intercept at particular frequency band of the saline recordings.

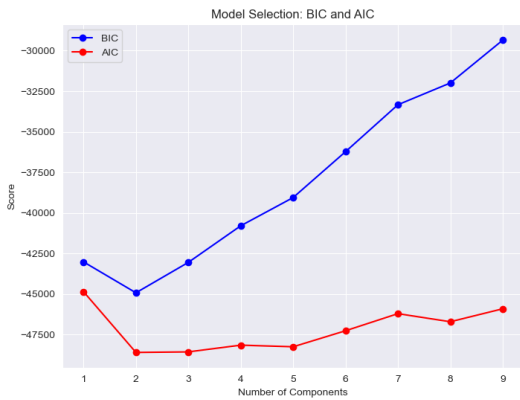
5 EEG Variations Across Pharmacological Groups

The section aims to investigate the differences in EEG activity across distinct pharmacological groups. For each rat, we calculated the difference in EEG activity between the baseline (before drug administration) and the six temporal segments after the drug was administered, across five spectral bands. This resulted in a 30-dimensional feature space for each measurement, with the data consisting of five spectral bands and six temporal segments per rat. We analysed whether EEG responses can reliably differentiate between these groups by administering proportionally scaled doses of various substances to rats within each group. For this analysis, we used data only from Run 1 and Run 2, excluding the "Stimulants - Balanced" group to ensure that doses were proportionally scaled for substance group determination. Runs 3 and 4 were deemed unsuitable for this purpose.

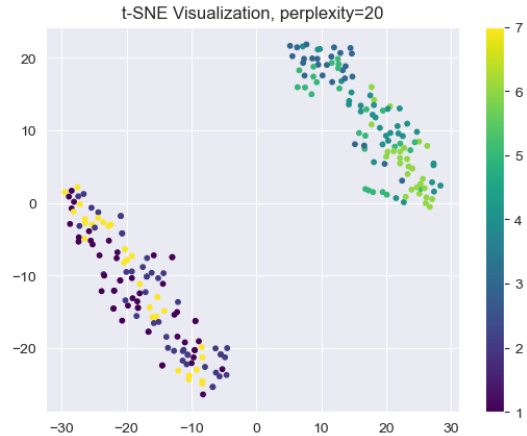
In this section, we focus on methods such as the Expectation-Maximization (EM) algorithm, Quadratic Discriminant Analysis (QDA), and Linear Discriminant Analysis (LDA). All these methods assume that each group follows a normal distribution. We hypothesize that our data primarily encode information in the mean structure rather than more complex separation patterns. Consequently, with this assumption and a limited dataset, we aim to achieve better results by reducing variance, even at the cost of potential increases in bias error due to partial assumption normality violation.

5.1 Unsupervised clustering using EM algorithm

Unsupervised clustering was applied to EEG data to investigate whether distinct pharmacological groups could be identified based solely on their EEG activity. We employed the Expectation-Maximization (EM) algorithm to cluster the data without prior knowledge of group labels. To determine the optimal number of components for the EM algorithm, we used the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), both of which indicated that the data could be best partitioned into two clusters (Figure 10a). The first cluster primarily included substances such as PMMA, ketamine, AMP, MDMA, MXP, heroin, fentanyl, and MDPV, which are associated with an increase in EEG power. In contrast, the second cluster included substances such as CBD, HHC, THC, FCBD, 25ENBOH, and psilocybin, which generally elicited a decrease in EEG power. While these clusters corresponded partially to pharmacological groupings, the method failed to resolve the full diversity of all seven groups. This result suggests that the clustering captures a broad distinction in EEG responses—amplification versus suppression of activity—but lacks the granularity to identify more subtle differences between groups. Additionally, the two clusters were visualized using t-SNE, providing a clear depiction of the separation in a lower-dimensional space (Figure 10b).



(a) AIC and BIC values for different numbers of components in the EM algorithm.



(b) t-SNE visualizes two clusters in data.

Figure 10: Figures showing AIC/BIC values and t-SNE visualization of clusters.

5.2 Supervised Discriminant Analysis

To further explore group separability, we applied supervised discriminant analysis methods, which leverage known group labels to classify the EEG data. For this analysis, the data was split into training and testing sets in a 70:30 ratio.

5.2.1 Quadratic Discriminant Analysis

Quadratic Discriminant Analysis (QDA) was tested as an initial supervised approach due to its ability to model non-linear boundaries between groups with different covariances. However, QDA suffered from overfitting in this context, as the dataset had a limited number of samples relative to the 30-dimensional feature space. This resulted in poor generalization performance, rendering QDA unsuitable for this analysis.

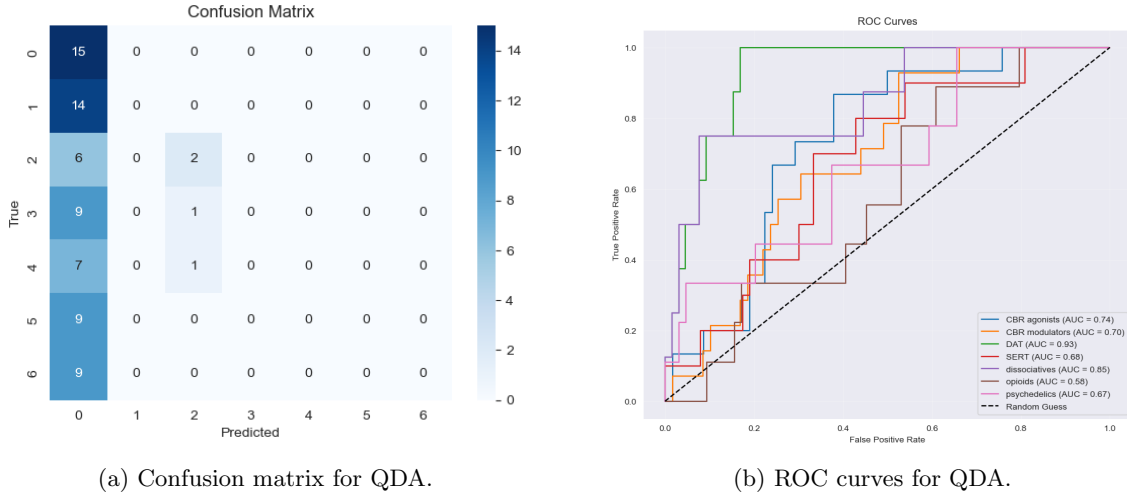


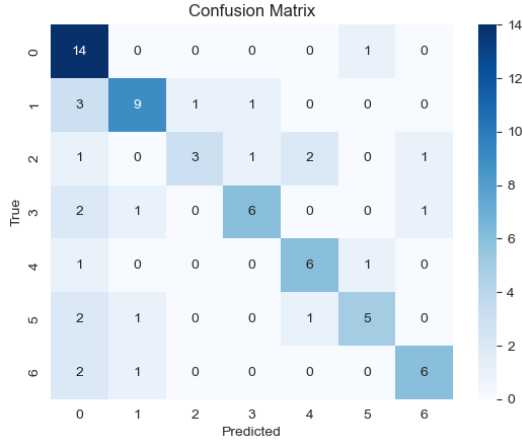
Figure 11: Quadratic Discriminant Analysis.

Encoded Label	Group	Precision	Recall	F1-Score
0	CBR agonists	0.22	1.00	0.36
1	CBR modulators	0.00	0.00	0.00
2	DAT	0.50	0.25	0.33
3	SERT	0.00	0.00	0.00
4	Dissociatives	0.00	0.00	0.00
5	Opioids	0.00	0.00	0.00
6	Psychedelics	0.00	0.00	0.00

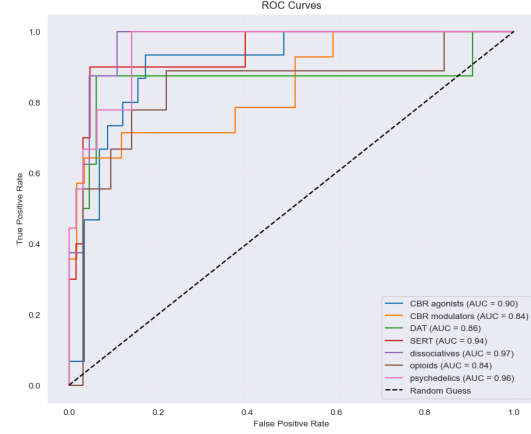
Table 1: Classification per each group for QDA.

5.2.2 Linear Discriminant Analysis with Ledoit-Wolf Shrinkage

Linear Discriminant Analysis (LDA) without shrinkage was initially used to distinguish between the seven pharmacological groups. While this method achieved reasonable performance, it showed limitations in accurately classifying some of the groups, especially in the high-dimensional feature space of EEG data. The results for LDA without shrinkage showed a high recall for the "CBR agonists" group (0.93), but with a lower precision (0.56), leading to an F1-score of 0.70. Other groups, such as "DAT," also showed suboptimal performance with lower recall and F1-scores. Figures 12a, 12b and table 2 illustrate the results for LDA without shrinkage.



(a) Confusion matrix for LDA with no shrinkage.



(b) ROC curves for LDA with no shrinkage.

Figure 12: Linear Discriminant Analysis with no shrinkage.

Encoded Label	Group	Precision	Recall	F1-Score
0	CBR agonists	0.56	0.93	0.70
1	CBR modulators	0.75	0.64	0.69
2	DAT	0.75	0.38	0.50
3	SERT	0.75	0.60	0.67
4	Dissociatives	0.67	0.75	0.71
5	Opioids	0.71	0.56	0.62
6	Psychedelics	0.75	0.67	0.71

Table 2: Classification per each group for LDA with no shrinkage.

To address these challenges and improve classification accuracy, Ledoit-Wolf shrinkage was applied to regularize the covariance matrix in LDA. The Ledoit-Wolf shrinkage estimator stabilizes the covariance matrix by shrinking the sample covariance toward a more robust target, often the identity matrix. This reduces the impact of noise and small sample sizes, leading to more reliable and robust estimates in high-dimensional spaces minimizing mean squared error. [Ledoit and Wolf, 2004]

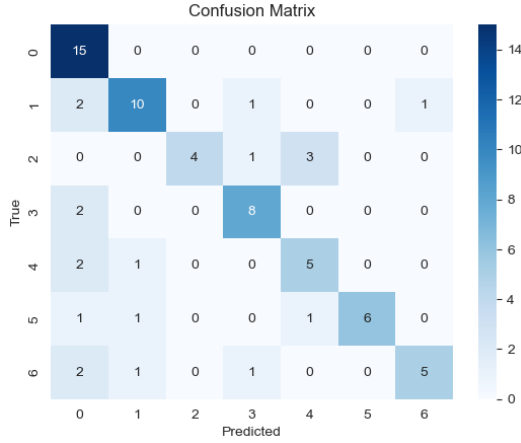
With Ledoit-Wolf shrinkage applied, LDA achieved significant improvements in classification performance. For example, the "CBR agonists" group achieved a higher precision of 0.62 with perfect recall (1.00), resulting in a higher F1-score of 0.77. Other groups also showed improved classification metrics, such as the "Opioids" group, which increased its precision to 1.00 and its F1-score to 0.80. Also Ledoit-Wolf shrinkage improves AUC results for classes (Figure 13b).

The following figures and table present the results for LDA with Ledoit-Wolf shrinkage:

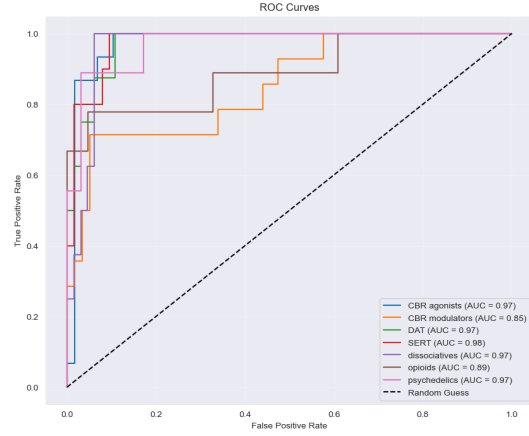
Encoded Label	Group	Precision	Recall	F1-Score
0	CBR agonists	0.62	1.00	0.77
1	CBR modulators	0.77	0.71	0.74
2	DAT	1.00	0.50	0.67
3	SERT	0.73	0.80	0.76
4	Dissociatives	0.56	0.62	0.59
5	Opioids	1.00	0.67	0.80
6	Psychedelics	0.83	0.56	0.67

Table 3: Classification per each group for LDA with Ledoit-Wolf Shrinkage.

In summary, while unsupervised clustering was limited in its ability to separate all groups, supervised discriminant analysis methods, particularly LDA with Ledoit-Wolf shrinkage, demonstrated strong performance in distinguishing between the EEG activity of the different pharmacological groups.



(a) Confusion matrix for LDA with Ledoit-Wolf Shrinkage.



(b) ROC curves for LDA with Ledoit-Wolf Shrinkage.

Figure 13: Linear Discriminant Analysis with Ledoit-Wolf Shrinkage.

However, for certain substances, precision and recall were notably lower, making this classifier unsuitable for real-world tasks where high reliability is required. Nevertheless, given the limited dataset and the exclusive use of linear methods, the results provide valuable insight into the feasibility of using EEG responses for group classification. The findings highlight that EEG responses to each pharmacological group are distinct and can be used to predict the group to which a substance belongs. This approach holds promise for future applications, such as the rapid identification of unknown or novel pharmacological substances requiring urgent treatment.

6 Summary

This study investigates the neurophysiological effects of classic and various novel psychoactive substances on the central nervous system using electroencephalography. The work is based on data collected from adult male Wistar rats injected with different NPS, including psychedelics, dissociatives, stimulants, opioids, and cannabinoids. The primary goal was to understand the spectral power changes in the EEG, which reflect the pharmacological effects of the substances.

The cross-over design of the experiment was validated by analyzing baseline EEG power across different factors, such as experimental arenas, weeks, and runs. The validation confirmed that while there were minimal differences in baseline EEG power between experimental arenas and across different weeks, there were significant temporal trends in spectral power across the 10-week period, suggesting that factors such as rat adaptation or changes in the experimental conditions might influence baseline measurements. These findings highlight the importance of normalizing EEG data to baseline recordings to avoid potential biases in subsequent analyses of substance-induced effects.

Key findings include distinct EEG signatures for each class of drugs. Analysis revealed temporal dynamics of EEG power changes post-substance administration. Linear mixed models (LMM) were employed to model time-dependent changes in spectral power, revealing that certain substances, such as ketamine and MDMC, caused initial increases in EEG power followed by a return to baseline levels. In contrast, THC consistently decreased power over time. Psychedelics such as psilocybin induced global decreases in EEG power, while dissociatives like ketamine caused dose-dependent increases in delta and gamma power. Stimulants showed a range of effects based on their receptor affinities, with dopamine-affinitive drugs enhancing low-frequency power, and serotonin-affinitive drugs reducing EEG activity. Opioids, particularly fentanyl, increased slow-frequency EEG power, while cannabinoids like THC and CBD generally decreased spectral power, with THC notably affecting theta oscillations.

Lastly, machine learning techniques, including unsupervised clustering and supervised discriminant analysis, were applied to explore the ability to differentiate pharmacological groups based on EEG patterns. The results indicated that while clustering successfully identified broad categories of substances based on EEG responses (i.e., power amplification versus suppression), supervised models,

particularly Linear Discriminant Analysis with Ledoit-Wolf shrinkage, performed well in classifying substances within specific pharmacological groups.

It is important to note that substances which typically increase EEG power, such as stimulants, are also known to enhance arousal and overall locomotor activity in rats. In the Linear Mixed Models (LMM) analysis, substances that increased EEG power were also those that likely influenced locomotion and arousal. This suggests that the increases in arousal and locomotion could be related to the observed EEG changes, as the recordings were obtained on freely moving rats.

In the machine learning analysis, substances that increased locomotion and arousal, such as stimulants, were grouped together in one cluster. In contrast, substances that typically suppress psychomotor activity, such as certain cannabinoids and psychedelics, formed a separate cluster. This clustering result supports the hypothesis that substances enhancing locomotion and arousal may share common neurophysiological effects that also impact EEG patterns. These findings suggest that drug-induced changes in behavioral states could be influencing the observed EEG power changes and that these factors may be interrelated.

These findings highlight the potential of EEG as a tool for mapping the neurophysiological effects of NPS and identifying biomarkers for toxicity and therapeutic effects, while also emphasizing the importance of considering behavioral factors such as locomotion and arousal in future analyses of substance-induced changes in brain activity.

7 Contribution Statement

We state that all three of us contributed to the work in roughly equal manner. Each one of us definitely put more hours into their part than the expected allotted time.

Timur Abragimovich worked on section 5 *EEG Variations Across Pharmacological Groups*.

Jakub Benetin worked on section 4 *Modeling Time-Dependent EEG Power Across Substance Doses*.

Čestmír Vejmla worked on sections 1 *Introduction*, 2 *Dataset*, 3 *Validation of Cross-over Design*, and 6 *Summary*. He also took part in collecting the data and provided us with this dataset.

References

- [Buonamici et al., 1982] Buonamici, M., Young, G., and Khazan, N. (1982). Effects of acute 9-thc administration on eeg and eeg power spectra in the rat. *Neuropharmacology*, 21:825–829.
- [Greenwald and Roehrs, 2005] Greenwald, M. K. and Roehrs, T. A. (2005). Mu-opioid self-administration vs passive administration in heroin abusers produces differential eeg activation. *Neuropsychopharmacology*, 30(1):212–221.
- [Hansen et al., 2019] Hansen, I. H., Agerskov, C., Arvastson, L., Bastlund, J. F., Sørensen, H. B. D., and Herrik, K. F. (2019). Pharmacoelectroencephalographic responses in the rat differ between active and inactive locomotor states. *European Journal of Neuroscience*, 50(2):1948–1971.
- [Ledoit and Wolf, 2004] Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.
- [Luethi and Liechti, 2020] Luethi, D. and Liechti, M. E. (2020). Designer drugs: mechanism of action and adverse effects. *Springer Berlin Heidelberg*, 94(4).
- [Polikov et al., 2005] Polikov, V. S., Tresco, P. A., and Reichert, W. M. (2005). Response of brain tissue to chronically implanted neural electrodes. *Journal of Neuroscience Methods*, 148(1):1–18.
- [Páleníček et al., 2011] Páleníček, T., Fujáková, M., Brunovský, M., et al. (2011). Electroencephalographic spectral and coherence analysis of ketamine in rats: Correlation with behavioral effects and pharmacokinetics. *Neuropsychobiology*, 63(4):202–218.
- [Páleníček et al., 2013] Páleníček, T., Fujáková, M., Brunovský, M., et al. (2013). Behavioral, neurochemical and pharmacoeeg profiles of the psychedelic drug 4-bromo-2,5-dimethoxyphenethylamine (2c-b) in rats. *Psychopharmacology*, 225(1):75–93.
- [Rudin et al., 2021] Rudin, D., Liechti, M. E., and Luethi, D. (2021). Molecular and clinical aspects of potential neurotoxicity induced by new psychoactive stimulants and psychedelics. *Experimental Neurology*, 343(June):113778.
- [Shokry et al., 2019] Shokry, I. M., Sinha, V., Silva, G. D., et al. (2019). Comparison of electroencephalogram (eeg) response to mdpv versus the hallucinogenic drugs mk-801 and ketamine in rats. *Experimental Neurology*, 313:26–36.
- [Skosnik et al., 2018] Skosnik, P. D., Hajós, M., Cortes-Briones, J. A., et al. (2018). Cannabinoid receptor-mediated disruption of sensory gating and neural oscillations: A translational study in rats and humans. *Neuropharmacology*, 135:412–423.
- [Tao et al., 2015] Tao, R., Shokry, I. M., Callanan, J. J., et al. (2015). Mechanisms and environmental factors that underlie the intensification of 3,4-methylenedioxymethamphetamine (mdma, ecstasy)-induced serotonin syndrome in rats. *Psychopharmacology*, 232(7):1245–1260.
- [Vejmola et al., 2021] Vejmola, , Tylš, F., Piorecká, V., et al. (2021). Psilocin, lsd, mescaline, and dob all induce broadband desynchronization of eeg and disconnection in rats with robust translational validity. *Translational Psychiatry*, 11.
- [Zanettini et al., 2019] Zanettini, C., Scaglione, A., Keighron, J. D., et al. (2019). Pharmacological classification of centrally acting drugs using eeg in freely moving rats: an old tool to identify new atypical dopamine uptake inhibitors. *Neuropharmacology*, 161.