The Influence of Brown and Pink Noise on Working Memory Performance in Students

Group 12 Final Report - 02455 Experiment in Cognitive Science

Weronika Ewa Straczek
Technical University of Denmark
Lyngby
s222754@student.dtu.dk

Shakir Maytham Shaker
Technical University of Denmark
Lyngby
s230553@student.dtu.dk

Siddhanta Gupta
Technical University of Denmark
Lyngby
s210217@student.dtu.dk

1 INTRODUCTION

In the realm of cognitive science, understanding how external factors impact cognitive abilities remains an important task. Among these factors, the influence of environmental stimuli on cognitive processes, particularly working memory, has been a subject of many studies throughout the years. Working memory, often described as the mental workspace responsible for temporarily holding and manipulating information essential for cognitive tasks, plays a crucial role in everyday functioning, including learning, problem-solving, and decision-making. [2] Assessing working memory can be done using a multitude of ways, one of which is by using the N-back task. As Conway et al. (2005) state, "the n-back task is arguably the current gold standard measure of WMC in the cognitive neuroscience literature" [4] (WMC meaning working memory capacity).

A significant aspect of this exploration lies in the examination of various environmental stimuli, such as noise, and their potential effects on cognitive performance. Noise, an omnipresent environmental factor, encompasses diverse types, each carrying its unique acoustic properties and effects on human cognition. In this study, specific attention is directed towards two types of noise: pink and brown noise. Brown (or otherwise called red) noise can be defined as having a power density that diminishes by 6 dB for each octave increase in frequency. In contrast, pink noise exhibits a frequency spectrum that appears linear when viewed on a logarithmic scale, with the spectral power density of pink noise decreasing by 3 dB per octave [9]. They will be both placed in contrast with a normal, silent environment.

1.1 Literature Review

Effects of background noise on cognitive processing have been widely studied throughout the years. Although it has been believed that noise acts as a distractor and has a negative effect on task performance, some studies show that under certain conditions noise could benefit task performance. [3]. As Awada et al. (2022) write, "The Moderate Brain Arousal (MBA) postulates that moderate levels of external auditory white noise introduce internal noise to the neural systems which allow undetected neural signals to pass the detection threshold thus leading to better cognitive performance." [1] However, most of the studies on effects of noise focus on individuals with Attention Deficit Hyperactivity Disorder (ADHD), as it has been already proven that auditory white noise has a positive effect on their cognitive processing [9].

Prior to designing the experiment, a review of already existing studies involving colored noise was conducted to contextualize the experiment idea within the existing body of knowledge. Among the studies identified as relevant to the research question were those by Guo et al. (2022) [6] and Lu et al. (2020) [9], both of which explored the impact of colored noise on human cognition.

Guo et al. (2022) conducted a between-subjects experiment involving 81 healthy college students, examining the effects of silence, white noise, and pink noise on attention. Participants of the study were instructed to execute a search task, which involved clicking a hexagon figure that was placed among pentagons, while their reaction times were recorded. Their findings indicated that a white noise environment led to significantly shorter reaction times among participants, while pink noise had a relatively minor impact compared to a silent environment. However, it's important to note that this study had certain limitations, such as its between-subjects design, which limited precise comparisons between different environments - there is no way of separating differences between noises from differences in individuals.

Lu et al. (2020) conducted a similar study, inspecting the effects of colored noise on work efficiency. Their study additionally introduced red (brown) noise, resulting in four levels of the independent variable - silence, white, pink and brown. Moreover, the experiment was a within-subjects study. Twenty-two participants were tasked with executing four different tests: Psychomotor Speed Test, Continuous Performance Test - Identical Pairs (CPT-IP), Trail Making Test (TMT) and Taiwans Odd-Even Number Sequencing Test (TOENST). Their experiment revealed that all types of colored noise (white, red, and pink) produced better results in these cognitive tests compared to a silent environment. Moreover, participants reported increased comfort when exposed to noise, as indicated by questionnaire responses.

The role of white noise on its own in influencing human cognition has been a subject of extensive research over the years. Among these studies, Helps et al. (2014) conducted research on the impact of white noise on children with varying levels of attentiveness. Their findings revealed that moderate levels of white noise could be beneficial for sub-attentive children but might have a disruptive effect on children with good levels of attention [7]. Baum and Chadda (2020) delved into the effects of white noise on the cognitive performance of healthy, neurotypical adults. Their study provided "preliminary evidence in support of the use of WN in healthy individuals to improve cognitive performance." [3]

In a separate study focusing on the effects of colored noise, Singh et al. (2023) explored how different types of colored noise influenced children's anxiety levels during a dentistry procedure. However, it's important to note that their study did not involve any task-solving assessments. [10]

1.2 Research Gap and Hypotheses

The review of existing literature reveals a significant gap in understanding the specific impact of colored noise, particularly pink and brown noise, on the working memory of students. While various studies have delved into colored noise's effects on attention and work efficiency, there is limited research into its influence on working memory. Furthermore, the prevailing body of research primarily involves white noise, lacking a focused exploration of colored noise's potential impact on human cognition. Lack of detailed research on this topic alongside the researchers' keen interest in different factors that can enhance working memory, were the motivation behind this study.

This paper addresses the research question: Does the exposure to colored noise, specifically pink and brown noise, have a significant impact on working memory in students? The proposed study adopts an experimental approach to examine the relationship between exposure to pink and brown noise and working memory in students while comparing it to the baseline silent environment. The hypotheses postulated for this study are based on the existing literature and are formulated as follows:

(1) Exposure to pink noise has a measurable effect on working memory in students.
(2) Exposure to brown noise has a measurable effect on working memory in students.

2 METHODS

2.1 Participants

The participants selected for this study were students above the age of 18, primarily between 20 and 26 years old. Gender wasn't a consideration in the recruitment process, resulting in a predominantly male participant pool. They were all full-time students, most of them enrolled in programs at the Technical University of Denmark or other danish universities, ensuring a consistent academic background within the group.

Participants were chosen based on their student status and general health without diagnosed physical or mental health issues. The study did not factor in language proficiency as the task involved letter recognition rather than full-word comprehension, so English proficiency was not a criterion for selection.

Power analyses were conducted to determine the necessary sample size for detecting effects within a repeated measures ANOVA design with an alpha level of 0.05 and a desired power of 0.95. Analyses were performed for effect sizes (Cohen's f) of 0.25, 0.15, and 0.3, with two groups and four measurements. For an effect size of 0.25, a total sample size of 36 participants was required to achieve a power of approximately 95.2%. When the effect size was reduced to 0.15, the required sample size increased to 98 participants to maintain the same level of power. Conversely, for a larger effect size of 0.3, the required sample size was reduced to 26 participants. These power calculations informed the experimental design, ensuring that the study was adequately powered to detect small to medium effects and minimizing the risk of Type II errors.

Thirteen participants were recruited for the study, a smaller number than anticipated based on the initial power analysis, which recommended a larger participant pool. However, limitations in resources, time constraints, and prior errors in experiment preparation dictated the smaller sample size.

Recruitment primarily occurred through personal contacts of the experiment conductors. All participants volunteered for the study and provided both verbal and written consent for their involvement. Personal data was not stored, ensuring anonymity, as participants were assigned random identification numbers for data storage purposes, with no linkage to individual test results.

2.2 Ethical Considerations

The study prioritized ethical guidelines to ensure participant well-being and avoid any potential harm or stress deliberately inflicted. The selection criteria excluded any involvement of unpleasant stimuli, invasive procedures, pain induction, deprivation of necessities like water, food, or sleep, drug administration, research on patients or at-risk individuals, and misleading or deceiving practices towards participants.

Participants were fully informed before their involvement in the experiment. Detailed information encompassing the experiment's purpose, procedures, duration, risks, anonymous data collection, voluntary participation, and contact information of experimenters was provided. They were made aware of the potential for noise to induce stress, anxiety, or fatigue.

To ensure participant well-being, the experiment was designed to avoid any distressing or harmful situations. Noise levels were maintained within safe limits to prevent auditory damage. Participants were monitored for signs of discomfort, and could take brakes while executing the tasks. They were also given the freedom to discontinue the experiment if they found the noise levels uncomfortable or distressing.

Measures were implemented to maintain participant anonymity and data confidentiality. No personally identifiable information beyond consent forms, which couldn't be linked to specific test results, was retained. Participants' data were stored under randomly assigned identification numbers to ensure privacy and confidentiality throughout the study.

2.3 Experimental Design

The experiment followed a within-subject design, where participants experienced different noise environments: pink noise, brown noise, and silence, played through headphones as the independent variable. The aim was to measure how these noise types affected participants' reaction time and accuracy in tasks, which served as the dependent variables.

Participant	First noise	Second noise	Third noise
P1	Brown	Pink	Silence
P2	Pink	Silence	Brown
P3	Silence	Brown	Pink
P4	Brown	Pink	Silence
P5	Pink	Silence	Brown
P6	Silence	Brown	Pink
P7	Brown	Pink	Silence
P8	Pink	Silence	Brown
P9	Silence	Brown	Pink
P10	Brown	Pink	Silence
P11	Pink	Silence	Brown
P12	Silence	Brown	Pink
P13	Brown	Pink	Silence

Table 1: Stratification table for independent variable order

The task itself was adapted from a study involving memory tests called 1-back and 2-back tasks. [5] Participants completed four tasks of each type for each noise environment. Within each task, they were presented with 12 letters on screen, and they had to recall if the current letter matched the one presented either one step back (1-back) or two steps back (2-back) in the sequence. To facilitate participant familiarity, a practice run including one 1-back and one 2-back test was provided. The 0-back task from the referenced paper was not included, as it doesn't measure working memory.

To avoid any potential bias, the order of the tests (1-back or 2-back) was randomized for each run. Additionally, the order of the noise environments was stratified to ensure fairness in the data collected, as seen in Table 1.

The entire test, including explanations, signing consent forms, the actual tasks, and post-task debrief, took approximately 15-20 minutes per participant. This time frame encompassed the entire process from start to finish for each participant involved in the study.

The experiment was designed using Psychopy Builder and incorporated code written by the experimenters, leveraging online resources for assistance. Efforts were made to control for specific variables:

- the time of day - although in the end, it varied from participant to participant, it was typically scheduled during study hours between 3 pm to 11 pm,
- individual differences - to attempt to mitigate them, averages of reaction times and correctness per participant per noise environment were considered, rather than analyzing all individual results,
- noise levels - they were maintained around 70 dB SPL (as recommended for cognitive tasks [8] across all experiment runs.

However, certain confounding factors remained. The experiment was conducted using three different laptops by three separate experimenters, potentially introducing variations in the participant briefing process, despite the attempts to standardize it through a participant briefing sheet. Additionally, the experiment took place in slightly different environments, although efforts were made to ensure a quiet setting with minimal disturbances from other individuals present.

The 1-back and 2-back tasks are cognitive assessments commonly used to evaluate working memory. In these tasks, participants are presented with a sequence of letters, each displayed on the screen for 0.5 seconds, followed by a 1.5-second fixation cross. Participants' objective is to determine if the current letter matches the one presented either one step back (1-back) or two steps back (2-back) in the sequence. For each letter, participants must press the space bar if it matches the target letter from the specified position in the sequence. These tasks demand continuous comparison and updating of letters in working memory, assessing participants' ability to retain and manipulate letter information in a time-constrained environment. The sequence of rapidly presented letters, each followed by a short pause, challenges memory retention, attention, and quick decision-making.

Despite testing and debugging of the experiment process, errors still occurred. Upon completing the initial round of data collection involving 14 participants, it became evident that the participants were expected to press the space bar significantly fewer times in each test than had been anticipated. The objective was to achieve a 30-40% rate of 'correct' space presses per test, yet some participants needed only one or zero presses. This discrepancy stemmed from an unnoticed code error, which was promptly addressed. However, this issue rendered the previously collected data unsuitable and it had to be discarded. Consequently, fewer data points were available for the intended analysis than originally planned.

2.4 Experimental Procedure

Participants were taken into a quiet room and received both verbal explanations and an information sheet detailing the study's objectives, procedures, and contact information for further inquiries. They were briefed on the nature of the tasks they would undertake, the expected duration of the study, and the nature of data collection, comprising reaction times and accuracy. Upon providing informed consent, participants completed a test session without headphones, engaging in one 1-back task involving 15 letters and one 2-back task with 15 letters.

Once participants demonstrated comprehension, they put on headphones through which the designated noise environment (or silence) was played. They then performed the main test, where they completed sequences of 1-back and 2-back tasks, responding to letter matches as instructed. The noise environment was altered for each subsequent test iteration, with participants having the option to take breaks between each 1-back or 2-back task to ensure comfort and readiness.

Post-test sessions involved debriefing participants, providing them with an opportunity to ask any remaining questions or share feedback on their experience throughout the study. Both the information sheet and consent form are available in the appendices for reference. This process facilitated a clear understanding of the study's purpose, task execution, and ensured participants' comfort and informed participation throughout the study.

2.5 Pre-processing

The preprocessing stage of our experiment involved several critical steps to ensure that the data collected were accurate, consistent, and suitable for analysis. These steps are outlined below:

2.5.1 *Data Collection Review.*

Initially, we conducted a thorough review of the data collected from each participant. This review aimed to identify any immediate inconsistencies, such as incomplete sessions or improperly recorded data points.

2.5.2 *Data Cleaning.*

In this stage of preprocessing, we focused on identifying and handling missing data and outliers, which are crucial for ensuring the integrity and reliability of our analysis.

Handling Missing Data: Due to the nature of our cognitive tests (1-back and 2-back), it was common for some trials not to yield a response or reaction time. This absence of data was primarily attributed to the participant's non-response in certain trials. We identified these instances, ensuring that missing data were only marked in cases where a response was expected but not recorded. This distinction was vital to maintain the accuracy of our data analysis.

Outlier Detection and Removal: To identify outliers, we employed both graphical and statistical methods. Initially, we generated boxplots for reaction times and accuracy scores across all participants and conditions. These plots provided a visual means to quickly identify data points that stood out from the rest due to their extreme values. Boxplots are particularly effective in highlighting outliers by displaying data spread and variability. Following the visual inspection, we applied a more quantitative approach. We calculated the mean and standard deviation for reaction times and accuracy scores within each condition. Data points lying beyond three standard deviations from the mean were flagged as potential outliers. This criterion is a widely accepted method in statistical analysis for identifying outliers, as it captures data points that are unusually distant from the typical range of values.

Evaluation of Outliers: Each identified outlier was then evaluated on a case-by-case basis. This evaluation involved examining the context of the data point such as the participant's overall performance pattern and any potential experimental errors or anomalies during data collection. This careful scrutiny ensured that only data points representing true deviations from normal participant behavior were removed, while preserving legitimate extreme values that are part of natural variability in human performance.

2.5.3 *Accuracy Calculation.*

Accuracy was calculated as the proportion of correct responses out of the total number of trials where a response was expected. Trials with incorrect responses were counted as errors in this calculation. This measure provided a straightforward assessment of participants' performance accuracy across different noise conditions.

2.5.4 *Aggregation of Data.*

An essential step in our preprocessing was the aggregation of data, which involved calculating the mean reaction time and accuracy for each participant under each of the noise conditions (silence, pink, and brown). This step was integral to the analysis for several reasons. By aggregating the data, we simplified the complex raw data into a more manageable format. This was crucial for the planned statistical tests, as it allowed us to apply methods like the Friedman Test. Aggregated data provided a clear, summarized view of each participant's performance across different conditions.

Mean Reaction Time Calculation: For reaction time, we computed the mean across all trials for each participant in each noise condition. This approach provided a singular, representative value of the participant's average response speed under each type of noise. It was vital to consider all valid responses to capture the participant's overall reaction tendency in each condition.

Accuracy Computation: Similarly, we calculated the accuracy as the proportion of correct responses to the total number of trials in each condition for each participant. This measure offered a concise yet comprehensive representation of the participant's performance accuracy, integrating their success rate across numerous trials.

Aggregation helped balance the data across different conditions. Given the within-subjects design, each participant experienced all noise conditions. Aggregating data ensured that each condition was equally represented and that the comparison across conditions was fair and balanced. By focusing on mean reaction time and accuracy, we also reduced the impact of variability in individual performance, allowing us to observe the effects of noise conditions more clearly. The aggregated data were then structured suitably for the subsequent statistical analysis. By having a consistent data format, we ensured that our statistical tests could be applied uniformly across all conditions and participants.

2.5.5 *Quality Checks.*

Finally, we conducted a series of quality checks. These included verifying the consistency of data entry, ensuring that all participants were exposed to the same experimental conditions, and confirming that the timings of the noise exposures were accurately recorded.

2.6 Data Analyses

The data analysis phase of our study involved several crucial steps, systematically applied to investigate the effects of different noise conditions on cognitive performance. These steps are detailed below. Our primary objective was to determine whether the type of noise (silence, pink, brown) had a statistically significant effect on participants' cognitive performance, as measured by reaction time and accuracy in 1-back and 2-back tests.

2.6.1 *Initial Data Exploration.*

Our initial exploration of the dataset served as a crucial first step in understanding the underlying trends and patterns in cognitive performance under different noise conditions. We computed descriptive statistics (mean, median, standard deviation) for reaction times and accuracy, which provided a quantitative overview of participants' performances across silence, pink, and brown noise conditions. Additionally, visual tools like box plots highlighted potential outliers and distribution shapes, line graphs depicted trends over time or conditions, and scatter plots revealed potential correlations or groupings. This visual and statistical exploration offered valuable insights and helped formulate hypotheses for more detailed analyses.

2.6.2 *Data Cleaning and Standardization.*

The integrity of our study hinged on the cleanliness and standardization of our data. In this phase, we meticulously filtered out anomalies and inconsistencies that could potentially skew our results. We standardized reaction times and accuracy values across participants and conditions to ensure uniformity in measurement scales and comparability. This step was particularly crucial in cognitive experiments, where even minor variations in data processing can lead to significant misinterpretations. Standardization also facilitated a more straightforward and meaningful comparison across different noise conditions.

2.6.3 *Assessment of Data Distribution.*

Following data cleaning, we turned our attention to assessing the distribution patterns of our dataset. This assessment was instrumental in guiding the selection of the most appropriate statistical tests. For data segments that adhered to the principles of normal distribution and homogeneity of variance, we leaned towards parametric tests. However, for parts of the dataset that deviated from these assumptions a common occurrence in behavioral and cognitive data we considered non-parametric alternatives or more complex models. This careful scrutiny of data distribution was a critical step in ensuring the validity and accuracy of our statistical interpretations.

2.6.4 *Laying the Groundwork for Analysis.*

The preliminary stage of our analysis, often understated in its importance, set the foundation for all subsequent analyses. By ensuring the data was clean, standardized, and categorized appropriately based on its distributional properties, we established a robust base for our analytical journey. This foundation wasn't solely about prepping data for statistical analysis. It was more like laying the groundwork for a deep investigation into how hearing things and thinking are linked together. We therefore aligned our analytical methods with the complex nature of cognitive processes, ensuring that our approach was as nuanced and multifaceted as the phenomena we aimed to explore.

2.7 Statistical Analysis

In the realm of cognitive science, where the subtleties of human perception and memory intertwine with quantifiable data, the role of statistical analysis becomes paramount. Our study, centered on evaluating the impact of colored noise on working memory, embarks on this analytical journey with a clear objective: to decipher the intricate relationship between auditory stimuli and cognitive performance. The statistical analysis thus helps us to look into the finer details of cognitive processes.

The unique nature of our data, encompassing varied responses from participants under different auditory conditions, necessitates a dual approach in analysis. This bifurcation is driven by the nature of the data itself and the underlying questions we seek to answer. On one hand, we have the averaged data analysis, a method that allows us to grasp overarching trends and general patterns. This approach is instrumental in painting a broad picture of how noise conditions collectively influence cognitive metrics like reaction time and correctness.

On the other hand, delving into individual response data unveils a different layer of understanding. Here, the granular details come into play, highlighting individual variabilities and specific interactions that might be obscured in a more aggregated analysis. It is in this detailed examination that the nuances of cognitive responses to auditory stimuli are most vividly observed.

In embarking on this two-pronged analytical approach, our study aligns itself with the rigorous standards of cognitive science research. We are not only looking to confirm or refute our hypotheses but also to contribute to the broader understanding of cognitive processes. The statistical analysis, in its comprehensive form, stands as the backbone of this endeavor, ensuring that our conclusions are not just data-driven but are also steeped in scientific rigor and relevance.

2.7.1 *Shapiro-Wilk Test for Normality.*

In our statistical analysis, a fundamental step involved assessing the normality of our data, specifically for reaction time and accuracy across different noise conditions (brown, pink, silence). To accomplish this, we employed the Shapiro-Wilk Test, a widely recognized method for testing the normality of data.

Importance of Normality Testing

Normality testing is a crucial prerequisite in many statistical analyses, as it determines the suitability of parametric statistical tests. These tests, such as ANOVA, rely on the assumption that data are normally distributed. Normality in data ensures the validity of inferences drawn from these tests, making this assessment an essential part of the analytical process.

Shapiro-Wilk Test Procedure

The Shapiro-Wilk Test is particularly effective for small to moderate sample sizes, making it a suitable choice for our study. The test operates by comparing the data in question to a perfectly normal distribution. It tests against a null hypothesis that the data follows a normal distribution.

Application to Reaction Time and Accuracy Data

We conducted the Shapiro-Wilk Test separately for reaction time and accuracy data within each of the noise conditions (brown, pink, silence). This allowed us to evaluate the distribution characteristics unique to each condition. The test computes a W statistic a measure of how closely the data conform to a normal distribution. Accompanying this statistic is a p-value, which is used to determine the significance of the results:

- A high p-value (typically >0.05) suggests that the data do not significantly deviate from normality, supporting the null hypothesis.
- A low p-value (0.05) indicates a significant deviation from normality, leading to the rejection of the null hypothesis.

Interpretation of Results

The results from the Shapiro-Wilk Test guided our subsequent choice of statistical methods. For data exhibiting normality, parametric tests were considered appropriate. Conversely, for data deviating from normality, non-parametric alternatives were explored to ensure the robust

Not centered?

2.7.2 Theoretical Foundations of Statistical Methods.

ANOVA (Analysis of Variance)

ANOVA is a fundamental statistical method for comparing means across multiple groups, crucial in determining if there are significant differences among them. It decomposes data variance into components due to various sources.

Assumptions: ANOVA assumes that variances are equal across groups (homogeneity of variance), evaluated using Levene's test. A significant result in Levene's test indicates a violation of this assumption, potentially requiring data transformation or a switch to non-parametric methods.

Components of ANOVA: ANOVA segregates the total observed variance into between-group (systematic) and within-group (unsystematic or error) variances. The F-ratio, crucial in ANOVA, is the ratio of systematic to unsystematic variance, measuring the experimental effect's relative size. The p-value associated with the F-ratio indicates the probability of observing such a ratio purely by chance. A low p-value (<0.05) suggests that the observed differences between group means are statistically significant.

Friedman Test

The Friedman Test is a non-parametric alternative to ANOVA, suitable for ordinal or non-normally distributed data across multiple matched groups or repeated measures.

Usage Context: This test is essential when normality assumptions required for ANOVA are violated, particularly in repeated measures designs.

Test Procedure: The Friedman Test ranks data across groups and compares these ranks, making it effective for data not meeting normality assumptions. The test generates a Chi-square statistic, and the associated p-value helps determine the likelihood of observing the data if the null hypothesis were true. A low p-value suggests significant differences across the groups.

Generalized Linear Mixed Models (GLMM)

GLMMs are sophisticated models that extend linear models to accommodate complex data structures, such as non-normal response distributions, and include both fixed and random effects.

Flexibility: GLMMs are versatile, capable of modeling various data distributions (like binomial for binary data, Poisson for count data) and incorporating fixed effects (like experimental treatments) and random effects (like subject variability).

Application: In cognitive science, GLMMs are valuable for modeling individual response variations under different experimental conditions, providing insight into individual differences and interactions. The model outputs include coefficients for fixed and random effects, standard errors, and p-values for each coefficient. The p-values help determine the statistical significance of each predictor in the model, indicating whether a particular factor has a significant impact on the response variable.

Rationale for Statistical Tests

Our study employs a sophisticated approach to statistical testing, designed to explore the intricate effects of pink and brown noise on cognitive performance. To capture the intricacies of our data and ensure a robust analysis, we have selected appropriate statistical methods that align with the distribution and characteristics of our collected data.

The Dependent Variables

In investigating the influence of noise on cognitive performance, we focus on two central dependent variables: accuracy and reaction time.

Accuracy Rate: Accuracy is the cornerstone of working memory assessment, indicating how effectively information is retained and utilized over short intervals. It serves as a critical gauge for the cognitive load and efficacy of information processing under varying auditory conditions.

Reaction time: offers insights into the swiftness and efficiency of cognitive operations, reflecting the immediacy of response to stimuli. It provides a complementary dimension to accuracy, painting a fuller picture of cognitive performance in the presence of different background noise types.

These variables were carefully selected for their proven sensitivity to cognitive load and their relevance to the domain of working memory. They allow us to delve into the cognitive impacts of noise exposure, providing a dual lens through which to view the participants' performance.

2.7.3 Selecting Statistical Tests for Averaged Data.

Condition	Metric	Test Statistic	P-value	Normality
Brown	Reaction Time	0.9689	0.8806	Yes
Brown	Accuracy	0.8131	0.0098	No
Pink	Reaction Time	0.9518	0.6252	Yes
Pink	Accuracy	0.8216	0.0125	No
Silence	Reaction Time	0.9232	0.2766	Yes
Silence	Accuracy	0.9240	0.2838	Yes

Table 2: Shapiro-Wilk Test Results for Normality

Rationale Behind Test Selection

In the context of our cognitive science experiment, the selection of appropriate statistical tests for the averaged data was guided by the specific structure of our dataset and the nature of the variables involved. With the averaging of data across n-back values of 1 and 2, the distinction between these two levels was integrated into a unified metric. This averaging effectively transformed our data, removing 'n-back value' as a contributing factor, and focusing our analysis on the overarching impact of noise conditions. The simplification of the data structure to these essential elements was instrumental in determining the most fitting statistical tests.

Evaluating Dependent Variables: Normality, Sphericity, and Homoscedasticity Tests

The integrity and suitability of our dependent variables reaction time and correctness for specific statistical tests were assessed through a series of diagnostic checks :

Reaction Time (rt): We performed normality, sphericity, and homoscedasticity tests on the reaction time data. These tests were crucial in confirming that the data met the assumptions required for parametric testing, justifying the use of the One-Way Repeated Measures ANOVA (Analysis of Variance). The ANOVA's ability to handle repeated measures and within-subject variations made it an ideal choice for analyzing reaction time across different noise conditions.

Correctness (is_correct): In contrast, the correctness data underwent the same series of tests, which revealed a non-normal distribution. This finding necessitated the shift to a non-parametric approach, leading to the selection of the Friedman Test. The test's suitability for ordinal or non-normally distributed data across multiple conditions made it the optimal choice for assessing correctness scores under varying auditory stimuli.

Application of Statistical Tests on Averaged Data

After thorough preparatory evaluations, we embarked on applying the selected statistical tests to our averaged dataset. This phase was critical in translating the theoretical underpinnings of our methods into practical, empirical analysis.

One-Way Repeated Measures ANOVA for Reaction Time

The One-Way Repeated Measures ANOVA was meticulously applied to the reaction time data. The essence of this analysis was to investigate if, and to what extent, the different noise conditions silence, pink noise, and brown noise influenced the participants' cognitive processing speed. This test is particularly adept at handling situations where the same subjects are exposed to multiple conditions, as was the case in our study. It accounts for individual variability among participants, allowing us to isolate the effects of the noise conditions from other influencing factors.

The ANOVA's capability to analyze within-subjects effects was crucial for our investigation. It enabled us to discern whether the variations in reaction times were significantly attributable to the different auditory stimuli or merely random fluctuations. By comparing the mean reaction times across each noise condition, the ANOVA provided a robust statistical framework to determine the presence and strength of any auditory effects on cognitive speed.

Friedman Test for Correctness

In parallel, the Friedman Test was employed to analyze the correctness data. This non-parametric test was particularly suited for our data, which, as established earlier, did not follow a normal distribution. The Friedman Test compared the correctness of participants responses across the three noise conditions, offering a statistical lens through which to view the impact of auditory stimuli on task accuracy.

This test is valuable in scenarios where data distribution deviates from normality, as it does not rely on parametric assumptions. Instead, it ranks the data and compares these ranks across different groups in our case, the different noise conditions. By doing so, the Friedman Test provided insights into whether the type of background noise had a measurable effect on the participants' ability to perform cognitive tasks correctly. It allowed us to understand if certain noise conditions led to higher accuracy in responses, highlighting the potential influence of auditory environments on cognitive precision.

2.7.4 *Detailed Analysis with Individual Response Data.*

Utilizing Generalized Linear Mixed Models (GLMM)

The journey of our cognitive science experiment advanced into a detailed examination of individual response data, a crucial phase for unraveling the subtleties hidden within the averaged data. To achieve this, we employed Generalized Linear Mixed Models (GLMM), an analytical tool adept at handling complex data structures.

Selection of GLMM: Informed by Assumption Testing

The choice of GLMM was informed by a rigorous evaluation of our data against key statistical assumptions. This evaluation included tests for normality, sphericity, and homoscedasticity, each integral in determining the appropriateness of statistical models for our dataset.

Normality Testing: We observed that individual response data exhibited non-normal distribution patterns. This deviation from normality is a common occurrence in behavioral data, indicating that traditional linear models might not be suitable.

Sphericity and Homoscedasticity: The assessment of sphericity and homoscedasticity revealed complexities in the data that required a more nuanced approach. The presence of these conditions suggested that our data contained intricacies that could be best captured through models accounting for random effects and non-normal distributions.

Given these findings, GLMM emerged as the optimal choice, offering the flexibility to accommodate the unique characteristics of our dataset.

GLMM for Reaction Time and Correctness Data

Reaction Time: Utilizing GLMM for reaction time data enabled us to model the effects of various noise conditions while accounting for random variability among participants. This approach provided a comprehensive understanding of how individual differences interact with auditory stimuli to affect cognitive processing speed.

Correctness: For the correctness data, GLMM was instrumental in exploring the accuracy of cognitive task performance across different noise environments. It allowed us to delve into the variability of task correctness among individuals, revealing the nuanced ways in which the auditory environment could influence cognitive accuracy.

The Interaction Effect Analysis

A pivotal aspect of our GLMM analysis was investigating interaction effects. This investigation was crucial for understanding how different factors - such as noise types, individual participant characteristics and type of n-back test jointly influenced cognitive performance. In cognitive science, where human cognition is influenced by a myriad of factors, such an analysis provides invaluable insights into the complex workings of the mind.

Through this in-depth analysis with individual response data, our study gained a richer, more detailed perspective. By employing GLMM, we could not only corroborate our initial findings from the averaged data but also illuminate the diverse individual responses and complex interactions that shape cognitive performance in varied auditory settings.

3 RESULTS

Our study embarked on an investigative journey to unravel the effects of different auditory environmentsspecifically, silence, pink noise, and brown noiseon cognitive performance. This exploration was critical in understanding how subtle variations in background noise could potentially influence cognitive tasks involving memory and attention. To achieve a comprehensive analysis, our approach integrated both averaged and individual response data, employing a robust combination of statistical techniques including ANOVA, the Friedman Test, and Generalized Linear Mixed Models (GLMM).

The primary aim was to ascertain whether these auditory conditions exerted a measurable impact on two key aspects of cognitive performance: reaction time and accuracy. These metrics were carefully chosen as indicators of cognitive processing speed and effectiveness, respectively, in response to the 1-back and 2-back tests administered to participants. Our analysis was designed not only to test our initial hypotheses regarding the influence of noise conditions but also to delve deeper into the nuances of individual responses and potential interactions within the data.

The statistical methods were selected to align with the specific characteristics of our data and the nature of our research questions. The One-Way Repeated Measures ANOVA was applied to reaction time data to explore within-subject differences across noise conditions. In contrast, the Friedman Test was employed for accuracy data, given its suitability for non-normally distributed or ordinal data typically encountered in cognitive science research. Furthermore, the GLMM was utilized to dissect the intricacies of individual response patterns, providing a more granular understanding of how each participant's cognitive performance varied under different auditory stimuli.

Our analytical strategy entailed synthesizing data across different levels of analysis, from a broad overview of trends across all participants to a detailed examination of individual differences. This multifaceted approach allowed us to capture both the collective and individual impacts of noise conditions, offering a holistic view of their effects on cognitive performance.

3.1 Descriptive Statistics

The preliminary analysis of our data involved an exploration of descriptive statistics, crucial for establishing a baseline understanding of participants' cognitive performance under varying noise conditions. Here, we focused on two primary metrics: reaction time and accuracy, both integral to assessing cognitive processing and effectiveness.

Reaction time

The mean reaction time observed in the silence condition was 0.95913 seconds. This value serves as a baseline, representing participants' performance in the absence of auditory distraction or enhancement.
Pink Noise: Under pink noise, the mean reaction time was 0.96955 seconds. Comparatively, this measure provided an initial indication of how a consistent, homogeneous auditory background influences cognitive response speed.
Brown Noise: In the brown noise condition, participants exhibited

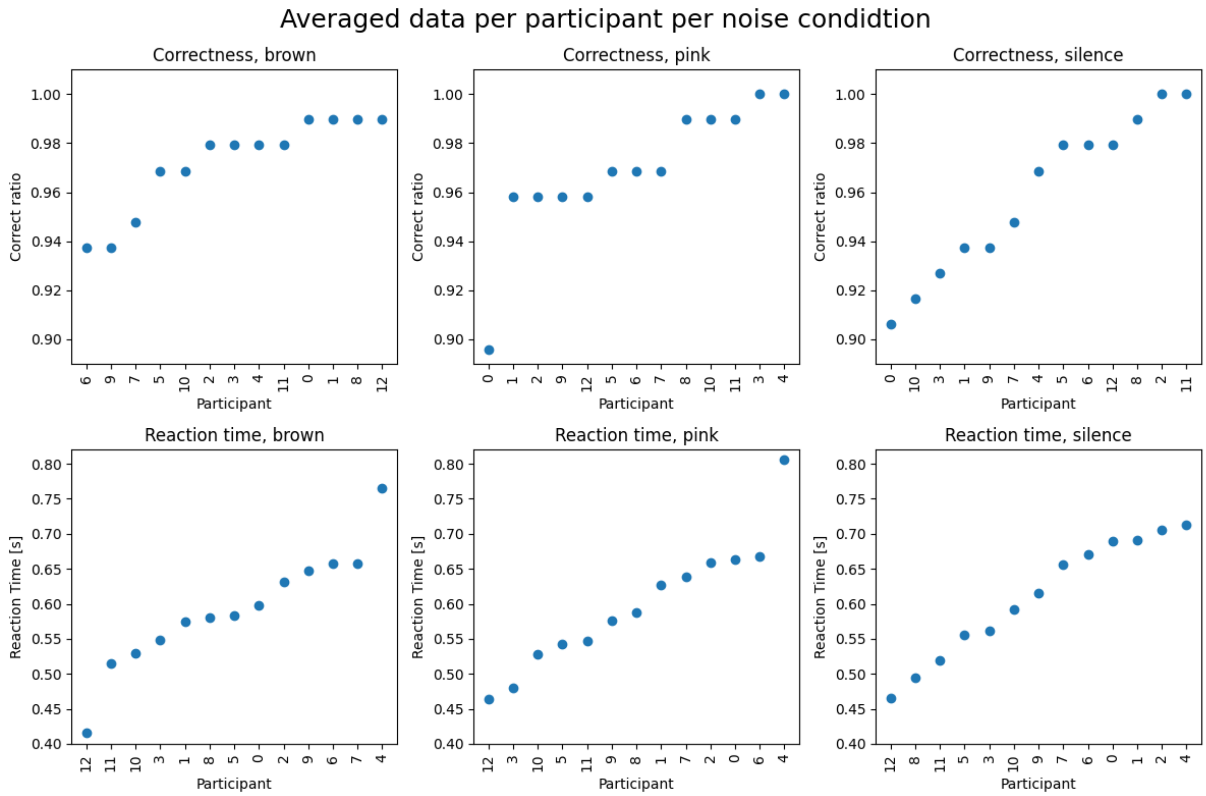

Figure 1: Pre-processed, averaged data per participant per noise condition

Session	Count	Mean	Std	Min	25%	50%	75%	Max
Brown	13	0.5926	0.0847	0.4165	0.5490	0.5837	0.6477	0.7649
Pink	13	0.5992	0.0920	0.4645	0.5428	0.5875	0.6590	0.8063
Silence	13	0.6100	0.0851	0.4650	0.5561	0.6158	0.6898	0.7134

Table 3: Descriptive Statistics for Reaction Time

Session	Count	Mean	Std	Min	25%	50%	75%	Max
Brown	13	0.9720	0.0192	0.9375	0.9688	0.9792	0.9896	0.9896
Pink	13	0.9696	0.0274	0.8958	0.9583	0.9688	0.9896	1.0000
Silence	13	0.9591	0.0320	0.9063	0.9375	0.9688	0.9792	1.0000

Table 4: Descriptive Statistics for Correctness

a mean reaction time of 0.97195 seconds. Brown noise, known for its deeper frequency, offered a contrast to pink noise in its potential cognitive impact.

Accuracy rate

The accuracy rate in the silence condition was 0.959%. This percentage reflects the participants' ability to correctly respond to the cognitive tasks when in a standard, noise-free environment.
Pink Noise: With pink noise, the accuracy rate was 0.969%. This measure was pivotal in understanding whether a steady auditory background aids or hinders concentration and correct information processing.
Brown Noise: The accuracy rate for brown noise stood at 0.971%. The deeper quality of brown noise provided an interesting variable to assess its distinct influence on cognitive accuracy compared to pink noise or silence.
These descriptive statistics were foundational in providing an overarching view of how different noise conditions might affect cognitive performance. Notably, the variations in mean reaction times and accuracy rates across conditions began to paint a picture of the auditory stimuli's potential impact. The comparison between the

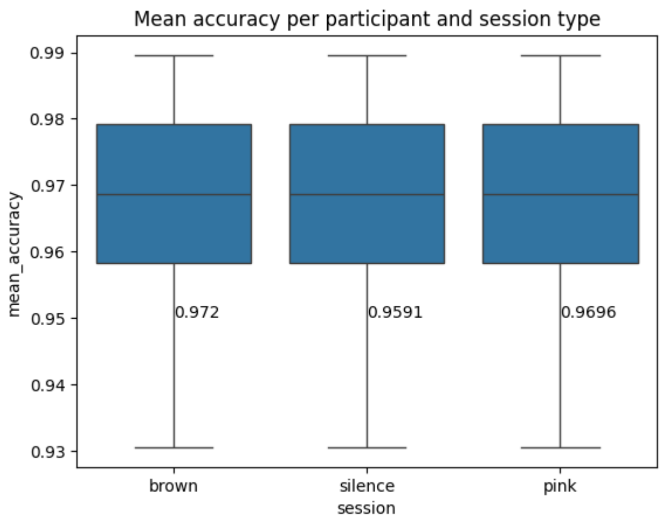

Figure 2: Average Accuracy Per Participant

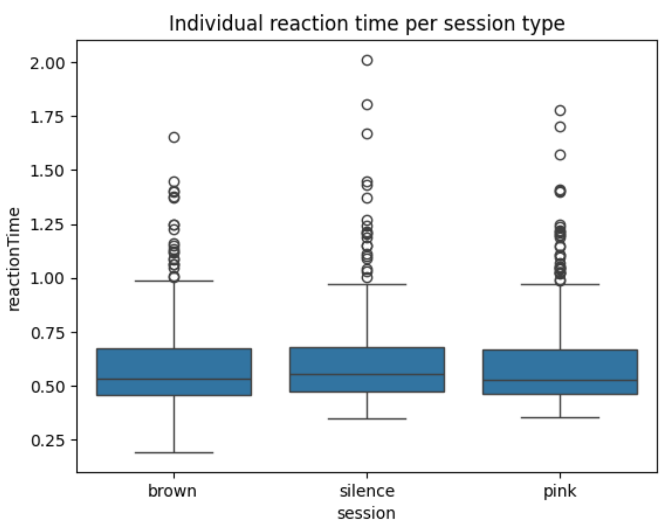

Figure 3: Average Reaction Time per Participant

mean reaction times and accuracy rates in different noise environments laid the groundwork for our subsequent, more detailed statistical analysis. It was crucial to understand these baseline trends to contextualize our findings within the broader scope of cognitive response to auditory stimuli.

While these initial statistics did not yet delve into the significance or causality of the observed differences, they were essential in guiding our hypothesis testing. They provided a quantitative snapshot of the study's raw data, setting the stage for the in-depth analyses that followed.

3.2 Assumptions testing

In our study, the Shapiro-Wilk tests played a crucial role in evaluating the normality of our data, a fundamental assumption for certain statistical analyses:

Test Results for Brown Noise

Reaction Time: The test statistic was 0.9689 with a p-value of 0.8806. This indicates that the reaction time data under brown noise conditions followed a normal distribution. Accuracy: The test statistic was 0.8131 with a p-value of 0.0098. This suggests that the accuracy data under brown noise conditions did not follow a normal distribution. Test Results for Pink Noise:

Reaction Time: The test statistic was 0.9518 with a p-value of 0.6252. This implies that the reaction time data under pink noise conditions were normally distributed. Accuracy: The test statistic was 0.8216 with a p-value of 0.0125. This indicates a deviation from normal distribution in the accuracy data under pink noise conditions. Test Results for Silence:

Reaction Time: The test statistic was 0.9232 with a p-value of 0.2766, suggesting that the reaction time data in silence conditions followed a normal distribution. Accuracy: The test statistic was 0.9240 with a p-value of 0.2838, indicating that the accuracy data in silence conditions were normally distributed.

Interpretation and Implications

These results informed our selection of subsequent statistical methods. The normality in reaction time data across all conditions allowed us to use parametric tests like the One-Way Repeated Measures ANOVA. The non-normal distribution of accuracy data in brown and pink noise conditions led us to choose the Friedman Test, a non-parametric method, for analyzing accuracy data.

Overall Significance

The Shapiro-Wilk test results played a pivotal role in ensuring that our statistical analysis was grounded on valid assumptions. By accurately identifying the distribution characteristics of our data, we were able to select the most appropriate statistical tests, bolstering the reliability and validity of our findings.

Reaction Time Data Normality

The Shapiro-Wilk tests applied to the reaction time data indicated a general adherence to a normal distribution across the different noise conditions (silence, pink, and brown noise). These results were critical as they suggested that parametric tests, which assume normal distribution, could be appropriately used for analyzing reaction time data.

The normality in reaction times across conditions also implied a degree of consistency in participants' response speeds, regardless of the auditory environment.

Accuracy Data Normality

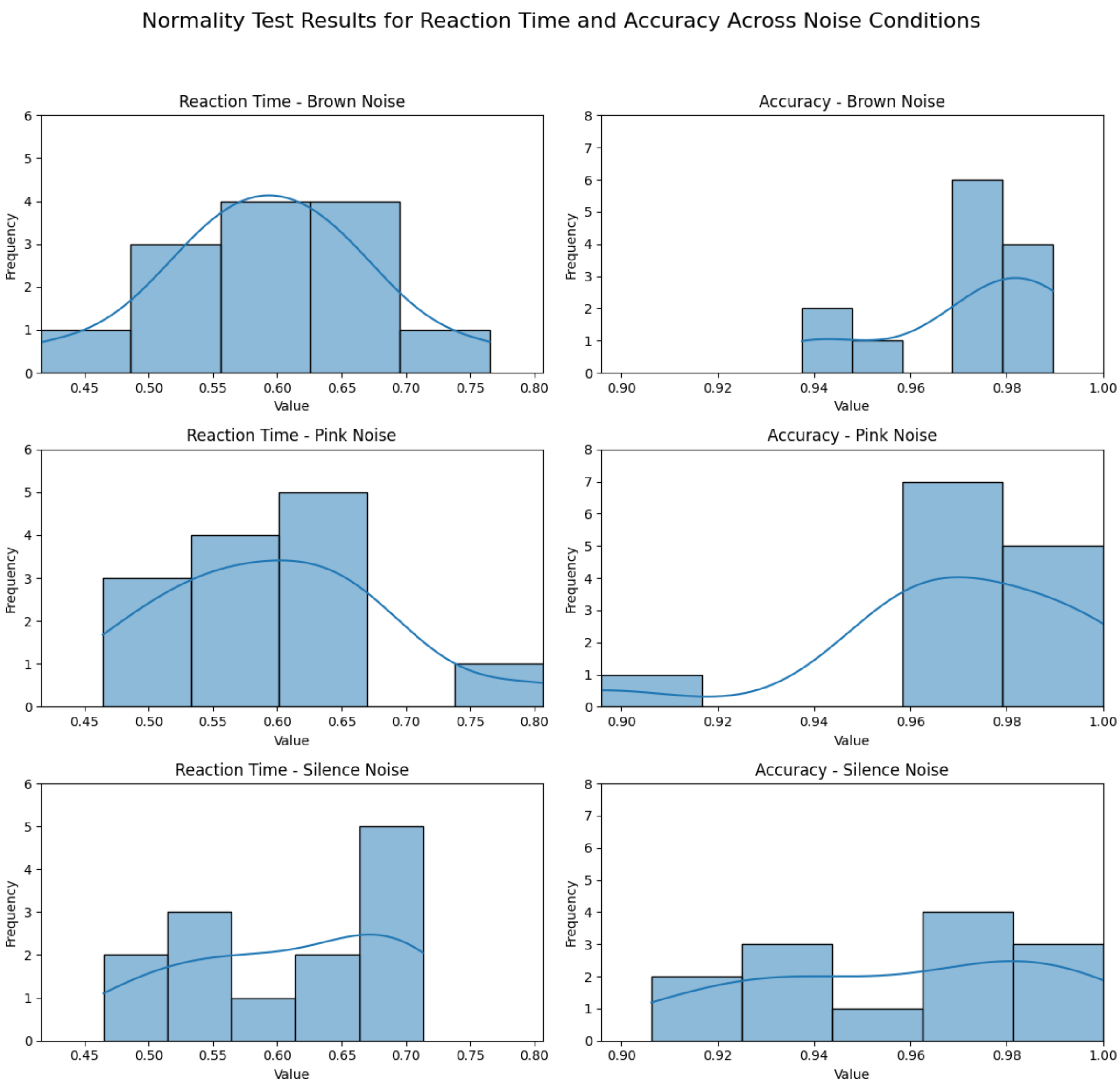

Figure 4: Distribution Plots

Condition	Metric	Test Statistic	P-value	Normality
Brown	Reaction Time	0.9689	0.8806	Yes
Brown	Accuracy	0.8131	0.0098	No
Pink	Reaction Time	0.9518	0.6252	Yes
Pink	Accuracy	0.8216	0.0125	No
Silence	Reaction Time	0.9232	0.2766	Yes
Silence	Accuracy	0.9240	0.2838	Yes

Table 5: Shapiro-Wilk Test Results for Normality

Conversely, the Shapiro-Wilk tests for accuracy data yielded different results. Specifically, in the pink and brown noise conditions, the data did not conform to a normal distribution. This deviation from normality was particularly notable, as it suggested that the

Data Type	Homoscedasticity test p-value
Reaction Time	0.8288
Correctness	0.1481

Table 6: Homoscedasticity Test Results for Reaction Time and Correctness

ANOVA Results	
F Value	0.7039
Num DF	2.0000
Den DF	24.0000
Pr > F (p-value)	0.5046
Friedman Test Results	
Statistic	0.3111
p-value	0.8559

Table 7: ANOVA and Friedman Test Results

accuracy of cognitive task performance was subject to greater variability under these noise conditions. The lack of normal distribution in accuracy data under pink and brown noise necessitated the use of non-parametric tests, such as the Friedman Test, which do not require the assumption of normality. Implications of Normality Test Outcomes:

These findings were instrumental in guiding our choice of statistical methods. For reaction time, the normality supported the use of One-Way Repeated Measures ANOVA, while for accuracy, the non-normal distribution led us to opt for the Friedman Test. The difference in normality between reaction time and accuracy data also provided an early indication of the distinct ways in which these cognitive performance measures might respond to different auditory stimuli.

3.3 One-Way Repeated Measures ANOVA (Reaction Time)

In analyzing the impact of noise conditions on reaction time, we employed the One-Way Repeated Measures ANOVA, a statistical test particularly suited for within-subjects designs like ours.

Statistical Outcome

The ANOVA results indicated no statistically significant difference in reaction times across the three noise conditions: silence, pink noise, and brown noise. This lack of statistical significance suggests that the variations in reaction times observed across the different auditory conditions could be attributed to natural variability in human performance rather than the specific type of background noise.

Interpretation of the ANOVA Result

The F Value in ANOVA (ref. Table 7) represents the ratio of the variance explained by the model (between groups) to the variance within groups (unexplained by the model). In this case, an F Value of 0.7039 suggests a relatively low ratio of variance explained by the noise conditions ('session') compared to the variance within each condition.

The p-value of 0.5046, which is greater than the standard significance level of 0.05, indicates that the differences in means across the different noise conditions (sessions) are not statistically significant. In simpler terms, there's no strong evidence to suggest that the noise conditions have a significant effect on the dependent variable being measured.

The absence of a significant difference implies that the auditory environment, whether quiet or infused with pink or brown noise, did not alter the speed with which participants responded to the cognitive tasks. This finding is particularly intriguing as it challenges some common assumptions about the influence of ambient noise on cognitive processing speed. It suggests a level of resilience or adaptability in cognitive performance, irrespective of the auditory backdrop.

Contextualizing the Findings

These results contribute to the ongoing discourse in cognitive science regarding the role of environmental factors in shaping cognitive abilities. While some studies have highlighted the potential distractions or enhancements caused by background noise, our findings suggest a more complex interplayor lack thereofbetween auditory stimuli and reaction time in cognitive tasks.

Considerations and Limitations

It is important to consider these results within the limitations of our study, including the sample size and the specific nature of the tasks employed. The lack of significant difference does not necessarily negate the potential impact of noise on cognitive performance in different contexts or with different tasks. Additionally, the consistency in reaction times across noise conditions could be indicative of the tasks' design, possibly not being sufficiently sensitive to detect subtle changes in cognitive speed due to auditory stimuli.

3.4 Friedman Test (Accuracy)

To assess the impact of different noise conditions on the accuracy of participants in cognitive tasks, we applied the Friedman Test. This non-parametric test was particularly chosen due to its appropriateness for the ordinal or non-normally distributed nature of our accuracy data.

Results of the Friedman Test

The Friedman Test did not reveal statistically significant differences in accuracy across the three noise conditions: silence, pink noise, and brown noise. This lack of significant difference suggests that the type of background noise present during the cognitive tasks did not have a discernible impact on the participants' ability to respond correctly.

Interpreting the Lack of Significant Differences

The Friedman test is a non-parametric test used for comparing more than two groups when the dependent variable is ordinal or the data do not meet the assumptions necessary for ANOVA. The test statistic (0.3111) here is derived from comparing the ranks of the data across the groups.

The p-value of 0.8559 is much higher than the conventional alpha level of 0.05, suggesting that there are no significant differences in the ranks across the different noise conditions. Essentially, this result implies that the noise conditions do not significantly affect the dependent variable under consideration.

This finding indicates a level of robustness in cognitive accuracy against the variations in auditory conditions tested. It implies that participants' ability to correctly recall and process information in the 1-back and 2-back tasks was not noticeably affected by the presence or type of background noise. The absence of a significant effect of noise on accuracy is noteworthy. It challenges certain expectations about the cognitive load that different auditory environments might impose, suggesting that the type of noise may not be as influential on task correctness as might be assumed.

Contextualizing Within Cognitive Science Research

These results contribute to a nuanced understanding of the relationship between environmental stimuli and cognitive task performance. While numerous studies have explored the influence of environmental factors on cognition, our findings highlight the potential for cognitive task performance to remain consistent across varied auditory environments. It's important to consider these findings in light of the specific tasks and noise conditions used in our study. The nature of the tasks or the characteristics of the noise types may have influenced this outcome.

3.5 Generalized Linear Mixed Models (GLMM)

In our study, Generalized Linear Mixed Models (GLMM) were utilized to delve deeper into the intricacies of individual responses and to explore potential interaction effects between various factors. This advanced statistical approach allowed us to account for both fixed effects (such as noise conditions) and random effects (like individual participant variability).

GLMM Analysis for Reaction Time

The GLMM is used to analyze the effect of various independent variables on Reaction Time.

- The model includes a constant term (`const`) and three independent variables (`x1`, `x2`, `x3`).
- The dependent variable is Reaction Time (`y`).
- The analysis is based on 1307 observations grouped into 13 clusters.

Interpretation of Coefficients

- **const coefficient:** The constant term `const` represents the intercept of the model. It is approximately 0.452. This indicates the estimated baseline reaction time when all independent variables are set to zero.
- **x1 coefficient:** The coefficient for `x1` is approximately 0.095. It is statistically significant ($p < 0.05$), suggesting that the nback variable (represented by `x1`) has a positive effect on reaction time. As the nback variable increases, reaction time tends to increase as well.
- **x2 coefficient:** The coefficient for `x2` is approximately 0.010. It is not statistically significant ($p > 0.05$), indicating that `x2` does not have a significant effect on reaction time.
- **x3 coefficient:** The coefficient for `x3` is approximately 0.018. It is also not statistically significant ($p > 0.05$), suggesting that `x3` does not have a significant effect on reaction time.

Our GLMM analysis of reaction time data revealed insightful findings. We found that there were no significant main effects of the noise conditions on reaction times (p-values for all conditions > 0.05). Additionally, the model did not indicate any significant interaction effects between the noise conditions and other variables in the study. This suggests that the speed of cognitive processing among participants remained consistent across different auditory backgrounds. The lack of significant findings in reaction times across noise conditions, as revealed by GLMM, further reinforces the notion that the type of background noise might not play a significant role in influencing the speed of cognitive responses in tasks like the ones used in our study.

GLMM Analysis for Accuracy

Similarly, the GLMM applied to accuracy data showed that there were no significant differences in accuracy attributable to the different noise conditions (p-values for all conditions > 0.05).

Model Information

The GLMM is used to analyze the effect of various independent variables on the Correctness of Responses.

- The model structure is similar to the Reaction Time analysis, including a constant term (`const`) and three independent variables (`x1`, `x2`, `x3`).
- The dependent variable is the Correctness of Responses (`y`).
- The analysis is based on 1307 observations grouped into 13 clusters.

Interpretation of Coefficients

- **const Coefficient:** The constant term `const` is extremely close to zero and not statistically significant. This suggests that the intercept does not significantly affect the correctness of responses.
- **x1 Coefficient:** The coefficient for `x1` is approximately -0.013. However, it is not statistically significant ($p > 0.05$), indicating that the nback variable (`x1`) does not have a significant effect on correctness. In other words, the correctness

Table 8: GLMM Analysis Results for Reaction Time and Correctness

			Reaction Time		
Coefficient	**Estimate**	**Std. Error**	**z value**	**P>|z|**	**95% CI**
Constant	0.452	0.028	16.031	<0.001	[0.396, 0.507]
x1	0.095	0.010	9.189	<0.001	[0.074, 0.115]
x2	0.010	0.013	0.823	0.410	[-0.014, 0.035]
x3	0.018	0.013	1.427	0.154	[-0.007, 0.043]
Group Variance	0.006	0.014			
			Correctness of Responses		
Coefficient	**Estimate**	**Std. Error**	**z value**	**P>|z|**	**95% CI**
Constant	-0.000	726263.856	-0.000	1.000	[-1423451.001, 1423451.001]
x1	-0.013	0.010	-1.332	0.183	[-0.032, 0.006]
x2	0.015	0.012	1.243	0.214	[-0.008, 0.038]
x3	-0.003	0.012	-0.273	0.785	[-0.027, 0.020]
Group Variance	0.000				

of responses is not significantly influenced by changes in nback.

- **x2 Coefficient:** The coefficient for x2 is approximately 0.015. Like x1, it is not statistically significant ($p > 0.05$), suggesting that x2 does not have a significant effect on correctness.
- **x3 Coefficient:** The coefficient for x3 is approximately -0.003 and is also not statistically significant ($p > 0.05$). This implies that x3 does not have a significant effect on correctness.

This consistency in accuracy rates across silence, pink noise, and brown noise suggests that the correctness of cognitive task performance was not affected by the auditory environment. The absence of significant variance in accuracy, as per the GLMM results, aligns with the findings from our other statistical tests, painting a cohesive picture of the impact of noise conditions on cognitive task performance.

In summary, based on the GLMM results:

- **For Reaction Time:**
  - The nback variable (x_1) has a statistically significant positive effect.
  - This indicates that as nback increases, reaction time tends to increase.
- **For Correctness of Responses:**
  - The nback variable (x_1) does not have a statistically significant effect.
  - Other variables (x_2 and x_3) also do not significantly affect correctness.

The GLMM results provide a nuanced view of the relationship between noise conditions and cognitive performance, taking into account individual differences and potential interactions. The lack of significant effects in both reaction time and accuracy highlights the resilience or adaptability of cognitive processes in the face of varying auditory stimuli. These findings contribute to the broader understanding of cognitive flexibility and adaptability, suggesting that under certain conditions, human cognitive performance may be less susceptible to environmental changes than previously thought. Broader Context and Future Research Directions:

The use of GLMM in our study underscores the complexity of cognitive processes and the importance of considering individual variability in cognitive science research. Future studies may benefit from exploring other types of cognitive tasks, different noise conditions, or varying levels of task difficulty to further understand the dynamics between environmental stimuli and cognitive performance.

4 DISCUSSION

This subsection delves into the detailed interpretation of our study's findings. Contrary to our initial hypothesis, we observed no statistically significant differences in reaction times and accuracy across the different noise conditions (silence, pink, and brown noise). This outcome suggests that the type of background noise, within the parameters tested, did not significantly influence the cognitive performance of participants in 1-back and 2-back tasks.

4.1 Interpretation of Findings

The study explored the effect of brown and pink noise on cognitive performance, focusing on reaction time and accuracy. Descriptive statistics showed:

- **Reaction Time:** Mean reaction times were slightly higher in pink (0.96955s) and brown noise (0.97195s) compared to silence (0.95913s).
- **Accuracy Rate:** Accuracy was marginally higher in pink (0.969%) and brown noise (0.971%) environments compared to silence (0.959%).

These results suggest a minimal influence of pink and brown noise on cognitive performance, which is contrary to the hypothesis predicting a significant impact. The minor differences observed do not strongly support the initial assumption of a measurable effect of noise types on working memory.

The One-Way Repeated Measures ANOVA for reaction time showed no significant difference across conditions (*F Value*: 0.7039, *p-value*: 0.5046). Similarly, the Friedman Test for accuracy data also indicated no significant differences (*Statistic*: 0.3111, *p-value*: 0.8559).

Statistical Test	Assumption Checked	Outcome
Shapiro-Wilk Test	Normality	Normal distribution for reaction time; non-normal distribution for accuracy in pink and brown noise conditions.
One-Way Repeated Measures ANOVA	Normality	Applied to reaction time data, based on the normality assumption.
Friedman Test	Non-Normality	Employed for accuracy data, due to the non-normal distribution in pink and brown noise conditions.
Generalized Linear Mixed Models (GLMM)	Individual Variability, Non-Normality	Utilized for analysis of both individual accuracy and reaction time.

Table 9: Overview of Statistical Tests Used and Their Assumptions

The minimal differences in reaction times and accuracy rates across noise conditions suggest a potential resilience in cognitive performance to auditory stimuli. However, the small sample size and the nature of the cognitive tasks used may limit the generalizability of these findings. The tasks might not have been sensitive enough to detect subtle influences of auditory stimuli on cognitive speed or accuracy.

4.2 Study Limitations

- **Sample Size and Diversity**: The study's small sample size of 13 participants, mainly male students from the Technical University of Denmark, presents a limitation. This affects the generalizability of the results, as the sample may not represent the broader population effectively.

- **Linking to Literature Review and Hypotheses**:
  - The literature review highlighted varied effects of different types of noise on cognitive performance. However, our study's focus on brown and pink noise in a specific population (students) and using specific tasks (1-back and 2-back) may not fully capture the broader effects observed in previous research.
  - The hypothesis was formulated based on existing literature, which suggested potential cognitive impacts of different noise types. However, the constraints of our study design and sample might have limited our ability to detect these effects.

- **Methodological Constraints**:
  - *Experimental Design*: The within-subjects design reduces variability due to individual differences but may limit external validity. The study's tasks and setup, based on the reviewed literature, might have influenced the outcomes.
  - *Task Specificity*: The tasks used, while relevant, may not encompass the full spectrum of cognitive activities affected by auditory stimuli. This limitation is significant, especially considering the broad range of effects discussed in the literature review.

- **Statistical Power and Data Analysis**:
  - Constraints in resources and time led to a smaller participant pool than initially planned based on power analysis, possibly affecting the study's statistical power.
  - The reaction time data followed a normal distribution, but accuracy in brown and pink noise did not, necessitating different analytical approaches. This distinction aligns with the varied impacts of noise types discussed in the literature.

- **Environmental Consistency**: The study was conducted across different laptops and settings, potentially introducing minor variations, despite efforts to standardize the experimental environment. This factor is crucial, considering the controlled conditions used in many studies reviewed in the literature.

4.3 Contributions to Science

4.3.1 Refinement of Cognitive Noise Theory. This study contributes to the nuanced understanding of how different types of noise impact cognitive performance. Contrasting with some studies suggesting significant effects of ambient noise, these findings propose that pink and brown noise might not markedly affect certain cognitive tasks.

4.3.2 Insights into Cognitive Resilience. The research underscores the adaptability of human cognition to different auditory environments, challenging existing theories that suggest a straightforward negative or positive impact of ambient noise on cognitive performance.

4.3.3 Expansion of Research Scope. The study broadens the scope of research in cognitive science by focusing on pink and brown noise, which have been less explored compared to white noise. This contributes to a more comprehensive understanding of how various noise types affect cognitive functions.

4.4 Future Work

Based on the findings and limitations of the current study, future research directions are suggested:

- **Expanding Sample Size and Diversity:** Future studies should aim for a larger and more diverse participant pool to enhance the generalizability of findings. Incorporating participants from different demographics, academic backgrounds, and noise sensitivity levels would provide a more comprehensive understanding of the effects of ambient noise on cognitive performance.

- **Examining Different Noise Types and Levels:** Exploring a broader range of noise types, including white noise and real-world ambient sounds, could provide insights into their varying impacts on cognition. Additionally, examining the effects of different noise levels would contribute to understanding the dose-response relationship between noise exposure and cognitive performance.

- **Longitudinal Studies:** Conducting longitudinal research could reveal how prolonged exposure to different noise types affects cognitive functions over time, providing insights into chronic effects and potential adaptation mechanisms.

- **Incorporating a Wider Range of Cognitive Tasks:** Utilizing a variety of cognitive tasks, including those that measure other aspects of memory, attention, and executive function, would help in understanding the specific cognitive domains affected by noise exposure.

- **Technological Enhancements in Experimental Design:** Implementing advanced technologies like virtual reality could create more controlled and immersive environments for studying the effects of noise, allowing for the replication of real-world settings.

- **Investigating Individual Differences:** Further research should focus on individual differences in noise sensitivity, exploring how personal characteristics like age, gender, and pre-existing cognitive abilities influence the impact of noise on cognitive performance.

5 CONCLUSION

In our quest to understand how pink and brown noise affect working memory in students, we conducted an experiment rooted in extensive literature.The experiment, grounded in a comprehensive literature review, aimed to fill a gap in cognitive science research by focusing on less-studied noise types and their potential cognitive impacts. Despite rigorous methodological approaches and statistical analysis, our findings revealed no significant differences in reaction times and accuracy across noise conditions. These results challenge prevailing assumptions about the influence of ambient noise on cognitive performance, suggesting a level of cognitive resilience to auditory stimuli.

This study fills a research gap by focusing on less-studied noise types, thereby broadening the understanding of environmental influences on cognitive abilities. It also underscores the need for further research, particularly with larger and more diverse participant samples and a broader range of cognitive tasks. The limitations identified, such as sample size and methodological constraints, provide a roadmap for future studies to build upon our findings and explore the nuances of cognitive performance in varied auditory environments.

In conclusion, while our results do not show a direct effect of pink and brown noise on working memory, they open new avenues for research, emphasizing the complexity of environmental influences on cognitive processes.

REFERENCES

[1] Mohamad Awada, Burcin Becerik-Gerber, Gale Lucas, and Shawn Roll. 2022. Cognitive performance, creativity and stress levels of neurotypical young adults under different white noise levels. *Scientific Reports* 12, 1 (2022), 14566.
[2] Alan Baddeley. 1992. Working memory. *Science* 255, 5044 (1992), 556–559.
[3] Nicole Baum and Jasleen Chaddha. 2020. The Impact of Auditory White Noise on Cognitive Performance. *The Journal of Science and Medicine* (2020), 1–15. Issue 3(Special Issue). https://doi.org/10.37714/josam.v3i0.82
[4] Andrew RA Conway, Michael J Kane, Michael F Bunting, D Zach Hambrick, Oliver Wilhelm, and Randall W Engle. 2005. Working memory span tasks: A methodological review and users guide. *Psychonomic bulletin & review* 12 (2005), 769–786.
[5] Amy Garrett, Ryan Kelly, Rowena Gomez, Jennifer Keller, Alan F Schatzberg, and Allan L Reiss. 2011. Aberrant brain activation during a working memory task in psychotic major depression. *American Journal of Psychiatry* 168, 2 (2011), 173–182.
[6] Kejing Guo, Yanzhuo Wu, and Hairu Zhang. 2022. The Effects of Color Noises on Attention. *Proceedings of the 2022 International Conference on Science Education and Art Appreciation (SEAA 2022)*, 576–583. https://doi.org/10.2991/978-2-494069-05-3_71
[7] Suzannah K Helps, Susan Bamford, Edmund J S Sonuga-Barke, and Göran B W Söderlund. 2014. Different Effects of Adding White Noise on Cognitive Performance of Sub-, Normal and Super-Attentive School Children. *PLOS ONE* 9 (11 2014), e112768–. Issue 11. https://doi.org/10.1371/journal.pone.0112768
[8] Nora A. Herweg and Nico Bunzeck. 2015. Differential effects of white noise in cognitive and perceptual tasks. *Frontiers in Psychology* 6 (2015). https://doi.org/10.3389/fpsyg.2015.01639
[9] Shih-Yi Lu, Yuan Huang, and Kuei-Yi Lin. 2020. Spectral Content (colour) of Noise Exposure Affects Work Efficiency. *Noise Health* 22 (2020), 19 – 27. https://api.semanticscholar.org/CorpusID:227181019
[10] D. Singh, A. Jain, D. Jain, and V. Goyel. 2023. Effect of white, brown and pink noises on anxious pediatric dental patients. *CARDIOMETRY* (2 2023), 1252–1258. Issue 25. https://doi.org/10.18137/cardiometry.2022.25.12521258

A INFORMATION SHEET

Appendix I: Participant information letter

Project title: Investigating the effect of coloured noise on human cognition.

Dear participant,

In this letter we will inform you about the purpose of the experiment and the procedures. It is important that you read the letter carefully. If you have any questions, do not hesitate to contact us (Weronika Strączek s222754@dtu.dk, Siddhanta Gupta s210217@dtu.dk, Shakir Maytham Shaker s230553@dtu.dk) for clarification.

Your rights as participant

Your participation in this experiment is voluntarily. This means that you can leave the experiment at any point in time without consequences for yourself, and without having to give a reason. We will ask you for an informed consent to participate after you have been informed about the experiment.

Purpose of the research project

In this research project we are investigating how different noise types affect short-term memory and sustaining attention. You will perform two tasks in three different noise environments (wearing headphones with pink noise, brown noise or silence) while we will record the accuracy of your answers.

Data storage and handling

All recorded data (test answers, gender and age) will be anonymized, that is, stored not in connection with your name, address, CPR number, or any other information that would allow to identify you. For this, your data will be stored under a handle (key code). Personal information such as your name and email address will not be stored.

Coloured noise

During the test you will hear different types of coloured noise from the headphones, namely pink and brown noise. Both are like white noise, as in they all sound like TV static or hum of a waterfall. The main difference between them is the amount of energy in different frequencies. While executing the tasks, the noise will be playing as a background noise, and you do not need to pay any special attention to it. Between each set of tasks there will be a period of rest.

If you have any questions about the experiment, the methods or your safety and rights, do not hesitate to contact us per email (written in the first paragraph) or telephone (50350385, 91875372 50414287).

With kind regards,
Weronika & Siddhanta & Shakir

B CONSENT FORM

Informed consent form

Experiment: Investigating the effect of coloured noise on human cognition.
Investigator: ____

I confirm that:

- I was satisfactorily informed about the study concerned both verbally and in writing by means of the subject information letter.
- I have had the opportunity to put forward questions regarding the study and that these questions have been answered satisfactorily
- I have carefully considered my participation in the experiment
- I participate of my own free will

I agree that:

- My data will be acquired and stored for scientific purposes as mentioned in the subject information letter in anonymized form
- My anonymized data (detached from any personal information) will be stored and may be shared openly through the DTU data repository.

I understand that:

- My participation is voluntary, and I have the right to withdraw from the experiment at any time without having to give a reason
- My privacy is protected according to Danish law and European guidelines (GDPR; EU 2016/679)

I give my consent to take part in this experiment:

____
Full name

______________________________________ ______________________________________
Date, Place Signature

C EXPERIMENT CODE

Experiment has been prepared in Psychopy Builder and can be found in this repository under folder "psychopy_new":
https://github.com/shakirmshaker/CognitiveScience