



UNIVERSIDAD
DE LOS ANDES
VENEZUELA

PROYECTO DE GRADO

PRESENTADO ANTE LA ILUSTRE UNIVERSIDAD DE LOS ANDES COMO
REQUISITO PARA OBTENER EL TÍTULO DE INGENIERO DE SISTEMAS

**DISEÑO DE UN MODELO PARA PREDECIR EL RENDIMIENTO
DE LA MEMORIA DE TRABAJO EN ESTUDIANTES
UNIVERSITARIOS SEGÚN LOS NIVELES DE ATENCIÓN Y
ESTRÉS MENTAL**

POR:

BR. DANIELLA MARYUSKA ZAMBRANO ROJAS

TUTOR: PROF. JESÚS A. PÉREZ A.

MÉRIDA, JUNIO 2025



**PROYECTO DE GRADO
CALIFICACIÓN FINAL**

Título del Proyecto de Grado: "Diseño de un Modelo para predecir el rendimiento de la memoria de trabajo en Estudiantes Universitarios según los niveles de Atención y Estrés Mental".

Bachiller: **Zambrano Rojas Daniella Maryuska** C.I 26.558.153

1) Calificación del (de la) Profesor(a) Tutor(a):	20	(20%)
2) Calificación del manuscrito final:	20	(40%)
3) Calificación de la defensa oral:	20	(40%)
Calificación final:	20	(puntos)

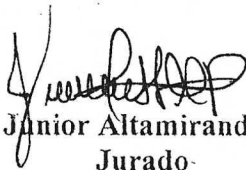
Los suscritos miembros del Jurado asignan como calificación final del Proyecto de Grado la nota de:

Observación: Los Miembros del jurado calificador sugerimos que el presente trabajo sea accesible a la comunidad universitaria en general a través de la publicación de un artículo científico, donde se destaquen los principales resultados de este trabajo. En consecuencia, le otorgamos **MENCIÓN PUBLICACIÓN**

20 (puntos)
Números


Prof. Jesús Pérez Angulo
Tutor

Veinte (puntos)
Letras


Prof. Junior Altamiranda
Jurado


Prof. Alejandro Mujica
Jurado

Mérida, 23 de junio de 2025.



Resumen

Este estudio presenta un modelo computacional diseñado para predecir el rendimiento de la memoria de trabajo a partir de variables fisiológicas (relacionadas a la atención y el estrés) registradas mediante señales de electroencefalograma (EEG) captadas con un dispositivo de bajo costo (NeuroSky MindWave). Esta investigación incluye recolección de datos y comparación de tres modelos: modelo de Interferencia de Klaus Oberauer (modelo base), modelo base extendido con atención y estrés (modelo neuro 2), y modelo base extendido con atención (modelo neuro 1). Los datos fueron recolectados bajo condiciones experimentales controladas. El modelo base alcanzó la mayor exactitud general (78,31 %) y una alta sensibilidad (92,19 %), aunque con baja especificidad (47,23 %). El modelo neuro 2 logró un mejor equilibrio entre sensibilidad (77,04 %) y especificidad (68,96 %), mientras que el modelo neuro 1 alcanzó la mayor especificidad (76,80 %), aunque sacrificando sensibilidad (67,86 %). Estos resultados confirman el valor de la atención y el estrés para predecir el rendimiento de la memoria, lo cual tiene gran potencial para construir sistemas adaptativos en entornos educativos personalizados.

Palabras clave: Modelo predictivo, Memoria de trabajo, Niveles de atención, Estrés mental.

Índice

Índice de Figuras.....	9
Índice de Tablas.....	9
Introducción.....	10
Capítulo I - Contextualización.....	12
1.1. Antecedentes.....	12
1.1.1. Procesos Cognitivos.....	12
1.1.2. Memoria de Trabajo.....	13
1.1.3. Memoria de Trabajo, Atención y Estrés.....	14
1.1.4. Modelos Predictivos.....	15
1.1.5. Medición de Actividad Cerebral.....	16
1.2. Planteamiento del Problema.....	17
1.3. Objetivos.....	18
1.3.1. Objetivo General.....	18
1.3.2. Objetivos Específicos.....	18
1.4. Justificación.....	19
1.5. Alcance.....	20
1.6. Limitaciones.....	20
Capítulo II - Marco Teórico.....	22
2.1. Fundamentos de la Memoria de Trabajo.....	22
2.2. Modelos cognitivos de memoria de trabajo.....	23
2.2.1. El Modelo Multicomponente de Baddeley y Hitch.....	23
2.2.2. El Modelo de Memoria Activada y Foco Atencional de Cowan.....	24
2.2.3. Modelo de Interferencia de la Memoria de Trabajo de Oberauer.....	25
2.2.3.1. Mecanismos de Interferencia.....	26

2.3. Modelo de Interferencia para la Memoria de Trabajo Visual y Verbal.....	27
2.3.1. Formulaciones Matemáticas Clave y Parámetros.....	27
2.3.2. Validación y Relevancia del Modelo de Oberauer.....	34
2.4. Neurociencia Cognitiva: Atención y Estrés Mental.....	34
2.4.1. Atención: Constructo y Modelos Neurocognitivos.....	35
2.4.2. Estrés Mental y su Impacto en los Procesos Cognitivos.....	35
2.5. Electroencefalografía (EEG) y Medición de Señales Cerebrales.....	36
2.5.1. Bandas de Frecuencia EEG y Estados Mentales Asociados.....	37
2.5.1.1. Alpha (α) (8-13 Hz).....	38
2.5.1.2. Beta (β) (13-30 Hz).....	38
2.5.1.3. Theta (θ) (4-8 Hz).....	39
2.5.2. Correlatos EEG de la Atención.....	39
2.5.3. Correlatos EEG del Estrés Mental.....	40
2.6. Dispositivo Neurosky Mindwave.....	41
2.6.1. Casos de Uso en Investigación Cognitiva.....	43
2.7. Construcción y Análisis de Datasets en Neurociencia Cognitiva.....	44
2.7.1. Buenas prácticas para la recolección de datos experimentales.....	44
2.7.2. Procesamiento de Señales y Estructuración del Dataset.....	45
2.7.3. Consideraciones Éticas y Protocolos de Consentimiento.....	46
2.7.4. Herramientas de captura y análisis.....	46
2.8. Modelado Predictivo en Ciencias Cognitivas.....	47
2.8.1. Conceptos Básicos de Modelado Matemático.....	47
2.8.2. Ajuste de Modelos con Base en Datos Empíricos.....	48
2.8.3. Comparación de Modelos: Métricas de Evaluación y Validación.....	48

Capítulo III - Diseño e Implementación.....	52
3.1. Diseño Experimental.....	52
3.1.1. Tipo de Estudio.....	52
3.1.2. Participantes.....	52
3.1.2.1. Criterios de Inclusión.....	53
3.1.2.2. Criterios de Exclusión.....	53
3.1.3. Variables.....	54
3.1.3.1. Variables Predictoras (Independientes).....	54
3.1.3.2. Variable de Resultado (Dependiente).....	54
3.2. Consideraciones Éticas.....	55
3.2.1. Consentimiento Informado.....	55
3.2.2. Privacidad y Anonimización de Datos.....	56
3.2.3. Manejo de Datos EEG.....	56
3.3. Recolección de Datos.....	57
3.3.1. Adaptación de la Tarea de Memoria Letterwheel.....	57
3.3.2. Integración del Mindwave Neurosky.....	61
3.3.2.1. Procesamiento de la Señal EEG.....	62
3.3.3. Implementación de la Tarea de Memoria Letterwheel.....	63
a. Llegada y Explicación Inicial.....	63
b. Revisión del Consentimiento Informado.....	64
c. Colocación del Dispositivo.....	64
d. Verificación de la Señal.....	64
e. Encuesta Inicial.....	64
f. Explicación de la Tarea Letterwheel.....	64
g. Ejecución de la Tarea.....	65

h. Encuesta Final.....	65
i. Retirada del Dispositivo y Agradecimiento.....	66
3.3.4. Esfuerzos de Estandarización.....	66
3.4. Construcción del Dataset.....	66
3.4.1. Generación de Archivos de Registro.....	67
3.4.2. Conversión de Datos: De Registros Python a Formato RData.....	67
3.4.3. Análisis Exploratorio de Datos para Identificación de Patrones.....	69
3.5. Desarrollo del Modelo Matemático Predictivo.....	71
3.5.1. Justificación para la Modificación del Modelo.....	71
3.5.2. Formulación Matemática del Modelo Modificado.....	71
3.5.2.1. Presentación de las Ecuaciones.....	72
3.5.2.2. Variables y parámetros.....	74
3.5.2.3. Modulación Fisiológica de Parámetros.....	74
3.6. Proceso de Entrenamiento.....	76
3.8 Diagrama General del Proceso de Diseño e Implementación del Estudio.....	77
Capítulo IV - Resultados.....	80
4.1. Participantes.....	80
4.1.1. Encuesta Inicial y Final.....	80
4.1.1.1. Niveles de Fatiga Pre sesión y Pos sesión.....	81
4.1.1.2. Niveles de Ánimo Pre sesión y Pos sesión.....	83
4.2. Dataset.....	85
4.2.1. Análisis Exploratorio de Datos.....	86
4.3. Rendimiento.....	89
4.3.1. Matriz de Confusión de Cada Modelo.....	90
4.3.1.1. Modelo Base.....	90

4.3.1.2. Modelo Neuro 2.....	91
4.3.1.3. Modelo Neuro 1.....	92
4.3.2. Tabla Comparativa de Modelos.....	93
4.3.3. Gráfico de Barras Verticales: Exactitud por Sujeto.....	96
Capítulo V - Conclusiones y Recomendaciones.....	98
5.1. Conclusiones.....	98
5.2. Recomendaciones.....	100
5.3. Trabajos Futuros.....	101
Referencias Bibliográficas.....	103
Anexos.....	113
Anexo 1: Carta de Consentimiento Informado.....	113
Anexo 2: Encuesta Inicial y Final.....	116

Índice de Figuras

Figura 2.5: Posiciones del lóbulo frontal y la ubicación del Fp1.....	39
Figura 2.5.1: Bandas de frecuencia alpha beta y theta.....	40
Figura 2.6: Neurosky Mindwave 1.....	45
Figura 3.8: Diagrama general del proceso de diseño e implementación del estudio...81	
Figura 4.1.1.1: Cambio de fatiga por participante y por sesión.....	83
Figura 4.1.1.2: Cambio de ánimo por participante y por sesión.....	86
Figura 4.2.1: Distribuciones de las variables atención (izq) y estrés (der).....	89
Figura 4.2.2 Correlación entre las variables de memoria, atención y estrés.....	91
Figura 4.3.1.1: Matriz de confusión del Modelo Base.....	93
Figura 4.3.1.2: Matriz de confusión del Modelo Neuro 2.....	94
Figura 4.3.1.3: Matriz de confusión del Modelo Neuro 1.....	95
Figura 4.3.3: Gráfica de Barras Verticales: Exactitud por Sujeto.....	99

Índice de Tablas

Tabla 2.8.3: Matriz de confusión (Clase 0 positiva).....	51
Tabla 4.3.2.1: Comparativa de las métricas derivadas.....	95
Tabla 4.3.2.2: Métricas resaltantes derivadas de la matriz de confusión.....	96

Introducción

Las funciones ejecutivas, en particular la memoria de trabajo, son esenciales para el aprendizaje significativo y la resolución de problemas, pero su rendimiento puede verse alterado significativamente por fluctuaciones en el estado mental del individuo. En los entornos educativos actuales, caracterizados por una alta demanda cognitiva, resulta fundamental comprender cómo los estados mentales influyen en el rendimiento académico, así se plantea un desafío: ¿Cómo adaptar modelos computacionales para predecir el rendimiento de la memoria de trabajo integrando variables como la atención y el estrés?

Es así, como esta investigación tiene la premisa de comprender cómo la atención y el estrés, medidos objetivamente a través de señales EEG, se relacionan con el desempeño de la memoria de trabajo, y utilizar ese conocimiento para predecir dicho rendimiento. La presente investigación propone la extensión de un modelo matemático capaz de predecir el rendimiento de la memoria de trabajo a partir de señales recogidas con el dispositivo NeuroSky MindWave, una tecnología de electroencefalografía portátil y accesible.

El estudio incluyó la integración de datos experimentales tomados en tiempo real mientras los participantes resolvían tareas tipo Letterwheel, adaptadas para estimular la carga y manipulación de información visual-espacial. Mientras resolvían dichas tareas se recolectaba una señal EEG que fue procesada para extraer métricas como la actividad en bandas alpha, beta y theta, y que luego fueron transformadas en índices cuantitativos de atención y estrés, según criterios descritos en la literatura.

También se aplicó un análisis exploratorio de datos para comprender la distribución y correlación de los niveles de atención y estrés, revelando patrones significativos que respaldan la relevancia del enfoque adoptado para la construcción del modelo. Finalmente, se realizó una comparación entre tres modelos para evaluar si la

inclusión de variables cognitivas derivadas del EEG mejoran la capacidad predictiva, en las que, las métricas evaluadas incluyeron exactitud, sensibilidad, especificidad, F1-score y la interpretación de matrices de confusión y su recall.

Así mismo, a lo largo de la investigación, se abordaron desafíos tanto metodológicos como computacionales: desde la depuración de señales EEG, el diseño del protocolo experimental y la elaboración de encuestas pre y pos tarea, hasta la construcción del dataset final, la normalización de variables y la aplicación de algoritmos de clasificación.

Esta investigación se estructura en cinco capítulos:

El Capítulo I contextualiza el problema de investigación, justifica su relevancia académica y social, y define los objetivos específicos.

El Capítulo II compila la base teórica que sustenta el modelo propuesto, incluyendo nociones de neurociencia cognitiva, análisis EEG, modelos computacionales de memoria de trabajo, y escalas de medición como la de Likert.

El Capítulo III describe el diseño experimental, los criterios de selección de los participantes, el uso del dispositivo NeuroSky, la integración de encuestas para obtener variables como el nivel de fatiga y el estado de ánimo, la recolección de los datos y el diseño del modelo matemático modificado con sus ecuaciones.

El Capítulo IV expone en detalle el análisis de los datos, presentando métricas de evaluación como resultados para comparar los modelos.

El Capítulo V sintetiza las conclusiones más relevantes, presenta recomendaciones prácticas y traza líneas futuras de investigación, orientadas a convertir estos hallazgos en herramientas funcionales dentro del entorno educativo.

Capítulo I - Contextualización

1.1. Antecedentes

1.1.1. Procesos Cognitivos

En los humanos, al igual que en otras especies, se distinguen varios niveles de funcionamiento cerebral; por ejemplo, según Jackson [1], existen tres: el inferior o medular, el nivel medio y el nivel superior. La característica distintiva de los seres humanos se encuentra en los procesos cognitivos, desarrollados en el nivel superior, particularmente en la corteza cerebral, donde se controlan funciones corticales complejas [2]. La corteza cerebral humana está más desarrollada para manejar procesos como el pensamiento abstracto, la planificación y la toma de decisiones conscientes, reflejando la evolución de habilidades cognitivas avanzadas [3].

La cognición humana, se ocupa principalmente, de procesos conscientes y respuestas voluntarias, aunque existe cierta superposición con procesos inconscientes e involuntarios. Entre las áreas específicas de la cognición se encuentran la atención, la percepción, la memoria, el lenguaje, y el pensamiento, entre los cuales la memoria figura como un componente crucial del sistema cognitivo humano [4].

La memoria humana, además, se compone de varios tipos, con funciones específicas y complementarias, donde una de las principales clasificaciones que se distinguen son la memoria a corto plazo y la memoria a largo plazo [5]. La memoria a corto plazo es responsable de la retención temporal de información durante un breve periodo, típicamente segundos o minutos, y es esencial para tareas inmediatas, como recordar un número de teléfono o seguir una conversación [4]. Por otro lado, la memoria a largo plazo almacena información durante periodos prolongados, desde días hasta décadas, y se subdivide en memoria explícita (o declarativa) e implícita (o no declarativa) [6].

En este contexto, la memoria de trabajo, que es una forma de memoria a corto plazo, juega un papel crucial al permitir la retención y manipulación temporal de información necesaria para tareas cognitivas complejas, como el aprendizaje, la resolución de problemas y la planificación [4].

1.1.2. Memoria de Trabajo

La memoria de trabajo (*WM, por sus siglas en inglés Working Memory*) es un sistema cognitivo fundamental. Según Eysenck [7], es de capacidad limitada, pero también es la responsable del almacenamiento temporal y la manipulación activa de la información necesaria para la ejecución de tareas cognitivas. De esta manera, la WM funciona como un "espacio de trabajo mental" o un "bloc de notas" que permite al cerebro manejar y procesar varias piezas de información simultáneamente [8]. Este sistema es crucial para una amplia gama de actividades intelectuales, incluyendo el aprendizaje, la lectura y comprensión, el razonamiento, la resolución de problemas y la toma de decisiones [7]. En el contexto universitario, donde las demandas cognitivas son elevadas y constantes, un rendimiento eficiente de la WM es un predictor significativo del éxito académico. La capacidad de los estudiantes para mantener y manipular información relevante, seguir instrucciones y organizar sus pensamientos depende en gran medida de la integridad de su WM.

Es por ello que, relacionando estos términos a la educación como lo señala Parra [9], la memoria y la atención son dos procesos cognitivos superiores implicados en las tareas educativas y en las tareas de la vida cotidiana del alumnado de educación infantil, primaria, secundaria, y en el ámbito universitario, ya que investigaciones como las de Ramos [10] han mostrado que tanto la memoria de trabajo como la atención son cruciales para el éxito académico. Estas habilidades cognitivas son esenciales para la gestión de la carga de trabajo académico y el procesamiento de información compleja, por lo tanto, los profesionales de la educación deben conocerlos para tener una

comprensión completa de ellos facilitando la valoración e intervención, dirigidas a la mejora del rendimiento académico.

Así mismo, entre las habilidades cognitivas que tienen un impacto significativo en el rendimiento académico también destaca el estrés mental. Según Sandi [11], el efecto del estrés sobre la cognición puede variar dependiendo de la intensidad, origen y duración del estrés, así como del tipo de función cognitiva en estudio, por ejemplo, mientras que el estrés leve puede facilitar funciones cognitivas simples, el estrés agudo o crónico puede deteriorar la formación de memorias explícitas y otras tareas que requieren razonamiento, afectando particularmente las funciones relacionadas con el hipocampo y la corteza prefrontal.

1.1.3. Memoria de Trabajo, Atención y Estrés

El rendimiento de la WM no es estático, está intrínsecamente modulado por estados cognitivos fluctuantes como la atención y estados emocionales como el estrés [12]. La atención es un componente crucial para la selección, el mantenimiento y la manipulación de la información dentro de la WM. Permite priorizar la información relevante y protegerla de la interferencia de estímulos irrelevantes, por otro lado, el estrés mental puede tener efectos perjudiciales sobre la WM, por ejemplo, niveles elevados de estrés pueden agotar los recursos cognitivos disponibles, interferir con el control ejecutivo necesario para las tareas de WM, o introducir "ruido" en las representaciones mentales. La interacción entre estos tres constructos es compleja y bidireccional, lo que subraya la necesidad de un modelo que pueda capturar estas dinámicas.

La predicción del rendimiento de la WM en estudiantes universitarios, considerando estos factores, tiene implicaciones directas para la neurociencia educativa, ya que los estudiantes universitarios constituyen una población sometida a una carga cognitiva y emocional considerable. Si un modelo puede predecir las fluctuaciones en el

rendimiento de la memoria de trabajo basándose en estados medibles como la atención y el estrés, podría conducir al desarrollo de intervenciones personalizadas o estrategias de estudio más eficaces.

1.1.4. Modelos Predictivos

Los modelos predictivos son herramientas estadísticas y computacionales diseñadas para estimar la probabilidad de que un fenómeno ocurra, a partir de la relación entre variables observadas en una muestra previa [13]. Los modelos predictivos sirven para estimar cómo se comportarán nuevas observaciones a partir de patrones aprendidos en datos previos, lo cual es fundamental en diversas áreas, por ejemplo, marketing o detección de fraudes [13, 14]. En el campo de la cognición, estos modelos predictivos se emplean para anticipar variables mentales como rendimiento atencional, o en memoria usando datos observables como registros de actividad cerebral o respuesta conductual.

Entre los enfoques aplicados a la memoria destacan, inicialmente, modelos probabilísticos de tipo Markoviano, como el *Prediction by Partial Matching* reducido por decaimiento (*PPM-Decay*). Este modelo extiende un clásico algorítmico de predicción de secuencias (n-gramas) incorporando un decaimiento en el peso de eventos pasados [15]. Este atributo lo hace adecuado para simular cómo el olvido afecta patrones de memoria temporal, integrando así una dimensión predictiva en entornos auditivos o lingüísticos, con resultados que mejoran la fidelidad frente a la versión sin decaimiento.

Por otra parte, los modelos basados en arquitecturas cognitivas ofrecen predicción detallada a nivel de ensayo, como por ejemplo, el modelo SAC (*source-of-activation confusion*), desarrollado por Reder y colaboradores [16], utiliza un sistema de nodos semánticos, perceptuales y episódicos interconectados. En este modelo, la probabilidad de recuperar un recuerdo depende de la activación propagada desde el contexto y de la confusión con eventos anteriores, una forma explícita de predecir errores de memoria.

Seguidamente, una teoría sobre la arquitectura y los mecanismos de la memoria de trabajo sobre modelos predictivos es el modelo de interferencia propuesto por Klaus Oberauer [17], el cual describe cómo se gestionan y mantienen activamente múltiples elementos de información bajo condiciones de carga cognitiva, centrándose particularmente en cómo se accede a la información y cómo la interferencia limita su capacidad y precisión.

1.1.5. Medición de Actividad Cerebral

El desarrollo de tecnologías de medición no invasiva como la electroencefalografía (*EEG*) ha abierto la posibilidad de monitorear en tiempo real estados mentales asociados a procesos atencionales y emocionales, mediante el análisis de la actividad en bandas de frecuencia específicas (delta, theta, alpha, beta, gamma) [18]. En particular, dispositivos de bajo costo como el NeuroSky MindWave han facilitado el acceso a este tipo de datos en contextos educativos reales.

Un ejemplo relevante se encuentra en el estudio de Ni [19], donde se utilizó el dispositivo MindWave para evaluar los niveles de atención de estudiantes universitarios durante sesiones de aprendizaje multimedia. En este experimento, el dispositivo fue colocado en la región frontal (*Fp1*) y se capturó un índice de atención derivado de las señales EEG mientras los participantes eran expuestos a materiales educativos en distintos formatos: texto plano, texto con imágenes y contenido audiovisual. Los resultados mostraron que el índice de atención variaba significativamente según el tipo de contenido presentado, siendo el video el estímulo que generó mayores niveles de atención sostenida.

Este tipo de aplicaciones demuestra la viabilidad del uso del MindWave para cuantificar variables cognitivas en entornos educativos reales y respalda la propuesta de este proyecto: integrar marcadores fisiológicos de atención y estrés, obtenidos de señales

EEG, con modelos computacionales que permitan predecir de manera más precisa el rendimiento de la memoria de trabajo.

1.2. Planteamiento del Problema

En el contexto de la educación universitaria, el rendimiento académico de los estudiantes es un aspecto crucial que puede verse afectado por diversos factores cognitivos y emocionales. Las investigaciones de Gathercole [20] y Engle [21] han mostrado interés por conocer los aspectos determinantes de la memoria de trabajo y las habilidades académicas, subrayando la necesidad de investigar estas relaciones para desarrollar estrategias efectivas que mejoren el rendimiento académico.

Existe una considerable variabilidad interindividual en el rendimiento de la memoria de trabajo entre los estudiantes universitarios, la cual no solo se debe a diferencias basadas en la capacidad de la misma, sino también a la influencia de diversos factores cognitivos y afectivos que modulan su eficiencia en tiempo real. Entre estos factores, la atención y el estrés mental emergen como moduladores particularmente significativos [2], y como mecanismos que dirigen los recursos cognitivos hacia información relevante [4].

Además, el estrés mental, una respuesta fisiológica y psicológica a demandas percibidas como excesivas, ha demostrado tener efectos perjudiciales sobre las funciones cognitivas, incluyendo la memoria de trabajo. Específicamente, el estrés puede comprometer la capacidad para mantener información y resistir la interferencia, llevando a una disminución en el rendimiento académico y cognitivo general [5].

A pesar de la existencia de estudios que abordan la memoria de trabajo, una revisión de los resultados de Leblanc [22], muestra que la mayoría de los trabajos referentes a procesos cognitivos utilizan mecanismos de evaluación de rendimiento cognitivo y académico basados en pruebas estandarizadas que no siempre reflejan la complejidad de las interacciones entre diferentes factores cognitivos y emocionales.

Actualmente y según nuestro conocimiento, no existen modelos matemáticos predictivos que integren los niveles de atención, estrés mental, y memoria de trabajo. Un modelo con esas características permitiría no sólo una mejor comprensión de las variables, sino también el desarrollo de intervenciones personalizadas que puedan mejorar significativamente el rendimiento académico, en consecuencia, este Proyecto de Grado se centrará en la propuesta de un modelo predictivo que relacione los niveles de atención y estrés con la memoria de trabajo en estudiantes universitarios.

1.3. Objetivos

1.3.1. Objetivo General

- Diseñar un modelo para predecir el rendimiento de la memoria de trabajo en estudiantes universitarios según los niveles de atención y estrés mental.

1.3.2. Objetivos Específicos

- Estudiar el Modelo de Interferencia de predicción de memoria planteado en las investigaciones de Klaus Oberauer.
- Construir un dataset para el Modelo de Interferencia utilizando el Mindwave 1 de la empresa Neurosky.
- Analizar los datos obtenidos sobre los niveles de atención, estrés mental, y memoria de trabajo.
- Proponer un Modelo de Interferencia Extendido que relacione los niveles de atención, estrés mental y memoria de trabajo.
- Realizar pruebas de rendimiento del modelo propuesto utilizando el dataset construido.

1.4. Justificación

La comprensión de los factores que influyen en el rendimiento de la Memoria de Trabajo en estudiantes universitarios es esencial para desarrollar estrategias de apoyo efectivas. En este sentido, la creación de modelos predictivos, que integren medidas fisiológicas objetivas, es una vía prometedora.

Aunado a esto, la electroencefalografía (EEG) es una técnica no invasiva que permite medir la actividad eléctrica cerebral asociada con estados cognitivos como la atención y el estrés [23]. Así, el uso de dispositivos EEG accesibles y de bajo costo, como el Neurosky Mindwave, democratiza la posibilidad de realizar investigaciones en neurociencia cognitiva, aplicada en entornos locales.

En este sentido, esta investigación propone una innovación al combinar datos EEG obtenidos con el Neurosky Mindwave, específicamente del canal FP1 (el cual es relevante para funciones ejecutivas [23]), con los constructos teóricos del modelo de interferencia de la memoria de trabajo de Klaus Oberauer [17]. La integración de estas dos fuentes de información (fisiológica y teórica) permitirá desarrollar un modelo matemático predictivo del rendimiento de la memoria de trabajo en función de los niveles de atención y estrés mental.

Dicho modelo no solo contribuiría al conocimiento científico sobre la interacción entre la memoria de trabajo, la atención y el estrés, sino que también podría sentar las bases para futuras herramientas de diagnóstico temprano o intervenciones personalizadas destinadas a optimizar el rendimiento cognitivo de los estudiantes.

1.5. Alcance

Esta investigación tiene como alcance el diseño y evaluación de un modelo predictivo del rendimiento de la memoria de trabajo (WM) en estudiantes universitarios activos, utilizando como variables predictoras las señales EEG relacionadas con los niveles de atención y estrés que serán recolectadas mediante el dispositivo NeuroSky MindWave, el cual registra la actividad cerebral desde el canal FP1, mientras los participantes realizan tareas de WM en un entorno controlado y se espera que este modelo, aunque exploratorio, brinde información útil para aplicaciones futuras.

El estudio se delimita a una muestra específica: estudiantes universitarios activos de la Universidad de Los Andes, específicamente de la carrera de Ingeniería de Sistemas Núcleo Mérida, con edades comprendidas entre 18 y 28 años, sin antecedentes de trastornos neurológicos o psiquiátricos y que no hayan consumido estimulantes en las dos horas previas a la prueba.

La tarea de memoria será presentada digitalmente, específicamente, la tarea diseñada por el profesor Klaus Oberauer, en su Experimento denominado “Letterwheel” [17], para evaluar la capacidad y el rendimiento de la Memoria de Trabajo.

1.6. Limitaciones

La principal limitación metodológica de este estudio radica en la disponibilidad de un único dispositivo EEG NeuroSky MindWave para la adquisición de datos. Esta restricción impone limitaciones logísticas importantes, como la imposibilidad de realizar sesiones paralelas o aumentar el tamaño muestral en un corto período de tiempo.

Otra limitación relevante se relaciona con la población seleccionada para el estudio, puesto que, la muestra estará compuesta exclusivamente por estudiantes universitarios de una única institución de educación superior, lo que restringe la

generalización de los hallazgos a otras poblaciones con características demográficas, académicas o culturales distintas.

Además, el tamaño de la muestra se limitará a 22 participantes, tanto por restricciones logísticas, como por el objetivo de mantener la coherencia con el experimento original de Klaus Oberauer, cuya base de datos también se fundamenta en una muestra de 22 personas.

Capítulo II - Marco Teórico

2.1. Fundamentos de la Memoria de Trabajo

La memoria de trabajo (*WM*) se rige como un constructo central en la psicología cognitiva, trascendiendo la noción de un simple almacén pasivo de información. Se define como un sistema cognitivo de capacidad limitada, responsable no solo del mantenimiento temporal de la información, sino, de manera crucial, de su activa manipulación y transformación para la consecución de tareas cognitivas complejas [24].

Estas tareas abarcan procesos de alto nivel como el razonamiento, la comprensión del lenguaje y el aprendizaje, que son fundamentales en el contexto del rendimiento académico universitario [25]. Asimismo, la psicología cognitiva, en su afán por desentrañar cómo el cerebro percibe, organiza, representa y utiliza la información, concede a la WM un papel protagónico. Cuando la información a mantener o la tarea a realizar alcanza una complejidad tal que satura la capacidad del sistema cognitivo, la memoria a corto plazo (*STM por sus siglas en inglés short-term memory*) resulta insuficiente, y es entonces cuando la WM despliega su funcionalidad esencial.

La evolución conceptual de la memoria de trabajo desde una perspectiva de mero almacenamiento temporal hacia un sistema dinámico y funcional es un hito en la neurociencia cognitiva. Este cambio de paradigma es fundamental porque posiciona a la WM [26] como un "espacio de trabajo mental" activo, un escenario donde la información no solo se sostiene, sino que se procesa, se transforma y se integra para guiar el pensamiento y la acción. Así, a diferencia de la STM, cuyo criterio definitorio es primordialmente temporal, la WM se caracteriza por un criterio funcional; es decir, su relevancia radica en lo que hace con la información, más que en cuánto tiempo la retiene.

En lugar de simplemente "retener", la WM "opera" sobre la información. Esto implica procesos sofisticados de mantenimiento activo, actualización constante de

representaciones y manipulación flexible de los contenidos mentales. Para este proyecto de investigación en donde se busca predecir el rendimiento de la WM, esta naturaleza activa es el núcleo de la cuestión, ya que el rendimiento no se reduce a la mera capacidad de almacenamiento, sino que refleja la eficiencia del procesamiento bajo condiciones variables, como los niveles de atención y estrés.

2.2. Modelos cognitivos de memoria de trabajo

La comprensión de la memoria de trabajo ha sido profundamente enriquecida por una variedad de modelos teóricos que buscan describir su estructura, mecanismos y limitaciones. Entre los enfoques más influyentes se encuentran el modelo multicomponente de Baddeley y Hitch [27] (Ejecutivo Central, Bucle Fonológico, Agenda Visoespacial, Buffer Episódico), el modelo de memoria activada y foco atencional de Cowan [28], y el modelo de interferencia propuesto por Klaus Oberauer [29].

A continuación, se presentará una breve descripción de cada uno, con énfasis especial en el modelo de Oberauer, el cual sirve como base para el desarrollo e implementación del modelo propuesto en esta investigación.

2.2.1. El Modelo Multicomponente de Baddeley y Hitch

La introducción del modelo multicomponente de Baddeley y Hitch [27], supuso un avance significativo al proponer una WM no unitaria, sino compuesta por subsistemas funcionalmente especializados. Esta modularidad permite investigar cómo se procesan diferentes tipos de información (verbal versus visoespacial) y, crucialmente, cómo la atención, a través del ejecutivo central, gestiona y coordina estos procesos, la distinción entre el bucle fonológico y la agenda visoespacial, por ejemplo, ayuda a explicar por qué

los seres humanos pueden realizar simultáneamente tareas que implican diferentes modalidades sensoriales (como escuchar una conferencia mientras se toman apuntes visuales) con menor interferencia que si se intentaran dos tareas que compiten por los mismos recursos de una única modalidad.

El ejecutivo central es responsable de la asignación de recursos cognitivos, la inhibición de información distractora o irrelevante, la actualización continua de la información relevante en la WM y la flexibilidad para cambiar entre diferentes tareas o estrategias mentales, estas son precisamente las funciones cognitivas que se ven comprometidas bajo condiciones de estrés, el cual puede generar rumiación o preocupación, consumiendo recursos ejecutivos y por déficits de atención que dificultan el mantenimiento del foco en la tarea y la supresión de interferencias.

2.2.2. El Modelo de Memoria Activada y Foco Atencional de Cowan

Nelson Cowan propuso un modelo alternativo que ofrece una perspectiva más unitaria e integrada de la memoria de trabajo [28], conceptualizándola no como un sistema de almacenamiento separado, sino como una porción activada de la memoria a largo plazo (*LTM por sus siglas en inglés long-term memory*).

A diferencia del modelo de Baddeley y Hitch, donde el ejecutivo central utiliza la atención como un recurso para gestionar los subsistemas, en el modelo de Cowan, el foco de atención es el núcleo mismo de la WM, el espacio donde la información es conscientemente procesada y manipulada, esta perspectiva es particularmente relevante ya que proporciona una base teórica directa para anticipar una modulación significativa del rendimiento de la WM por las fluctuaciones en la eficiencia atencional.

La concepción de la WM como una porción activada de la LTM, junto con la flexibilidad inherente al foco atencional (que puede ampliar o restringir su alcance según las demandas de la tarea, aunque siempre dentro de un límite de capacidad de *chunks*),

sugiere un sistema cognitivo más integrado y dinámico que el propuesto por el modelo multicomponente.

La flexibilidad del foco implica que la "capacidad" de la WM no es un valor fijo, sino que puede adaptarse a las demandas específicas de la tarea, esto podría explicar por qué algunas tareas parecen requerir más "espacio" o recursos de la WM que otras, o cómo la experiencia previa y el conocimiento almacenado en la LTM pueden influir en la eficiencia con la que opera la WM. Por ejemplo, un experto en un dominio particular puede manejar más información relevante para ese dominio en su WM, no porque su foco atencional sea intrínsecamente mayor, sino porque puede formar *chunks* más grandes y significativos a partir de su conocimiento consolidado en la memoria a largo plazo.

2.2.3. Modelo de Interferencia de la Memoria de Trabajo de Oberauer

El modelo de interferencia (*IM*, por sus siglas en inglés) propuesto por Klaus Oberauer y colaboradores [29] concibe que los contenidos de la WM son codificados mediante vinculaciones temporales a contextos específicos (por ejemplo, localizaciones espaciales o posiciones seriales), los cuales actúan subsecuentemente como claves de recuperación para acceder a dicha información.

El IM, más que una capacidad de almacenamiento fija o basada en "ranuras" (slots), se distingue por su énfasis en la interferencia como el principal factor limitante del rendimiento de la WM, en lugar de basarse únicamente en conceptos de decaimiento temporal o una capacidad de almacenamiento fija en términos de *slots* o un recurso continuo desplegable.

Oberauer extendió y refinó el modelo de Cowan, proponiendo una estructura jerárquica más diferenciada dentro de la WM. Una distinción clave es la introducción de un foco de atención aún más limitado que en el modelo de Cowan, postulando que solo

un *chunk* o ítem puede ser procesado activamente en un momento dado [30]. Sus investigaciones se han enfocado en dilucidar los mecanismos de selección, actualización y recuperación de la información en la WM, y cómo estos procesos son afectados por la interferencia. Esta perspectiva dinámica es particularmente adecuada para integrar estados fluctuantes como la atención y el estrés.

La elección del IM de Oberauer se justifica por dos razones: primero, su explícita modelización de la interferencia y la recuperación basada en claves contextuales lo hace particularmente adecuado para investigar cómo factores moduladores (como la atención y el estrés que plausiblemente afectan los niveles de interferencia o la precisión de las vinculaciones) impactan la WM; segundo, la tarea experimental seleccionada para este estudio, la tarea Letterwheel, se alinea directamente con los supuestos del IM sobre la recuperación de ítems a partir de contextos espaciales [17]. La conceptualización de la interferencia en el modelo de interferencia como un proceso activo y dinámico ofrece un marco más flexible para incorporar estados fluctuantes como la atención y el estrés, en comparación con modelos de capacidad más estáticos como los basados en ranuras [30].

El IM, con su énfasis en gradientes de interferencia y precisión de vinculación, ofrece parámetros específicos que podrían ser afectados por estados fisiológicos, constituyendo así un punto de partida teórico para las modificaciones propuestas en esta investigación.

2.2.3.1. Mecanismos de Interferencia

La interferencia es un concepto central en el modelo de Oberauer [17], en la que postula que la información en la WM es susceptible a la degradación y al olvido principalmente debido a la interferencia entre representaciones, más que a una simple decaída pasiva por el tiempo, de las cuales derivan dos conceptos a tomar en cuenta como lo son: la Interferencia Proactiva, en la que la información previamente procesada

o mantenida puede interferir con la codificación o recuperación de nueva información; así como también, se presenta la Interferencia Retroactiva, en la que la nueva información puede interferir con el mantenimiento o la recuperación de información más antigua.

2.3. Modelo de Interferencia para la Memoria de Trabajo Visual y Verbal

En el "*Interference Model for Visual and Verbal Working Memory*", Oberauer y Lin [31] asumen que los contenidos se codifican en la WM mediante uniones temporales a representaciones de contexto (por ejemplo, posiciones seriales, ubicaciones espaciales). Estas uniones tienen una precisión limitada tanto en la dimensión del contenido como en la del contexto y la recuperación se basa en el uso del contexto como clave.

La interferencia surge de la confusión entre uniones similares o de la activación de uniones incorrectas, el modelo explica fenómenos como la forma de la distribución de errores en tareas de reproducción continua (que suelen tener colas pesadas, indicando errores grandes ocasionales además de pequeños errores de precisión) y cómo esta distribución se ve afectada por el tamaño del conjunto de memoria (set size), también aborda los efectos de posición serial (primacía y recencia).

2.3.1. Formulaciones Matemáticas Clave y Parámetros

En el Modelo de Interferencia, no se modela de manera explícita el espacio de enlaces o asociaciones entre los ítems de memoria, en cambio, el modelo se enfoca en representar la distribución de activación que se genera nuevamente (o se recupera) dentro de una dimensión de características.

En este contexto, los valores posibles dentro de esa dimensión de características, o un subconjunto de ellos, constituyen los candidatos a ser recuperados por el sistema de memoria, así pues, la distribución de activación funciona como una medida de la evidencia a favor de cada uno de estos candidatos, indicando qué tan probable es que cada uno sea la respuesta correcta, según la información almacenada en la memoria.

Cada respuesta de recuerdo (recall) dentro del modelo se toma como una elección probabilística entre ese conjunto de candidatos, basada en la fuerza relativa de activación $A(x)$ de cada opción x en el momento de la recuperación.

La probabilidad de elegir una respuesta específica x , dentro de un total de N opciones, se calcula con la siguiente ecuación (Ec. 2.3.1.1), que normaliza la activación para que así las probabilidades sumen uno. Los detalles de estas ecuaciones se pueden encontrar en [31].

$$P(x) = \frac{A(x)}{\sum_{j=1}^N A(j)} \quad (2.3.1.1)$$

Donde:

- x : es una respuesta específica.
- $N \in [1, 19]$: es el número total de opciones a escoger.
- $A(x)$: representa la fuerza relativa de activación.
- $\sum_{j=1}^N A(j)$ suma todas las activaciones de las posibles respuestas, de $j = 1$ a N , para que la probabilidad total quede normalizada (sume 1).

Esta ecuación (Ec. 2.3.1.1), implica que la probabilidad de seleccionar una opción depende directamente de cuán activada esté en comparación con las demás.

Ahora bien, la distribución de activación total (Ec. 2.3.1.2) que representa la fuerza con la que cada posible respuesta es activada en el momento de recuperación, se compone de tres elementos principales, cada uno de los cuales refleja una fuente distinta de información en la memoria.

$$A(x|L_\theta) = c \cdot A_c(x|L_\theta) + a \cdot A_a(x) + b \cdot A_b(x) \quad (2.3.1.2)$$

Donde:

- $A(x|L_\theta)$: representa la activación total del ítem x (para el caso de Letterwheel representa un nonword), dada una localización L_θ .
- L_θ : representa la localización del nonword.
- a, b, c : son coeficientes de ponderación que determinan la influencia relativa de cada componente sobre la activación total.
- $A_c(x|L_\theta)$: es el componente de recuperación basada en ítems relacionadas con el nonword x y la posición del ítem L_θ durante la codificación.
- $A_a(x)$: es la activación basal del ítem x , es decir, una activación constante que representa su disponibilidad general en memoria independientemente de los ítems de recuperación.
- $A_b(x)$: es el componente de ruido o interferencia de fondo, que introduce variabilidad no sistemática en la activación de los ítems.

En este caso, la variable x representa un ítem candidato a ser recuperado por el modelo en un intento dado. En el caso específico del Letterwheel, estos ítems corresponden a las nonwords presentadas visualmente al participante durante la tarea.

Por lo tanto, los posibles valores de x pertenecen al conjunto de nonwords disponibles en cada intento, es decir, una nonword distinta ubicada en el círculo de respuesta, cada una de estas palabras es evaluada para calcular su probabilidad de ser seleccionada como respuesta en función de su activación total $A(x|L_\theta)$, la cual determina su probabilidad de ser seleccionada como respuesta en función de su grado de asociación con la localización proporcionada L_θ .

De esta manera, siguiendo la Ec. 2.3.1.2, $A(x|L_\theta)$ está compuesta por la suma ponderada de tres componentes que reflejan distintas fuentes de información durante el proceso de recuperación de la memoria.

La primera de ellas, $A_c(x|L_\theta)$, representa la activación que se genera a partir de la localización del ítem objetivo durante la tarea. Por ejemplo, “recuerda el nonword en la posición X). Dado que no solo se activa la característica objetivo, sino también otras que estuvieron en contextos similares (por ejemplo, ítems cercanos espacialmente), lo que puede llevar a errores por intrusión. Esta activación es ponderada por el parámetro c , que representa la fuerza de la recuperación basada en las localizaciones y cuya ecuación (Ec. 2.3.1.3) es la siguiente:

$$A_c(x|L_\theta) = \sum_{i=1}^n [w_{direct} \cdot R_1 + (1 - w_{direct}) \cdot R_2] \cdot VM(x, x_i, k) \quad (2.3.1.3)$$

Donde:

- n : representa el número de ítems en la memoria en ese intento.
- x_i : representa el ítem i codificado en el intento t .
- $VM(x, x_i, k)$: es la función de distribución de von Mises, que en este caso describe la similitud entre el candidato x y el ítem x_i , modulada por un parámetro de precisión k .

- w_{direct} : es un parámetro que controla el peso relativo de la ruta directa respecto a la mediada. Valores bajos de w_{direct} implican mayor dependencia de la recuperación mediada.

Ahora bien, la ruta directa R_1 , representa el acceso a la información en memoria a partir de una pista espacial inmediata, en este caso, la activación de un ítem depende exclusivamente de qué tan cerca se encuentra en términos espaciales. Cuanto más próxima sea la ubicación del ítem a la posición objetivo, mayor será su activación.

Esta ruta asume que la recuperación de la información es guiada directamente por la similitud espacial entre la pista y la posición original del ítem. La ruta directa se describe en la Ec. 2.3.1.4.

$$R_1 = \exp[-s_{space} \cdot D(L_i, L_\theta)] \quad (2.3.1.4)$$

Donde:

- $D(L_i, L_\theta)$: es la distancia espacial entre la posición del ítem i y la posición objetivo L_θ .
- s_{space} : es un parámetro de sensibilidad que controla qué tan rápido decae la activación con la distancia espacial.

Y por último, la ruta mediada R_2 , introduce un componente adicional al proceso de recuperación, considerando no sólo la cercanía espacial, sino también la cercanía temporal entre los ítems presentados. En este caso, se evalúan todas las posiciones posibles en las que un ítem pudo haber sido codificado y se pondera la activación combinando tanto la similitud espacial con la pista como la similitud temporal entre ítems.

Esta ruta modela un proceso más complejo de recuperación, en el que los participantes pueden confundir ítems no solo por su proximidad en el espacio, sino también por su cercanía en el tiempo de presentación, la ruta mediada se describe en la Ec. 2.3.1.5.

$$R_2 = \sum_{j=1}^n \exp[-s_{space} \cdot D(L_j, L_\theta)] \cdot \exp[-s_{time} \cdot D(i, j)] \quad (2.3.1.5)$$

Aquí:

- $\exp[-s_{space} \cdot D(L_j, L_\theta)]$: representa cuán fuerte es la activación de la posición de entrada j dada la localización L_θ .
- $\exp[-s_{time} \cdot D(i, j)]$: representa cuán fuerte es la activación del ítem i a partir de la posición j , en función de su cercanía en la dimensión temporal (es decir, en el orden de presentación).

Este modelo incorpora así tanto la proximidad espacial como la proximidad temporal como factores clave en la activación durante la recuperación. La inclusión de la ruta mediada permite capturar errores de tipo transposición basados en similitud serial, lo que da cuenta de fenómenos como los efectos de primacía y recencia debido a la forma del gradiente de activación en la dimensión temporal.

Seguidamente, la segunda componente $A_a(x)$, corresponde a la activación asociada al conjunto de memoria actual, independientemente del contexto en que los ítems fueron presentados, esto implica que todos los elementos que formaron parte del set de memoria reciben un nivel de activación simplemente por haber estado presentes, sin que se considere su ubicación específica o el orden de aparición. Este tipo de activación da cuenta de fenómenos de familiaridad, donde el sujeto reconoce una nonword como previamente vista, aunque no recuerde con exactitud dónde se encontraba. La influencia de esta fuente de activación está ponderada por el parámetro a

que cuantifica su peso relativo durante el proceso de recuperación y cuya ecuación descrita en la Ec. 2.3.1.6 es:

$$A_a(x) = \sum_{i=1}^n VM(x, x_i, k) \quad (2.3.1.6)$$

Por último, la componente $A_b(x)$, representa la activación de fondo o ruido. A diferencia de las anteriores, esta activación se distribuye de forma uniforme entre todos los candidatos posibles de respuesta, sin depender del contexto ni de la familiaridad del ítem, su función es modelar la incertidumbre basal del sistema de memoria, aportando una activación constante que puede explicar elecciones completamente incorrectas en ausencia de señales claras. Esta componente está modulada por el parámetro b , que determina el grado de influencia del ruido aleatorio en la activación total y cuya ecuación descrita en la Ec. 2.3.1.7 es:

$$A_b(x) = n \cdot U \quad (2.3.1.7)$$

Donde:

- n : representa el número de ítems en la memoria en ese intento.
- U : es una constante que representa activación uniforme o ruido basal que se reparte por igual entre todos los posibles candidatos.

Así, el modelo puede explicar los efectos del tamaño del conjunto de memoria y las curvas de posición serial sin recurrir a un número fijo de "slots" o un recurso general que se agota, basándose en cambio en la interferencia y la modulación de la fuerza de los enlaces.

2.3.2. Validación y Relevancia del Modelo de Oberauer

El modelo de Oberauer ha sido validado a través de hallazgos empíricos en tareas de memoria de trabajo (WM), tanto visuales como verbales, incluyendo distribuciones de error, efectos de posición serial y gradientes de transposición [32]. La hipótesis del enlace (*binding hypothesis*) sostiene que la capacidad central de la WM radica en la habilidad para construir, mantener y actualizar enlaces arbitrarios, una función esencial para el razonamiento [33]. Por ello, las tareas de actualización en WM se consideran buenas medidas de esta capacidad, ya que requieren la manipulación dinámica de estos enlaces temporales.

Para el presente proyecto, la relevancia del modelo se encuentra en que las variaciones en los niveles de atención y estrés podrían influir directamente en parámetros clave del mismo, en particular, se propone modificar el componente de recuperación basada en pistas A_c (Ec. 2.3.1.3), modulando el parámetro w_{direct} en función del nivel de atención (ya que una mayor atención favorecería una recuperación más directa), mientras que el estrés podría impactar las rutas R_1 y R_2 (Ec. 2.3.1.4 y Ec. 2.3.1.5), modulando las variables de s_{space} y s_{time} , alterando la precisión tanto del contexto espacial como del orden temporal para así permitir que esta integración pueda evaluar cómo factores fisiológicos afectan el proceso de recuperación en la memoria de trabajo.

2.4. Neurociencia Cognitiva: Atención y Estrés Mental

La memoria de trabajo no opera en un vacío cognitivo, sino que su eficiencia está intrínsecamente ligada a otros procesos mentales fundamentales, entre los que destacan la atención y el nivel de estrés mental [34]. Este apartado explora los constructos de atención y estrés desde una perspectiva neurocognitiva, detallando sus modelos teóricos, sus bases neurales y, crucialmente, su impacto modulador sobre el rendimiento de la

WM. Esta comprensión es esencial para este proyecto, que busca predecir el rendimiento de la WM en estudiantes universitarios precisamente en función de sus niveles de atención y estrés mental.

2.4.1. Atención: Constructo y Modelos Neurocognitivos

La atención es una función cognitiva multifacética y de capacidad limitada que desempeña un papel crítico en prácticamente todos los aspectos del procesamiento de la información y la conducta adaptativa. Se puede definir, en términos generales, como la capacidad del sistema cognitivo para seleccionar información relevante del flujo de estímulos internos y externos, focalizar los recursos de procesamiento en dicha información, y mantener ese foco durante un periodo de tiempo, mientras se inhibe o se filtra la información irrelevante o distractora [35]. La atención actúa, por tanto, como un mecanismo de filtro y de asignación de recursos, siendo fundamental para dirigir nuestros pensamientos, regular nuestras emociones y orientar nuestras conductas hacia la consecución de objetivos específicos en la vida diaria. Dada su complejidad, se han propuesto diversos modelos para explicar su funcionamiento, destacando entre ellos el modelo de redes atencionales de Posner y Petersen [36] y el modelo jerárquico de Sohlberg y Mateer [37].

2.4.2. Estrés Mental y su Impacto en los Procesos Cognitivos

El estrés mental es una respuesta psicofisiológica compleja del organismo ante situaciones o estímulos (estresores) que son percibidos como amenazantes, desbordantes o que exceden la capacidad de afrontamiento del individuo. Si bien una cierta cantidad de estrés puede ser adaptativa e incluso mejorar el rendimiento (eustrés), el estrés excesivo o crónico (distrés) tiene efectos perjudiciales como lo menciona Lukasik [38]

sobre la salud física y mental, y de manera particular, sobre el funcionamiento cognitivo, incluyendo la atención y la memoria de trabajo.

Los efectos del cortisol sobre la cognición también son complejos y a menudo siguen una curva en forma de U invertida (Ley de Yerkes-Dodson) [39]. Niveles basales o ligeramente elevados de cortisol pueden ser necesarios para un funcionamiento cognitivo óptimo y pueden incluso facilitar ciertos aspectos de la memoria (especialmente la consolidación). Sin embargo, niveles crónicamente elevados de cortisol, o picos agudos muy altos, tienden a ser perjudiciales para la memoria de trabajo, la atención, la flexibilidad cognitiva y la recuperación de memoria explícita.

2.5. Electroencefalografía (EEG) y Medición de Señales Cerebrales

La electroencefalografía (EEG) es una técnica no invasiva ampliamente utilizada en neurociencia para medir la actividad eléctrica cerebral a través de electrodos colocados sobre el cuero cabelludo. Su principal ventaja es su alta resolución temporal (milisegundos), lo que la hace especialmente útil para estudiar la dinámica de procesos cognitivos como la atención, la percepción o la memoria de trabajo [40]. No obstante, su resolución espacial es limitada en comparación con métodos como la resonancia magnética funcional (fMRI).

En este sistema, la ubicación Fp1 (frontal polar izquierda), está asociada con funciones ejecutivas, atención y regulación emocional [23]. Aunque en este estudio se utilizó un único canal (Fp1) provisto por el NeuroSky Mindwave, comprender esta localización permite contextualizar los registros obtenidos en relación con la actividad de la corteza prefrontal izquierda, una región clave en procesos atencionales y en la regulación del estrés. En la Fig 2.5 se muestra la ubicación Fp1.

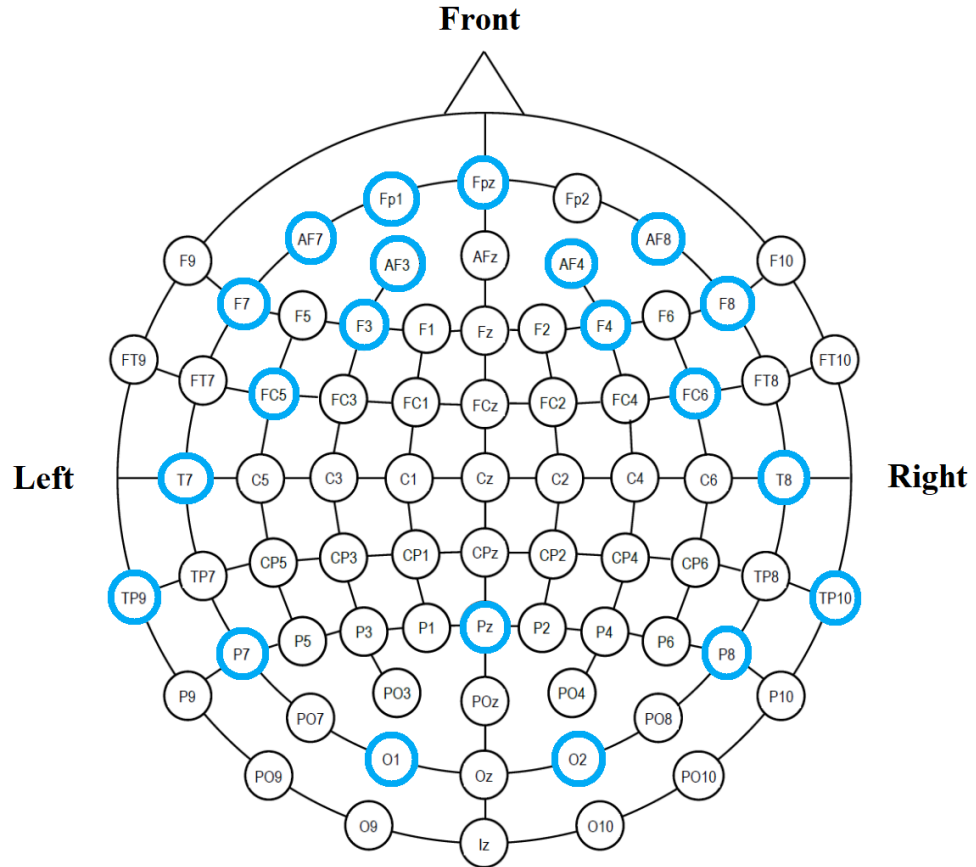


Figura 2.5: Posiciones del lóbulo frontal y la ubicación del Fp1.

2.5.1. Bandas de Frecuencia EEG y Estados Mentales Asociados

El EEG puede descomponerse en distintas bandas de frecuencia, cada una asociada con estados mentales y funciones cognitivas específicas como las que se mencionan a continuación:

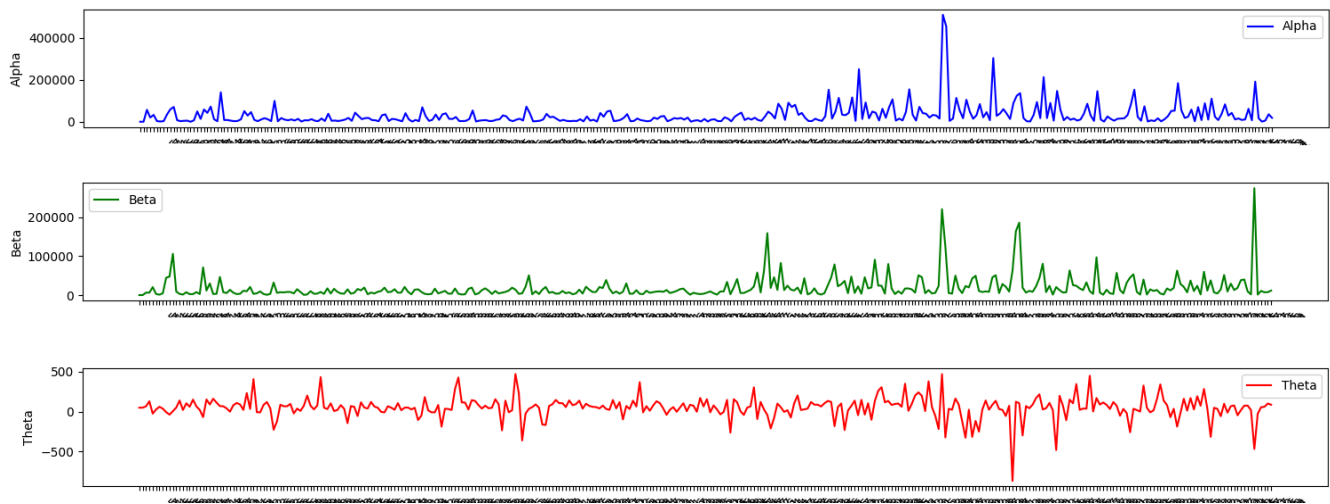


Figura 2.5.1: Bandas de frecuencia alpha beta y theta

2.5.1.1. Alpha (α) (8-13 Hz)

La actividad alpha es la oscilación dominante en el EEG de un adulto relajado con los ojos cerrados. Se asocia con estados de calma y atención pasiva. La supresión o desincronización alpha (disminución de la potencia alpha) se observa con la apertura de ojos, el aumento de la alerta, el esfuerzo mental y la atención activa, reflejando una mayor actividad cortical [41]. La supresión de alpha en regiones posteriores se ha vinculado con el mantenimiento activo de representaciones en la memoria de trabajo visual.

2.5.1.2. Beta (β) (13-30 Hz)

La banda beta está relacionada con estados de alerta, atención activa y enfocada externamente, concentración, actividad mental intensa y procesamiento cognitivo [42], también puede estar elevada en estados de ansiedad o estrés [43]. Incrementos en la potencia beta se han relacionado con la atención sostenida.

2.5.1.3. Theta (θ) (4-8 Hz)

La actividad theta está asociada con estados de somnolencia, relajación profunda, meditación, y también con procesos de memoria y aprendizaje, especialmente la codificación y recuperación de información. El aumento de la actividad theta frontal se ha relacionado con una mayor carga de memoria de trabajo y control cognitivo [43].

2.5.2. Correlatos EEG de la Atención

La atención sostenida y el control ejecutivo se reflejan en patrones específicos de actividad eléctrica cerebral, y una de las métricas más ampliamente utilizadas en este contexto es la razón Beta/Theta (β/θ) [43], la cual compara la potencia relativa de las bandas de frecuencia beta y theta. Esta relación ha demostrado ser un marcador confiable del estado atencional del individuo, ya que integra la dinámica de dos ritmos que responden de manera opuesta a los requerimientos cognitivos.

Las oscilaciones theta, especialmente en regiones frontales, están vinculadas con estados de relajación, fatiga cognitiva y disminución del esfuerzo mental. Por el contrario, la actividad en la banda beta se incrementa durante periodos de atención activa, concentración sostenida y procesamiento cognitivo exigente, así pues un aumento en la razón β/θ , por tanto, sugiere un predominio de la activación cortical orientada hacia el control de tareas y la atención sostenida, mientras que valores bajos suelen estar asociados a una menor activación cognitiva o a estados de somnolencia.

Esta relación ha sido ampliamente utilizada en estudios clínicos y experimentales como índice cuantitativo del nivel atencional, por ejemplo, se ha documentado su utilidad en la evaluación de sujetos con trastorno por déficit de atención e hiperactividad (*TDAH*), en donde una razón β/θ reducida aparece como patrón distintivo [44]. También se ha empleado en investigaciones sobre neurofeedback, donde el entrenamiento para

aumentar esta razón se asocia con mejoras en la capacidad de atención y rendimiento académico. Asimismo, en contextos de carga cognitiva elevada, un aumento sostenido de la razón β/θ se correlaciona con una mayor eficacia en la ejecución de tareas de memoria de trabajo, tiempos de reacción más cortos y menor susceptibilidad a la distracción [45].

Desde el punto de vista neurofisiológico, esta razón captura la dinámica entre una banda inhibidora (theta) y una excitadora (beta), reflejando el grado de activación cortical frente a demandas cognitivas [46]. En particular, cuando se registra actividad en regiones prefrontales, se interpreta como un reflejo del esfuerzo mental voluntario, la planificación y el control inhibitorio, funciones directamente asociadas al componente ejecutivo de la atención [47].

Dado lo anterior, en este estudio la razón β/θ fue empleada como indicador principal de atención, evaluando su comportamiento durante las distintas fases de la tarea cognitiva para identificar variaciones que pudieran reflejar fluctuaciones en el estado atencional del sujeto.

2.5.3. Correlatos EEG del Estrés Mental

El estrés mental agudo también puede ser cuantificado mediante correlatos neuroeléctricos, siendo la razón Alpha/Beta (α/β). Este es uno de los indicadores más robustos en contextos de evaluación psicofisiológica [42], porque esta relación compara la actividad en la banda alpha, tradicionalmente asociada a la relajación, la inhibición cortical y el control interno, con la actividad en la banda beta, la cual se incrementa ante estímulos externos, procesamiento cognitivo intenso y activación emocional.

Bajo condiciones de estrés, es común observar una disminución sostenida en la potencia alpha, particularmente en regiones frontales, acompañada de un aumento en la potencia beta. Este patrón genera una caída en la razón α/β , la cual ha sido asociada con estados de sobrecarga cognitiva, hiperactivación del sistema nervioso simpático y aumento de la tensión emocional [41]. En estudios experimentales que inducen estrés

mediante tareas cronometradas o retroalimentación negativa, esta razón ha mostrado sensibilidad para distinguir entre estados de calma, alerta y ansiedad.

Desde una perspectiva neurofuncional [48], la banda alpha cumple un rol modulador de la actividad cortical, inhibiendo regiones que no están involucradas en la tarea en curso. Su supresión refleja un desbalance entre los recursos cognitivos disponibles y las demandas de la tarea y al mismo tiempo, el incremento de la banda beta en situaciones de estrés indica un estado de alerta elevado que, si bien puede favorecer el rendimiento a corto plazo, en exceso puede conducir a fatiga y deterioro de la función ejecutiva.

En este proyecto, la razón α/β se utilizó como un estimador del nivel de estrés mental experimentado por los participantes, siendo analizada en función de la carga cognitiva de las tareas asignadas. Su comportamiento permitió establecer inferencias sobre el nivel de estrés durante el experimento y su relación con el rendimiento.

2.6. Dispositivo Neurosky Mindwave

El NeuroSky MindWave 1 (Fig. 2.6) es un dispositivo de electroencefalografía (EEG) de canal único, diseñado para capturar señales eléctricas cerebrales no invasivas a través de un sensor frontal seco ubicado en la posición Fp1 del sistema internacional 10–20, y una pinza de referencia en el lóbulo de la oreja izquierda. Este opera a una frecuencia de muestreo de 512 Hz [49], lo que le otorga una resolución temporal adecuada para capturar la dinámica rápida de las oscilaciones cerebrales. Internamente, incorpora un chip TGAM1 (ThinkGear ASIC Module), que se encarga tanto de la amplificación de la señal EEG como del procesamiento digital, este chip realiza filtrado de artefactos, extracción de características y segmentación en bandas de frecuencia mediante algoritmos propietarios, y produce métricas como los niveles de atención y meditación, así como la potencia relativa de las ondas delta, theta, alpha, beta y gamma [50].

Una de las ventajas del MindWave 1 es su conectividad inalámbrica, que se establece mediante Bluetooth estándar (versión 2.1 + EDR) en la frecuencia de 2.4 GHz. La comunicación se realiza mediante el protocolo de puerto serie (SPP), lo que permite una integración sencilla con múltiples plataformas, incluyendo Windows, macOS, Linux y dispositivos móviles Android mediante el uso de adaptadores Bluetooth compatibles o módulos USB-UART. Esta conectividad permite la transmisión en tiempo real de los datos EEG a una tasa de baudios estándar de 57600 bps, suficiente para garantizar una transferencia fluida de los paquetes de datos que incluyen tanto las métricas pre-procesadas como los valores crudos de la señal. Todos estos datos fueron obtenidos del manual de usuario propio de la marca [\[51\]](#).

El dispositivo se conecta automáticamente al puerto serie virtual del sistema operativo, permitiendo la adquisición de datos mediante scripts personalizados escritos en lenguajes como Python o MATLAB, a través de bibliotecas que interpretan el formato del protocolo ThinkGear. Cada paquete de datos transmitido incluye una estructura codificada que contiene valores de señal cruda, la potencia en bandas específicas y los niveles de atención o meditación calculados por el firmware del dispositivo.

Debido a su portabilidad, facilidad de uso y bajo costo, el NeuroSky MindWave 1 tiene potencial para ser utilizado en contextos educativos, investigaciones exploratorias, estudios de neurofeedback y aplicaciones interactivas. Si bien su diseño no alcanza la precisión clínica de sistemas multicanal profesionales, proporciona una herramienta válida para el estudio de correlatos neurofisiológicos en tareas cognitivas controladas.



Figura 2.6: Neurosky Mindwave 1.

Una de las principales limitaciones del MindWave 1 radica en su naturaleza de "caja negra". Los algoritmos que utiliza para calcular métricas cognitivas no son de dominio público, lo que impide una comprensión completa de cómo se derivan los datos proporcionados. Esta opacidad limita la capacidad de los investigadores para validar y replicar estudios basados en este dispositivo. Además, al ser un dispositivo de un solo canal, no permite la aplicación de técnicas avanzadas de procesamiento de señales.

2.6.1. Casos de Uso en Investigación Cognitiva

A pesar de ser un dispositivo de bajo costo, el Neurosky Mindwave ha sido empleado en numerosos estudios de investigación para evaluar estados cognitivos y afectivos como la atención, el estrés, la carga cognitiva, la fatiga, e incluso para aplicaciones de Interfaz Cerebro-Computadora (*BCI*) y la evaluación del rendimiento en tareas cognitivas [52, 53]. Algunos estudios han encontrado que el dispositivo puede medir cuantitativamente los niveles de atención [54] y que sus lecturas de atención

(posiblemente las métricas eSense) se correlacionan con medidas de autoinforme [55], y con el rendimiento en pruebas de atención estandarizadas, donde se ha utilizado para investigar el estrés [38].

Para este Proyecto de Grado, se trabajará con los datos crudos de EEG (RAW EEG) transmitidos por el dispositivo, en lugar de depender de las métricas propietarias de eSense ("Atención" y "Meditación"). El objetivo es desarrollar un modelo predictivo propio y transparente, basado en características EEG derivadas y procesadas, lo cual requiere acceso a los datos fundamentales de la señal cerebral. Confiar en las métricas eSense sería utilizar una "caja negra" cuyo algoritmo no es público y no permitiría una comprensión de la relación entre la actividad cerebral y el rendimiento en la memoria de trabajo.

2.7. Construcción y Análisis de Datasets en Neurociencia Cognitiva

El desarrollo de modelos predictivos confiables en neurociencia cognitiva depende de la calidad y estructura del dataset utilizado. Un dataset en este contexto se construye a partir de datos experimentales que capturan actividad fisiológica y conductual, como señales EEG y respuestas en tareas cognitivas. La rigurosidad en la recolección, procesamiento y documentación de estos datos garantiza su validez para el modelado computacional y para análisis posteriores.

2.7.1. Buenas prácticas para la recolección de datos experimentales

La recolección de datos en entornos neurocientíficos debe cumplir con estándares metodológicos que aseguren la fiabilidad y reproducibilidad de los resultados. Entre las buenas prácticas recomendadas se encuentra el control del ambiente experimental:

reducción de estímulos externos (ruido, luz intensa), homogeneidad en el horario de aplicación, y condiciones constantes entre sesiones [56].

Además, es esencial capacitar previamente al participante en la tarea que realizará, verificar el estado técnico del equipo de medición (como el Neurosky Mindwave 1) y registrar eventos que puedan interferir en la adquisición de datos, como interrupciones o movimientos bruscos.

Otra práctica recomendada es la recolección de datos basales antes de la tarea experimental, con el objetivo de establecer una línea base fisiológica del sujeto, lo cual facilita la interpretación de las variaciones durante la prueba.

2.7.2. Procesamiento de Señales y Estructuración del Dataset

Las señales EEG son altamente sensibles al ruido eléctrico, movimientos musculares y parpadeos, por lo que es necesario aplicar técnicas de limpieza y preprocesamiento. En este estudio, se utilizó un filtro paso banda entre 1–50 Hz para preservar las bandas alpha, beta y theta relevantes en contextos de atención y memoria [42,43]. Asimismo, se aplicaron transformaciones estadísticas como la **Normalización Min-Max** en un rango entre [0, 1] (Ec. 2.7.2.) para estandarizar las métricas de atención y estrés derivadas del dispositivo.

$$x_{norm} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (2.7.2)$$

Luego de esto, se ordena el dataset estructurado en formato tabular (CSV), incluyendo columnas como: identificador del participante, timestamp, nivel de atención, nivel de estrés, condición experimental, y rendimiento en tareas (aciertos/errores), para así permitir análisis exploratorios, estadísticos y modelado computacional eficiente [57].

Además, se conservaron los metadatos asociados a cada sesión, como la fecha, duración de la tarea, edad de los participantes, entre otros datos que sigan las recomendaciones para datos neurocientíficos reproducibles. La preservación sistemática de estos metadatos no solo permite contextualizar adecuadamente los resultados, sino que también facilita su interpretación, replicación y reutilización en estudios futuros. Según Peng [58], los metadatos bien documentados son esenciales para garantizar la trazabilidad del flujo de datos.

2.7.3. Consideraciones Éticas y Protocolos de Consentimiento

Las investigaciones con datos de origen humano, especialmente fisiológicos, implican la implementación de protocolos éticos rigurosos. En este estudio se aplicaron los principios del consentimiento informado, anonimización de los datos y libre retiro del estudio en cualquier momento [59].

El formulario de consentimiento contenía información detallada sobre los objetivos del estudio, los procedimientos involucrados, la confidencialidad de los datos, y las implicaciones del uso del dispositivo EEG. Se dejó claro que el dispositivo no tiene fines clínicos ni de diagnósticos.

2.7.4. Herramientas de captura y análisis

La captura de datos EEG se realizó a través del software NeuroPy [60], una biblioteca de Python compatible con el sensor MindWave de NeuroSky, que permite acceder en tiempo real a métricas como la de las ondas alpha, beta y theta.

Los datos obtenidos fueron procesados utilizando Python, aprovechando numpy, pandas y matplotlib para la limpieza, análisis y visualización preliminar de las señales. Posteriormente, se empleó R para análisis estadísticos más avanzados y ajuste de

modelos predictivos, dada su robustez para trabajar con modelos lineales generalizados, pruebas de hipótesis y validaciones cruzadas.

2.8. Modelado Predictivo en Ciencias Cognitivas

El modelado predictivo es una técnica que permite estimar el comportamiento de variables cognitivas a partir de datos empíricos, integrando teorías psicológicas y herramientas computacionales. Este enfoque resulta particularmente útil en neurociencia cognitiva, donde fenómenos como la memoria de trabajo pueden representarse matemáticamente e interpretarse en función de variables fisiológicas como la atención o el estrés.

2.8.1. Conceptos Básicos de Modelado Matemático

Un modelo matemático en ciencias cognitivas representa un proceso mental mediante ecuaciones que describen cómo las entradas (inputs) se transforman en salidas (outputs). En el presente proyecto se implementó una versión modificada del Modelo de Interferencia de Klaus Oberauer [17], el cual explica el deterioro de la memoria de trabajo debido a la interferencia de información previa.

Este modelo fue adaptado para incluir moduladores fisiológicos, donde los valores de atención y estrés influyen directamente en parámetros como el peso relativo de las rutas (w_{direct}) y los pesos de decaimiento de las rutas (s_{space} y s_{time}), de esta forma, se integra información de desempeño cognitivo y biomarcadores EEG en un mismo marco analítico.

2.8.2. Ajuste de Modelos con Base en Datos Empíricos

El procedimiento de ajuste de modelos implementado en esta investigación se basa en la metodología descrita por Oberauer y Lin [17]. En sus investigaciones, se exploraron variantes del Modelo de Interferencia con distintas restricciones sobre el componente de ruido de fondo (A_b). Específicamente, se comparó la versión original, en la que el ruido crece con el tamaño del conjunto, con una variante de ruido fijo, donde A_b permanece constante independientemente del número de ítems codificados. Esta variante se fundamenta en la hipótesis de que el ruido puede originarse durante la fase de recuperación y no necesariamente en la fase de codificación.

Adicionalmente, en el contexto de la Teoría de Detección de Señales (SDT), se evaluó una versión del modelo con la desviación estándar del ruido (σ) fijada en 1, como una forma de restringir la flexibilidad del modelo y controlar su complejidad.

El ajuste se realizó individualmente para cada participante utilizando un enfoque de máxima verosimilitud, debido a que los intentos iniciales en un marco jerárquico bayesiano (empleando JAGS y Stan) presentaron problemas de convergencia en las cadenas MCMC y para evitar mínimos locales, el modelo se ejecutó múltiples veces con valores iniciales distintos utilizando dos algoritmos: Nelder-Mead (función `nmkb` del paquete `dfoptim` en R) y L-BFGS-B (función `optim` del paquete base de R). Se seleccionó la mejor solución de entre las 20 ejecuciones según el menor valor de la Deviance ($-2 \times \log\text{-likelihood}$).

2.8.3. Comparación de Modelos: Métricas de Evaluación y Validación

En el contexto de la evaluación del desempeño de modelos predictivos, resulta fundamental el uso de métricas que permitan no solo cuantificar la efectividad general del modelo, sino también analizar su comportamiento frente a diferentes tipos de errores.

Estas métricas se calculan a partir de diversas ecuaciones que ayuden al análisis y resumen de los resultados de clasificación comparando las predicciones del modelo con las etiquetas reales.

Una de las herramientas fundamentales en la evaluación de modelos de clasificación es la Matriz de Confusión [61], la cual se trata de una representación tabular que muestra el número de aciertos y errores cometidos por el modelo al comparar sus predicciones con las clases reales.

Esta matriz permite distinguir entre los siguientes cuatro casos:

- **Verdaderos Positivos (VP):** instancias correctamente clasificadas como positivas.
- **Falsos Positivos (FP):** instancias incorrectamente clasificadas como positivas.
- **Falsos Negativos (FN):** instancias incorrectamente clasificadas como negativas.
- **Verdaderos Negativos (VN):** instancias correctamente clasificadas como negativas.

Y su estructura se expresa como se muestra en la Tabla 2.8.3:

Actual	1	Falsos Positivos (FP)	Verdaderos Negativos (VN)
	0	Verdaderos Positivos (VP)	Falsos Negativos (FN)
Matriz de Confusión		0	1
		Predicción del Modelo	

Tabla 2.8.3: Matriz de confusión (Clase 0 positiva).

Este enfoque gráfico no solo facilita la visualización del rendimiento global del modelo, sino también la identificación de posibles sesgos o debilidades específicas, por ejemplo, si tiende a confundir sistemáticamente una clase con otra.

A partir de los valores contenidos en la matriz de confusión, se calculan métricas de rendimiento fundamentales como:

La Exactitud (*accuracy*): es una de las métricas más utilizadas y se define como la proporción de predicciones correctas (tanto verdaderos positivos como verdaderos negativos) sobre el total de predicciones, si bien es útil como una medida global del desempeño, puede resultar engañosa en problemas con clases desbalanceadas. Se describe por medio de la Ec. 2.8.3.1:

$$accuracy = \frac{VP + VN}{VP + FP + FN + VN} \quad (2.8.3.1)$$

La Precisión (*precision*): mide la proporción de verdaderos positivos respecto al total de casos clasificados como positivos por el modelo. Esta métrica es especialmente relevante cuando el costo de una clasificación positiva errónea (falso positivo) es alto, como en contextos de diagnóstico. Se describe en la ecuación Ec. 2.8.3.2:

$$precision = \frac{VP}{VP + FP} \quad (2.8.3.2)$$

La Sensibilidad (*recall o true positive rate*): representa la capacidad del modelo para identificar correctamente los casos positivos. Es decir, mide cuántos de los casos que realmente pertenecen a la clase positiva fueron correctamente identificados por el modelo. Esta métrica es crítica cuando el objetivo es minimizar los falsos negativos, por ejemplo, al detectar estados de estrés elevados o bajo rendimiento cognitivo. Se describe en la Ec. 2.8.3.3:

$$recall = \frac{VP}{VP + FN} \quad (2.8.3.3)$$

La Especificidad (tasa de verdaderos negativos): indica la proporción de verdaderos negativos correctamente identificados. Es complementaria a la sensibilidad y cobra importancia cuando es crítico evitar clasificaciones positivas incorrectas. Se describe en la Ec. 2.8.3.4:

$$specificity = \frac{VN}{VN + FP} \quad (2.8.3.4)$$

El F1-score (medida armónica entre precisión y recall): es una medida que combina precisión y sensibilidad mediante su media armónica, y es útil cuando se desea lograr un balance entre ambas métricas, especialmente en conjuntos de datos con clases desbalanceadas. Un valor alto de F1-score indica que el modelo tiene tanto alta precisión como alta sensibilidad, lo cual es deseable en contextos experimentales donde ambos tipos de errores son relevantes. Se describe en la Ec. 2.8.3.5:

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (2.8.3.5)$$

Capítulo III - Diseño e Implementación

Este capítulo describe el diseño metodológico. Las fases abarcan desde la recolección de datos, pasando por el procesamiento, hasta la formulación del modelo matemático y su validación, lo que responde directamente a los objetivos específicos planteados.

3.1. Diseño Experimental

3.1.1. Tipo de Estudio

Este estudio adopta un enfoque metodológico cuantitativo experimental, con un diseño de un solo grupo, en el que se induce una tarea de memoria de trabajo computarizada (con estímulos tipo *nonwords*) mientras se registra la actividad cerebral mediante EEG con el fin de observar su efecto sobre variables fisiológicas (niveles de atención y estrés) y el rendimiento cognitivo.

3.1.2. Participantes

La muestra está conformada por 22 estudiantes universitarios activos de la carrera de Ingeniería de Sistemas de la Universidad de Los Andes, Núcleo Mérida, seleccionados mediante un muestreo no probabilístico por conveniencia, basado en la disponibilidad y accesibilidad de los sujetos. Esta muestra se enmarca dentro de lo que Fidias Arias [62] denomina población accesible, es decir, aquella fracción de la población a la cual el investigador puede acceder efectivamente en función de limitaciones geográficas, temporales, administrativas o de recursos.

El tamaño de la muestra se justificó tanto por restricciones logísticas (como la disponibilidad de equipos, espacio físico y tiempo de evaluación), como por la coherencia con estudios experimentales previos sobre memoria de trabajo, como el

desarrollado por Oberauer, en el que se utilizaron muestras pequeñas, pero controladas, para análisis detallados.

3.1.2.1. Criterios de Inclusión

- Edad comprendida entre 18 y 28 años.
- Estudiantes universitarios activos.
- Sin diagnóstico de trastornos neurológicos o psiquiátricos.
- Sin consumo de sustancias estimulantes en las 2 horas previas a la prueba.
- Disponibilidad para participar en tres sesiones de 45 minutos cada una.

3.1.2.2. Criterios de Exclusión

- Historial médico de trastornos neurológicos.
- Uso de marcapasos o diagnóstico de epilepsia.
- Problemas visuales que impidan la correcta lectura de los estímulos.
- Imposibilidad de comprometerse con la duración total del experimento.

La selección de los criterios de inclusión y exclusión se estableció con el fin de garantizar la validez interna del estudio y la homogeneidad de la muestra, aspectos fundamentales en investigaciones experimentales que involucran mediciones neurofisiológicas. El rango de edad entre 18 y 28 años se eligió para reducir la variabilidad asociada con el desarrollo neurológico temprano o con procesos de envejecimiento cognitivo [63], los cuales podrían afectar el desempeño en tareas de memoria de trabajo y la actividad cerebral registrada mediante EEG .

El requisito de ser estudiante universitario activo asegura un nivel educativo relativamente homogéneo, lo cual es importante para controlar posibles efectos de la escolaridad sobre las capacidades cognitivas evaluadas. Asimismo, la exclusión de participantes con antecedentes de trastornos neurológicos o psiquiátricos busca eliminar factores que podrían alterar significativamente las señales cerebrales o interferir con el procesamiento cognitivo típico [64].

La restricción sobre el consumo de sustancias estimulantes (como cafeína, bebidas energéticas o fármacos) [65] en las horas previas al experimento se fundamenta en la evidencia de que estas sustancias pueden modificar los niveles de activación cortical y, por ende, afectar las medidas neurofisiológicas, particularmente las asociadas con la atención y el estrés.

Por otro lado, la exclusión de personas con problemas visuales no corregidos o con dificultades para asistir a todas las sesiones responde tanto a criterios de calidad de los datos como a aspectos logísticos [66], ya que la lectura precisa de estímulos visuales y la participación completa en las sesiones son condiciones indispensables para la correcta implementación del protocolo experimental.

3.1.3. Variables

3.1.3.1. Variables Predictoras (Independientes)

En este estudio las variables predictoras son dos valores derivados de mediciones EEG: atención y estrés mental.

3.1.3.2. Variable de Resultado (Dependiente)

La variable dependiente es el Rendimiento en la Memoria de Trabajo, la cual fue cuantificada a través del desempeño en las tareas computarizadas diseñadas según los principios del Modelo de Interferencia de Oberauer [17]. El valor utilizado hace referencia a la proporción de ítems correctamente recordados en la posición serial exacta

en la que fueron presentados durante la fase de codificación. Esta medida es particularmente relevante en tareas de recuerdo ordenado, ya que implica no solo la recuperación del contenido, sino también de su estructura secuencial. Así, la precisión evalúa la fidelidad del recuerdo respecto al orden original, siendo sensible a errores de ubicación aun cuando el ítem recordado pertenezca al conjunto objetivo. Por ejemplo, si una secuencia es A-B-C-D, y el participante responde A-C-B-D, solo A y D se consideran precisos (correctos).

3.2. Consideraciones Éticas

La investigación con seres humanos requiere una atención meticulosa a los principios éticos para garantizar la protección, el bienestar y los derechos de los participantes, por lo cual antes de realizar la recolección de los datos se tomaron en cuenta algunos aspectos éticos importantes.

3.2.1. Consentimiento Informado

Antes de cualquier participación, se obtuvo un consentimiento informado (Ver Anexo 1) por escrito de cada estudiante. Este documento explicó de manera clara y comprensible el propósito del estudio, los procedimientos involucrados (colocación del dispositivo EEG, realización de tareas cognitivas), la duración estimada de la participación, los riesgos potenciales (que se anticipan como mínimos, como una posible ligera incomodidad por el uso del headset EEG o fatiga por las tareas) y los beneficios (directos, si los hubiera, o indirectos para el avance del conocimiento), el derecho del participante a retirarse del estudio en cualquier momento sin penalización alguna, y las medidas que se tomarán para asegurar la confidencialidad y el anonimato de sus datos.

Cabe destacar que, si bien los participantes fueron plenamente informados sobre el uso del dispositivo EEG y las tareas experimentales, no se les indicó explícitamente

que el estudio incluía una evaluación implícita de los niveles de atención y estrés mental. Esta decisión metodológica respondió a la necesidad de evitar sesgos cognitivos y conductuales inducidos por la expectativa del participante, los cuales podrían modificar su estado mental, alterar la actividad cerebral registrada y comprometer la validez ecológica de los datos recolectados. Este enfoque trata de manipular mínimamente las condiciones naturales de la cognición para preservar la autenticidad del rendimiento observado.

3.2.2. Privacidad y Anonimización de Datos

Todos los datos recolectados (EEG, rendimiento en tareas, información personal básica) fueron anonimizados. A cada participante se le asignó un ID único, y este ID fue el único identificador utilizado en los archivos de datos. Los archivos que contengan información personal identificable (como los formularios de consentimiento firmados) se almacenarán de forma segura y separada de los datos de investigación.

3.2.3. Manejo de Datos EEG

Se explicó a los participantes que las señales EEG registradas no permiten "leer pensamientos" ni acceder a información personal sensible, sino que reflejan patrones generales de actividad eléctrica cerebral asociados con estados cognitivos como la atención, el estrés o la relajación. Esta aclaratoria fue esencial para evitar malentendidos sobre la naturaleza de la tecnología utilizada, especialmente considerando que el EEG puede generar inquietudes debido a su asociación con el monitoreo de funciones cerebrales.

Los datos EEG fueron transmitidos de forma inalámbrica a través de Bluetooth y procesados en tiempo real mediante software especializado, utilizando algoritmos que segmentan la señal en bandas de frecuencia (delta, theta, alpha, beta, gamma), a partir de

las cuales se infieren métricas cognitivas. Sin embargo, se enfatizó que estas métricas son de carácter agregado y probabilístico, no deterministas, y no permiten identificar pensamientos individuales, emociones específicas ni diagnósticos médicos.

En cuanto al almacenamiento y gestión, los archivos EEG fueron codificados bajo identificadores únicos no vinculados directamente con los datos personales de los participantes. Se utilizaron formatos estándar como .csv para facilitar el análisis posterior y asegurar la interoperabilidad con Python y R. Estos archivos se almacenaron en discos duros cifrados y bajo respaldo en servidores seguros, accesibles sólo por los miembros autorizados del equipo de investigación.

Adicionalmente, se aplicaron principios de minimización de datos (solo se recolectó la información estrictamente necesaria para los objetivos del estudio) y limitación temporal del almacenamiento (los datos se conservarán sólo por el tiempo estipulado en el protocolo aprobado). Todo el tratamiento de los datos EEG se realizó conforme a las recomendaciones éticas en neurociencia cognitiva [59], promoviendo una gestión responsable, transparente y respetuosa con la privacidad de los participantes.

3.3. Recolección de Datos

3.3.1. Adaptación de la Tarea de Memoria Letterwheel

La tarea de memoria de trabajo utilizada en este estudio no fue creada desde cero, sino que corresponde a una adaptación directa del paradigma denominado Letterwheel, planteado por Oberauer y Lin, en su experimento titulado “Presentación secuencial de pseudopalabras (nonwords)” [17], (ver Fig. 3.3.1). El objetivo de dicho experimento es evaluar la precisión de las representaciones verbales en la memoria de trabajo mediante un procedimiento de reproducción continua (*continuous reproduction*), que permite observar cómo varía el recuerdo en función de la similitud fonológica entre estímulos.

Esta tarea evalúa específicamente la capacidad de recuerdo con clave (cued recall) de ítems (palabras clasificadas como nonwords en este caso) asociados a localizaciones espaciales específicas, lo cual se alinea con los mecanismos de codificación contextual y recuperación por claves del Modelo de Interferencia.

Se empleó software personalizado desarrollado en Python para la presentación de las *nonwords*, basado directamente en los scripts proporcionados por el Profesor Klaus Oberauer y el Dr. Hsuan-Yu Lin, obtenidos mediante comunicación por correo electrónico. Estos scripts originales están diseñados, entre otras cosas, para simular el rendimiento de la memoria de trabajo en tareas cognitivas específicas, como la tarea Letterwheel y para estimar los parámetros del modelo a partir de datos comportamentales. Los parámetros centrales del IM tradicional, tal como se implementan en estos scripts y se describen en las publicaciones de Oberauer & Lin [17], incluyen aquellos relacionados con la precisión de la vinculación contexto-contenido (por ejemplo, el parámetro κ), las tasas de decaimiento u olvido, la capacidad o eficacia del foco de atención, y la susceptibilidad a la interferencia de diversas fuentes.

Para lograr esto, los autores diseñaron un conjunto de 19 pseudopalabras (nonwords) dispuestas en un bucle cerrado de relaciones de similitud fonológica, a diferencia de tareas con fonemas aislados, los investigadores crearon un anillo circular de similitud fonológica compuesto por 19 nonwords con estructura CVCV (consonante–vocal–consonante–vocal), donde cada estímulo difiere mínimamente de sus vecinos (un fonema y, a veces, una característica fonológica), lo cual permite mapear un espacio continuo de “ángulos de error fonológico”.

Esto genera una escala continua y homogénea de similitud, superando limitaciones de estudios previos como el de Hepner [67], donde los estímulos intermedios no correspondían a unidades lingüísticas válidas, generando sesgos en las respuestas hacia sonidos más familiares.

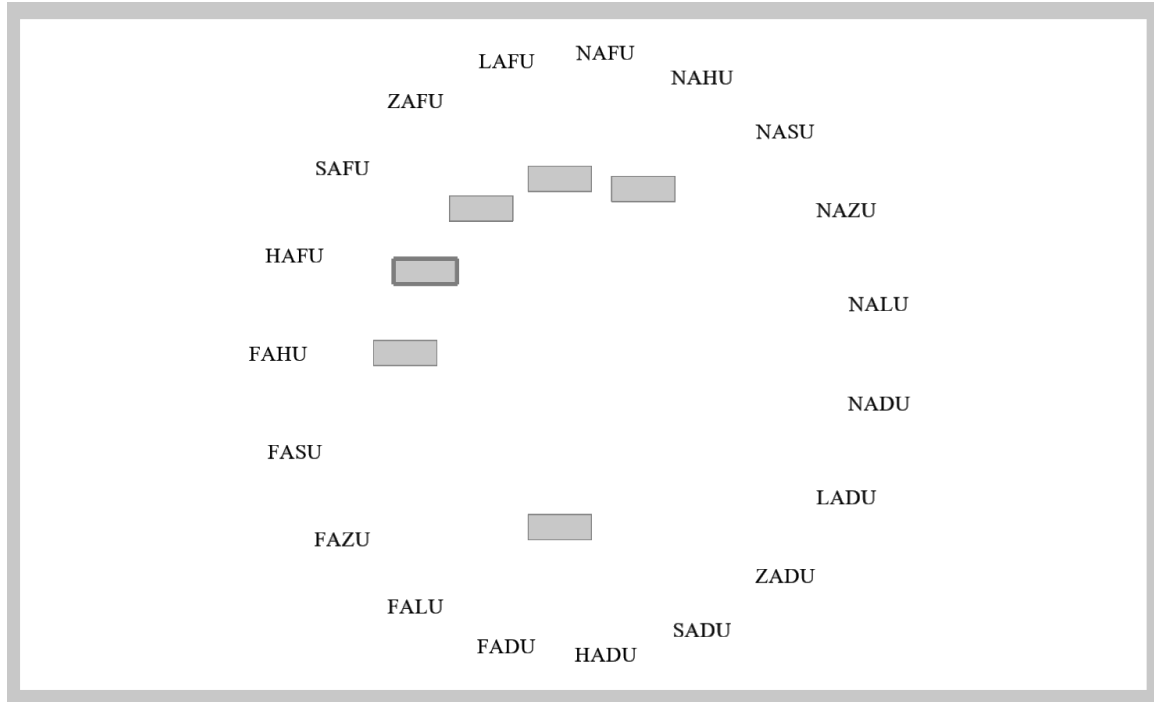


Figura 3.3.1: Ejemplo de la tarea de WM Letterwheel.

El uso directo de los scripts desarrollados por los propios autores del modelo asegura una implementación de alta fidelidad del IM tradicional. Esto no solo minimiza el riesgo de una interpretación errónea o una implementación incorrecta de las formulaciones matemáticas del modelo, sino que también proporciona una línea base sólida contra la cual se podrán comparar las modificaciones propuestas posteriormente.

De manera explicativa, el Letterwheel presenta al participante una secuencia aleatoria entre 1 y 6 nonwords, cada una ubicada en una posición específica de la pantalla. El objetivo es que el sujeto intente memorizar tanto el contenido como la ubicación de cada nonword. El número de palabras mostradas varía aleatoriamente entre intentos, y está relacionado con el nivel de dificultad de la tarea, siendo 6 el nivel más desafiante en términos de carga de memoria.

Posteriormente, se presenta un círculo de respuesta (ver Fig. 3.3.1), que contiene 19 nonwords dispuestas espacialmente en manera de círculo, entre las cuales se incluyen

las que fueron previamente presentadas durante la fase de codificación. Esta disposición permite al participante seleccionar la palabra que considera haber visto en una determinada ubicación previamente indicada.

Cabe destacar que:

- La posición que se indica para recordar no necesariamente corresponde al orden en el que las palabras fueron inicialmente mostradas.
- El participante realiza su selección mediante un clic sobre la *nonword* que considera correcta.
- Una vez seleccionada, la palabra no desaparece del círculo ni es resaltada en la casilla que se preguntó, sino que esta selección queda registrada internamente por el sistema.

En cuanto a la temporalidad del intento:

- El tiempo de exposición de cada nonword para memorización es aleatorio, pero no excede los 2 segundos por palabra.
- El tiempo para responder durante la fase de recuperación no está limitado pues depende del tiempo que el participante tarde en hacer clic en su elección.

Finalmente, la duración total de la sesión experimental varía en función de diversos factores, como la disposición del participante, la cantidad de palabras por intento y los tiempos de reacción individuales. No obstante, se estimó un tiempo promedio de aproximadamente 45 minutos, dentro de un rango general de entre 30 minutos y 1 hora.

3.3.2. Integración del Mindwave Neurosky

Para la adquisición continua de la actividad cerebral durante la ejecución de las tareas cognitivas, se empleó el dispositivo NeuroSky MindWave en combinación con la biblioteca NeuroPy desarrollada en Python. NeuroPy permite establecer una comunicación directa con el sensor EEG a través de Bluetooth, utilizando el protocolo estándar de 2.4 GHz, facilitando así la transmisión de datos en tiempo real al sistema de procesamiento.

Dado que la versión original de NeuroPy ofrece un conjunto limitado de funcionalidades, fue necesario adaptar el código fuente para ajustarlo a los requerimientos específicos del experimento. Las modificaciones incluyeron:

- Filtrado personalizado de métricas: se restringió la adquisición únicamente a las variables relevantes para el análisis tales como las bandas crudas correspondientes a theta, beta y alpha, eliminando señales innecesarias que pudieran introducir ruido o complejidad computacional.
- Sincronización temporal precisa: se integró un sistema de marcas de tiempo de alta resolución basadas en la función `time()` de Python, expresadas en segundos desde el inicio de la sesión. Esta sincronización permite correlacionar con exactitud los eventos experimentales (presentación de estímulos, tiempos de respuesta, etc.) con la señal EEG correspondiente.
- Estructuración del registro de datos: se diseñó una plantilla de salida en formato CSV estructurado, con campos separados para cada métrica relevante, incluyendo atención, meditación, valores de cada banda de frecuencia, y su correspondiente marca temporal. Esta organización de los datos facilita su posterior procesamiento en R y Python, permitiendo visualizaciones y modelado computacional.

Esta integración tecnológica posibilitó un monitoreo continuo y no intrusivo de la actividad cerebral de los participantes, con una latencia mínima, asegurando que las métricas registradas reflejaran con fidelidad la dinámica cognitiva durante las tareas de memoria de trabajo.

3.3.2.1. Procesamiento de la Señal EEG

Es importante reconocer que el dispositivo Neurosky Mindwave 1 realiza un preprocesamiento interno de la señal, que incluye ciertos niveles de filtrado. Sin embargo, los detalles específicos de estos algoritmos no son divulgados públicamente por el fabricante, pero se consigue una base en su manual y en el repositorio de NeuroPy [51, 60]. Esta naturaleza de "caja negra" del preprocesamiento inicial del Mindwave representa una limitación inherente al uso de este tipo de dispositivos comerciales de bajo costo en investigación.

Por esta razón, fue necesario el desarrollo y adaptación de scripts personalizados para establecer una comunicación efectiva con el dispositivo. Esto implicó la exploración y prueba de múltiples bibliotecas, especialmente en Python, como NeuroPy, y la implementación de rutinas de conexión vía Bluetooth que pudieran manejar errores comunes, latencias de respuesta y pérdida temporal de señal. Asimismo, se diseñaron mecanismos de control para sincronizar con precisión los registros EEG con los eventos del protocolo experimental, garantizando así la integridad de los datos recolectados. Todo este proceso requirió una fase de pruebas, depuración y documentación, dado que el soporte oficial del dispositivo es limitado y obsoleto, lo que aumentó la complejidad de su integración en un flujo de trabajo de adquisición y análisis de datos.

3.3.3. Implementación de la Tarea de Memoria Letterwheel

En cada intento, se presentaba una serie de estas palabras sin significado y se requería que el participante memorizara las palabras y su ubicación en la pantalla, posteriormente, se señalaba (sombreada) una parte específica de la pantalla donde había aparecido una palabra, y el participante debía seleccionar, utilizando el ratón, la palabra correcta de un conjunto de opciones que aparecían alrededor.

En el contexto de esta investigación, se adaptó dicha tarea para ser ejecutada en computadora y sincronizada con la recolección de datos EEG mediante el dispositivo NeuroSky MindWave. Durante su realización, los participantes debían memorizar secuencias de pseudopalabras que se presentaron de forma secuencial y, posteriormente, reproducir el orden correcto. Mientras tanto, se registraron ondas en las frecuencias de Alpha, Beta y Theta para luego procesarlas a los Niveles de Atención y Estrés mental.

Cada uno de los 22 participantes asistió a tres sesiones de recolección de datos de 45 minutos cada una, programadas en días diferentes para minimizar los efectos de la fatiga acumulada y obtener mediciones más estables. Se puso especial énfasis en mantener la consistencia del protocolo a lo largo de todas las sesiones y para todos los participantes. A continuación, se describe el protocolo utilizado por sesión:

a. Llegada y Explicación Inicial

A la llegada del participante, se le recibía y se le proporcionaba una breve explicación sobre las actividades de la sesión utilizando el siguiente guión: "Estoy investigando la actividad cerebral durante tareas de memoria. Por eso, necesito tu apoyo para participar en esta sesión. El objetivo es realizar una actividad relacionada con la memoria de trabajo, mientras registramos tu actividad cerebral con un instrumento no intrusivo, parece un cintillo. La duración aproximada de la sesión es de 45 minutos".

b. Revisión del Consentimiento Informado

Se revisaba la carta de consentimiento informado y si el participante estaba de acuerdo se procedía a firmar antes de la primera sesión. También se resolvían las dudas que el participante pudiera tener y se registraba dicha firma para mantener un control informativo de su participación.

c. Colocación del Dispositivo

El instructor ayudó al participante a colocar la diadema Mindwave 1, ajustandola cuidadosamente para asegurar un buen contacto del sensor en la posición Fp1 y del clip en el lóbulo de la oreja, garantizando tanto la comodidad del participante como la calidad de la señal.

d. Verificación de la Señal

Se verificó la correcta conexión del dispositivo vía Bluetooth con el ordenador y se comprobaba la calidad de la señal EEG a través de la interfaz escrita en Python (utilizando NeuroPy). Se procedía sólo si la señal era satisfactoria.

e. Encuesta Inicial

Se administró una encuesta pre-tarea (Ver anexo 2), para registrar el nivel de fatiga y el estado emocional. Esta herramienta tuvo como propósito controlar variables internas que pudieran influir significativamente en el desempeño cognitivo y en los patrones de actividad cerebral registrados mediante EEG. Las preguntas incluidas fueron formuladas siguiendo un formato estructurado, utilizando escalas de tipo Likert.

f. Explicación de la Tarea Letterwheel

Se explicó detalladamente las instrucciones de la tarea Letterwheel utilizando el siguiente guión: "En la actividad se te mostrará una serie de palabras sin significado y debes hacer el mayor esfuerzo por recordar, luego se te pedirá con un sombreado en una

parte específica de la pantalla la palabra que apareció ahí y debes seleccionarla con el mouse en un conjunto de palabras que estarán alrededor. No te preocupes si no recuerdas alguna, lo importante es que hagas el mejor esfuerzo posible".

Además, antes del inicio formal del experimento, se incorporó una fase inicial de familiarización compuesta por aproximadamente 10 intentos de prueba, con una duración total estimada de 5 minutos. Esta fase funcionaba como un período de calentamiento, en el cual se presentaba la tarea de manera idéntica a la versión real del experimento, incluyendo la presentación de las nonwords, su ubicación en pantalla, y la posterior selección desde el círculo de respuesta.

Durante esta etapa, se animaba al participante a realizar preguntas, aclarar dudas y familiarizarse tanto con la interfaz como con el procedimiento general. Esta estrategia permitió reducir posibles sesgos por desconocimiento del experimento y mejorar la calidad de los datos recolectados, ya que el participante entraba a la fase formal con una comprensión de la dinámica de la actividad.

g. Ejecución de la Tarea

El participante realizaba la tarea Letterwheel. Durante toda la ejecución, los datos EEG (ondas alpha, Beta, Theta) eran registrados continuamente por el script de Python y sincronizados con los eventos de la tarea (por ejemplo, inicio/fin de sesión, presentación de estímulos, respuestas).

h. Encuesta Final

Al finalizar la tarea Letterwheel, se administraba la encuesta post-tarea, idéntica a la encuesta inicial (Ver anexo 2), para comparar los datos antes y después de la prueba.

i. Retirada del Dispositivo y Agradecimiento

Para finalizar, se retiraba la diadema Mindwave 1 y se agradecía al participante por su tiempo y colaboración en el estudio.

3.3.4. Esfuerzos de Estandarización

Se tomaron medidas para asegurar la mayor consistencia procedural posible entre participantes y sesiones. Esto incluyó el uso de un guión estandarizado para las instrucciones, la monitorización de los tiempos de cada fase del experimento y, el mismo instructor (la autora de este Proyecto de Grado) en todas las sesiones.

El diseño multi-sesión implementado constituye una fortaleza metodológica significativa ya que permitir la recolección de datos en tres ocasiones distintas por participante no solo ayuda a evaluar la fiabilidad de las mediciones EEG y del rendimiento en la WM a lo largo del tiempo, sino que también abre la posibilidad de modelar efectos de aprendizaje o cambios en los niveles de atención y estrés entre sesiones. Al obtener datos en múltiples sesiones, se pueden lograr estimaciones más estables de la capacidad individual de memoria de trabajo y de su relación con la atención y el estrés, o incluso investigar cómo estas relaciones pueden variar, lo que enriquece la interpretabilidad del conjunto de datos construido.

3.4. Construcción del Dataset

La creación de un conjunto de datos estructurado y limpio fue un paso fundamental en el desarrollo de esta investigación, ya que constituyó la base sobre la cual se realizó el análisis cuantitativo posterior y la validación del modelo propuesto. Esta etapa no se limitó únicamente a la recopilación de datos, sino que abarcó un proceso

de planificación, ejecución, limpieza, y organización de los datos provenientes de múltiples fuentes, incluyendo los registros EEG generados por el dispositivo Neurosky Mindwave 1 y los resultados obtenidos durante la tarea LetterWheel.

3.4.1. Generación de Archivos de Registro

El script personalizado en Python fue diseñado para generar archivos de registro (.log), creando un archivo de registro general que compilaba los datos de todas las sesiones y participantes, así como archivos de registro individuales para cada uno de los 22 participantes, cubriendo sus 3 sesiones experimentales.

Estos archivos contenían múltiples campos de datos cruciales, incluyendo: identificador del participante, número de sesión, número de intentos, marcas de tiempo, eventos específicos de la tarea Letterwheel (por ejemplo, presentación del estímulo, clave de recuerdo, respuesta del participante), detalles del estímulo presentado, la respuesta emitida por el participante, el tiempo de respuesta (RT), la exactitud de la respuesta (correcta/incorrecta), y los valores sincronizados de potencia de las bandas alpha, Beta y Theta.

3.4.2. Conversión de Datos: De Registros Python a Formato RData

Para facilitar el análisis estadístico y el modelado, los datos contenidos en los archivos de registro generados por Python fueron convertidos al formato RData, nativo del entorno R. Esta conversión fue esencial para poder utilizar las capacidades de R, incluyendo la compatibilidad con los scripts de análisis del Modelo de Interferencia proporcionados por Oberauer y Lin [68].

La obtención de los archivos de registro ocurre de la siguiente manera: al finalizar la ejecución de cada sesión experimental, el sistema generaba un archivo de registro (log) en formato CSV, estructurado originalmente en 27 columnas. Este log base corresponde al diseño propuesto por Oberauer en sus experimentos previos con la tarea

LetterWheel, e incluye tanto información del intento como las respuestas del participante y la configuración espacial de los estímulos presentados. Las columnas contenidas en este log son:

ID, Session, Setsize, Target_index, Input_position, Output_position, Spatial_location, Index1 a Index6, Location1 a Location6, Text1 a Text6, Response, RT

Estas variables representan:

- El identificador del participante y la sesión correspondiente.
- El tamaño del set presentado (de 1 a 6 ítems).
- La ubicación del objetivo, su posición en pantalla, el índice de presentación y la posición de salida.
- Las nonwords y su disposición espacial, con campos específicos que indican tanto la ubicación como el contenido textual de cada estímulo.
- La respuesta seleccionada por el participante (Response) y el tiempo de reacción (RT).

Para integrar los datos fisiológicos, el presente proyecto extendió el log original mediante la adición de una columna número 28 con n puntos de muestreo de señales EEG recogidas durante cada intento. Cada punto de muestreo contiene 3 valores correspondientes a las señales EEG crudas por segundo para las bandas alpha, beta y theta. Esto dio lugar a logs considerablemente más grandes que los originales, ya que cada intento contenía en promedio decenas de registros fisiológicos de actividad cerebral.

Luego de ello, la conversión de los archivos de registro crudos a un objeto RData estructurado es un paso de preparación de datos crítico que consistió en preparar los datos para su análisis y modelado en el entorno estadístico R. Se diseñó un procedimiento de conversión a partir de los logs enriquecidos.

Este proceso fue implementado en un script encargado de:

- Leer el log CSV enriquecido, detectando la posición exacta de las columnas EEG a partir del valor de n.
- Calcular medidas agregadas de atención y estrés mental:
 - Atención promedio (*avg_attention*): se estimó a partir de la relación beta/theta, reconocida en la literatura como indicador del enfoque atencional.
 - Estrés promedio (*avg_stress*): se calculó con base en la relación alpha/beta, frecuentemente empleada como proxy de activación fisiológica asociada al estrés.
- Integrar estas métricas fisiológicas como nuevas columnas en el dataset final.
- Exportar los datos procesados a un archivo RData, manteniendo los 27 campos originales del log más las dos nuevas columnas *avg_attention* y *avg_stress*, resultando en un total de 29 columnas por intento.

Este archivo en formato RData sirvió como entrada directa para el posterior análisis estadístico y la ejecución del modelo computacional de memoria de trabajo, donde la conversión no sólo facilitó la compatibilidad con scripts de modelado previamente desarrollados por Oberauer y Lin, sino que también garantizó una estructura limpia, organizada y reproducible, facilitando la interpretación de los datos y su trazabilidad durante todo el flujo de la investigación.

3.4.3. Análisis Exploratorio de Datos para Identificación de Patrones

Una vez que el conjunto de datos estuvo en formato RData, se procedió a un análisis exploratorio de datos (*EDA por sus siglas en inglés*) utilizando tanto R como Python. Se emplearon paquetes como *dplyr* para la manipulación de datos y *ggplot2* en

R, junto con pandas, numpy, matplotlib y seaborn en Python para procesamiento y visualización gráfica.

El análisis inicial de los scripts en R fueron proporcionados por Oberauer y Lin [68], facilitando el punto de partida para comprender cómo se procesan los datos comportamentales de tareas como la Letterwheel y cómo se relacionan con su Modelo de Interferencia.

El objetivo principal de esta fase fue identificar patrones y rangos significativos que ofrecieran evidencia preliminar de relaciones predictivas entre los indicadores EEG de atención y estrés, y el rendimiento de la memoria de trabajo. Esta etapa fue esencial para sustentar empíricamente las modificaciones propuestas al Modelo de Interferencia de Oberauer.

En particular, se realizó un análisis sobre los indicadores fisiológicos extraídos del dispositivo NeuroSky (attention y stress). Estas variables fueron procesadas para su posterior modelado mediante una serie de transformaciones. Primero, se aplicó una transformación logarítmica natural ajustada (\log_{1p}) a la variable stress, con el objetivo de reducir su asimetría y estabilizar su varianza, dado que presentaba valores altamente concentrados en un rango bajo y esporádicos picos elevados.

Posteriormente, se procedió con la normalización Min-Max tanto de la atención como del estrés transformado. Este método es utilizado en análisis de datos y aprendizaje automático porque transforma los valores originales a un rango estándar de $[0, 1]$ mediante la Ec. 2.7.2. La elección de esta técnica respondió a la necesidad de estandarizar ambas variables en una escala común, de forma que se pudieran comparar directamente, evitar sesgos en los análisis estadísticos y permitir una integración armoniosa en modelos computacionales multivariados. Esta normalización fue clave para la generación de visualizaciones comparables entre sujetos.

Como parte de este análisis, se observó que la variable de atención, evaluada de forma aislada, presentaba una tendencia consistente a diferenciar niveles de desempeño

en las tareas de memoria. Esta observación motivó la comparación de tres enfoques: un modelo base (el IM), uno con las variables de atención y estrés (Modelo Neuro 2), y otro con sólo la atención como parámetro modulador (Modelo Neuro 1). Esta comparación fue útil para explorar el aporte individual y conjunto de los indicadores fisiológicos, enriqueciendo la interpretación de los resultados y la validación de la propuesta.

3.5. Desarrollo del Modelo Matemático Predictivo

3.5.1. Justificación para la Modificación del Modelo

Si bien el Modelo de Interferencia (IM) de Oberauer es una representación validada de los mecanismos de la memoria de trabajo, su formulación tradicional no incorpora explícitamente la influencia de las fluctuaciones individuales y momentáneas en estados cognitivos y afectivos como la atención y el estrés mental.

3.5.2. Formulación Matemática del Modelo Modificado

El modelo neuro propuesto, es una extensión del modelo original de Oberauer [17], con modificaciones que introducen la influencia de señales fisiológicas obtenidas en tiempo real mediante EEG, específicamente atención y estrés.

La función objetivo del modelo es calcular la probabilidad $P(x_t = k)$ de que el sujeto elija la opción k en el intento t , utilizando una combinación de rutas de recuperación directa e indirecta y modulaciones dinámicas basadas en variables cognitivas. En este contexto, la variable t representa el número de intentos dentro de una sesión experimental, por lo que toma valores enteros consecutivos según la cantidad de intentos presentados al participante, por ejemplo, si se administraron 40 intentos en total, entonces $t \in \{1,2,3,\dots,40\}$. Por su parte, k representa una de las posibles nonwords que el participante puede seleccionar como respuesta en cada intento. Dado que la tarea

Letterwheel presentaba 19 nonwords en cada rueda, los posibles valores de k corresponden al conjunto de nonwords mostradas en ese intento.

Esta formulación permite capturar cómo varía la probabilidad de seleccionar cada opción en función del intento específico y del estado cognitivo del participante en ese momento.

3.5.2.1. Presentación de las Ecuaciones

El desarrollo del modelo modificado parte de las ecuaciones fundamentales del Modelo de Interferencia (IM) propuesto por Oberauer, tal como se describe en la literatura y se implementan en los scripts originales utilizados para el entrenamiento y ajuste del modelo. En su formulación tradicional, el IM describe cómo se distribuye la activación de los ítems almacenados en memoria ante una clave contextual (por ejemplo, una ubicación), y cómo estas activaciones compiten entre sí durante el proceso de recuperación.

Por ejemplo, la activación total de un candidato de respuesta x dada una clave contextual L_θ en el IM tradicional (Ec. 2.3.1.2), donde A_c es la activación basada en el contexto (que depende de la precisión de la vinculación, k), A_a es la activación independiente del contexto, y A_b es el ruido de fondo (que puede depender del tamaño del conjunto n).

Las modificaciones propuestas en este trabajo integran las medidas fisiológicas obtenidas mediante EEG, en particular, la atención, y el estrés, para ser los moduladores de los parámetros del modelo.

Concretamente:

- **Modulación del Peso de Recuperación Directa por la Atención:**

Se propone que la atención del participante influye en la ruta de recuperación preferida dentro de la activación basada en la pista A_c , específicamente, se plantea que el parámetro w_{direct} , que en el modelo original representa la proporción de activación asignada a la recuperación directa R1 frente a la recuperación mediada R2, puede ser modulado por el nivel de atención.

Así, se introduce un nuevo parámetro modulado:

$$w'_{direct} = f(w_{direct}, attention) \quad (3.5.2.1)$$

donde f es una función que busca favorecer rutas directas, es decir, se basa más en la localización precisa del ítem objetivo.

- **Modulación del Ruido o Interferencia por el Estrés:**

De forma análoga, se postula que el estrés mental puede afectar la dispersión de $A_c(x|L_\theta)$, a través de los parámetros s_{space} y s_{time} que controlan qué tan rápidamente decae la activación en función de la distancia espacial y temporal, respectivamente.

Así, se propone:

$$s' = g(s, stress) \quad (3.5.2.2)$$

donde g es una función que busca producir un mayor grado de dispersión y, por tanto, más interferencia o errores de recuperación.

Las ecuaciones modificadas se presentarían formalmente, definiendo cada nuevo parámetro o función de modulación y justificando su inclusión basándose en los hallazgos exploratorios y la fundamentación teórica.

3.5.2.2. Variables y parámetros

Dentro de los parámetros que se integran para entrenar al modelo están:

- w_{direct} : peso relativo de la ruta directa frente a la indirecta.
- s_{space} , s_{time} : pesos de decaimiento en rutas directa e indirecta.
- $attention_t$, $stress_t$: valores fisiológicos obtenidos del EEG en el intento t .

3.5.2.3. Modulación Fisiológica de Parámetros

Con el objetivo de adaptar el comportamiento del modelo computacional a los estados mentales reales de los participantes durante la ejecución de la tarea Letterwheel, fue necesario introducir mecanismos de modulación dinámica sobre los tres parámetros centrales del modelo de memoria de trabajo: w_{direct} , s_{space} y s_{time} .

Estos representan, respectivamente, la proporción de atención dirigida al objetivo (w_{direct}) y los coeficientes de interferencia asociados a elementos distractores (s_{space} y s_{time}). Dichas transformaciones permiten que el modelo reaccione en función de los niveles de atención y estrés mental, captados en tiempo real con el dispositivo NeuroSky MindWave.

Así, en el caso de w_{direct} , se empleó la siguiente transformación:

$$w'_{direct_t} = \min(\max(w_{direct} \cdot (1 + attention_t), 0), 1) \quad (3.5.3.1)$$

Esta ecuación se encarga de darle el peso a la ruta directa frente a la ruta indirecta. Se toma el valor base de atención w_{direct} y se multiplica por un factor que depende del

nivel de atención medido en el intento t . Se utiliza la forma $(1 + attention_t)$ para garantizar que la atención nunca se anule completamente y que incluso niveles bajos de atención tengan un impacto atenuado pero positivo en el proceso. La estructura $\min(\max(...))$ se utiliza para limitar el resultado a un rango válido entre 0 y 1, ya que como w_{direct} representa una proporción (de atención dirigida al estímulo), no puede exceder ese rango sin perder su interpretación teórica.

Por otro lado, los parámetros s_{space} y s_{time} , que son los asociados a los pesos de decaimiento de la rutas, fueron modificados mediante las siguientes expresiones:

$$s'_{space_t} = \max\left(\frac{s_{space}}{1+5 \cdot stress_t}, \varepsilon\right) \quad (3.5.3.2)$$

$$s'_{time_t} = \max\left(\frac{s_{time}}{1+5 \cdot stress_t}, \varepsilon\right) \quad (3.5.3.3)$$

Aquí, el objetivo fue modelar la influencia negativa del estrés sobre la capacidad de recuperación de la información, donde un mayor nivel de estrés aumenta el denominador de la fracción, reduciendo así el valor de s_{space} y s_{time} , lo que representa un mayor decaimiento de los elementos almacenados en memoria de trabajo. La constante ε , de valor muy pequeño, se incorpora como protección contra divisiones por cero o valores extremadamente bajos que puedan generar inestabilidad numérica en el entrenamiento.

El valor 5 utilizado como factor multiplicador del estrés en el denominador de las Ec. 3.5.3.2 y Ec. 3.5.3.3, se seleccionó tras realizar pruebas empíricas con distintos coeficientes, su elección responde a la necesidad de amplificar de forma controlada el impacto del estrés, sin provocar una reducción excesiva de los parámetros ni pérdida de estabilidad en el modelo. Coeficientes menores (como 1 o 2) generaban variaciones muy

leves, poco sensibles a los cambios reales en los datos fisiológicos, mientras que valores mayores a 5 producían respuestas no realistas. De esta forma, el factor 5 se adopta como un valor intermedio razonable que permite una modulación perceptible, diferenciada entre individuos y coherente con la dinámica cognitiva esperada.

Cabe destacar que, a diferencia de w_{direct_t} , los parámetros s_{space} y s_{time} no están acotados superiormente porque no representan proporciones sino factores de interferencia con un comportamiento decreciente y restringirlos a un máximo artificial. De esta manera, se garantiza una respuesta diferenciada entre individuos con diferentes niveles de estrés, preservando la capacidad del modelo para adaptarse a condiciones cognitivas reales.

Estas transformaciones matemáticas fueron implementadas en el código fuente como funciones dinámicas que reciben como entrada los valores $attention_t$ y $stress_t$ estimados en cada intento t , permitiendo una integración directa entre la señal fisiológica en tiempo real y el procesamiento computacional del modelo.

3.6. Proceso de Entrenamiento

El modelo neuro 2 y el modelo neuro 1 constituyen una extensión del modelo base, introduciendo la modulación dinámica de parámetros en tiempo real mediante señales fisiológicas obtenidas del dispositivo NeuroSky MindWave. Específicamente, se integraron las variables de atención y estrés, obtenidas del EEG en formato de variables normalizadas según la Ec 2.7.2. Estas variables influyen directamente sobre los parámetros w_{direct} , s_{space} y s_{time} , de la siguiente manera:

- w_{direct} (peso de la ruta directa): se aumenta con la atención, simulando que una mayor concentración favorece el acceso directo a la información.

- s_{space} y s_{time} (decaimiento en las rutas de recuperación): se reducen con el estrés, modelando el efecto negativo de éste sobre la precisión de recuperación.

Durante el entrenamiento, el modelo lee intento por intento los valores empíricos de atención y estrés de cada sujeto. En cada prueba, se reescala el estrés, se ajustan dinámicamente los parámetros mencionados, y luego se ejecuta el cálculo de probabilidades de respuesta mediante la combinación de rutas directa e indirecta. Estas rutas se modelan con funciones exponenciales de decaimiento, que dependen de distancias temporales entre ítems.

El resultado de cada intento en el entrenamiento es un vector de probabilidades sobre las posibles opciones de respuesta, de donde se extrae la probabilidad asignada a la opción elegida por el participante.

3.8 Diagrama General del Proceso de Diseño e Implementación del Estudio

La Fig. 3.8, ilustra de manera integral el flujo metodológico seguido en el desarrollo de esta investigación, el cual busca diseñar un modelo matemático para predecir el rendimiento de la memoria de trabajo a partir de señales fisiológicas de atención y estrés mental registradas mediante EEG.

Esta representación gráfica agrupa las etapas fundamentales del estudio, desde la planificación experimental hasta la evaluación del modelo propuesto, ofreciendo una visión global del diseño implementado. En primer lugar, se muestra la fase de recolección de datos, que implicó el uso del dispositivo NeuroSky Mindwave para obtener medidas en tiempo real de los niveles de atención y estrés de los participantes durante tareas de tipo Letterwheel cuyos datos fueron almacenados, preprocesados y organizados cuidadosamente para permitir su posterior análisis computacional.

En una segunda etapa, se describe el ajuste de los modelos, en la cual se implementaron variantes del Modelo de Interferencia de Oberauer, realizando pruebas de sensibilidad sobre parámetros del modelo, como w_{direct} , s_{space} y s_{time} , con el fin de evaluar el impacto de los niveles de atención y estrés sobre el mecanismo de recuperación de información en memoria de trabajo. La calibración de estos parámetros se llevó a cabo mediante procedimientos de optimización basados en máxima verosimilitud, utilizando medidas de desempeño empírico como referencia. Este proceso permitió adaptar la estructura del modelo teórico a los datos fisiológicos observados, facilitando una representación más precisa del comportamiento cognitivo individual.

Finalmente, se representa la fase de evaluación del modelo, en la que se comparan las predicciones generadas con los datos reales de desempeño de los participantes. En este caso, la métrica utilizada para los análisis fueron las matrices de confusión junto con sus respectivas derivaciones estadísticas. El diagrama no solo resume las decisiones metodológicas adoptadas, sino que también facilita la comprensión del flujo de trabajo.

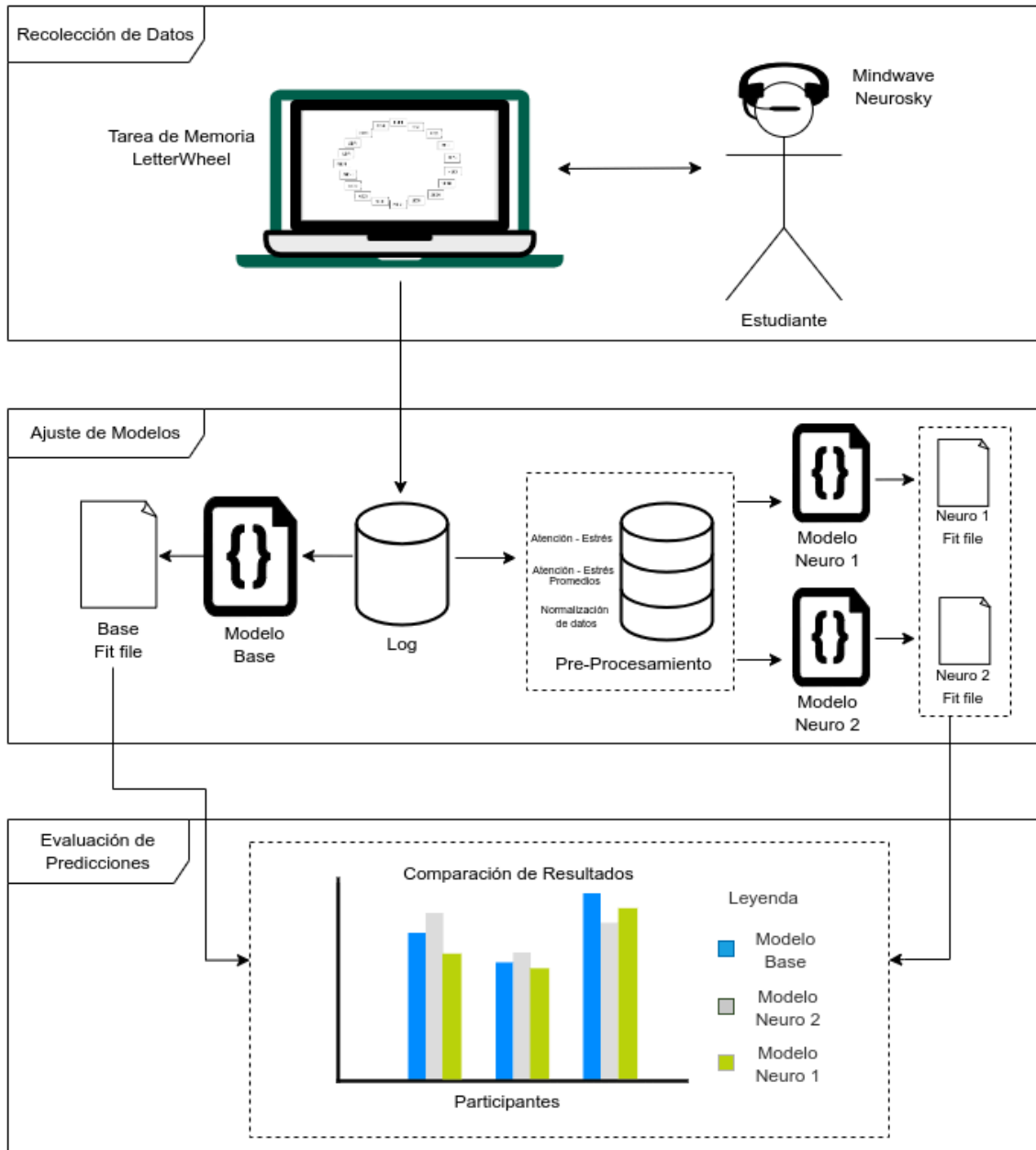


Figura 3.8: Diagrama general del proceso de diseño e implementación del estudio.

Capítulo IV - Resultados

4.1. Participantes

En esta sección se presentan los datos obtenidos a partir de la participación de los 22 estudiantes universitarios seleccionados conforme a los criterios previamente establecidos. Los participantes presentaron una edad promedio de 23 años, conformándose el grupo por 5 estudiantes de género femenino y 17 de género masculino. En cuanto al nivel académico, la muestra incluyó estudiantes de la carrera de Ingeniería de Sistemas de la Universidad de Los Andes desde el primer semestre hasta los últimos semestres de la carrera.

Como parte del protocolo experimental, cada participante completó una encuesta antes y después de las tareas cognitivas, las cuales se describen a continuación:

4.1.1. Encuesta Inicial y Final

Antes de iniciar las tareas experimentales, los participantes respondieron un cuestionario breve que permitió registrar variables subjetivas relevantes como el nivel de fatiga percibida y el estado emocional predominante. Estos factores son reconocidos en la literatura [69] como moduladores potenciales del rendimiento cognitivo, especialmente en tareas que requieren atención sostenida y memoria de trabajo. Su inclusión en el protocolo permite contextualizar las diferencias individuales en los datos recogidos y constituye un insumo valioso para el análisis exploratorio de covariables.

Una vez finalizada la sesión, se aplicó nuevamente un instrumento similar, con el objetivo de evaluar el impacto subjetivo de la tarea en los participantes. Se buscó identificar posibles cambios en la percepción de fatiga, tensión o carga mental, inducidos por la exposición prolongada a estímulos cognitivos bajo monitoreo EEG. Esta comparación pre y pos permite no solo observar el efecto inmediato del protocolo

experimental, sino también validar que la carga cognitiva administrada fue suficiente para generar una respuesta percibida.

Para su cuantificación, se empleó una escala tipo Likert de cinco puntos, de acuerdo con la propuesta de Rensis Likert, ampliamente utilizada para la medición de percepciones y estados subjetivos en contextos de investigación psicológica, educativa y de neurociencia aplicada. Esta escala permite capturar gradaciones en las respuestas, facilitando un análisis más fino de los datos.

4.1.1.1. Niveles de Fatiga Pre sesión y Pos sesión

Para evaluar el efecto de las sesiones experimentales sobre el nivel subjetivo de fatiga, se compararon los estados pre y pos de cada participante en tres momentos distintos. A continuación en la Fig. 4.1.1.1, se presentan las variaciones observadas.

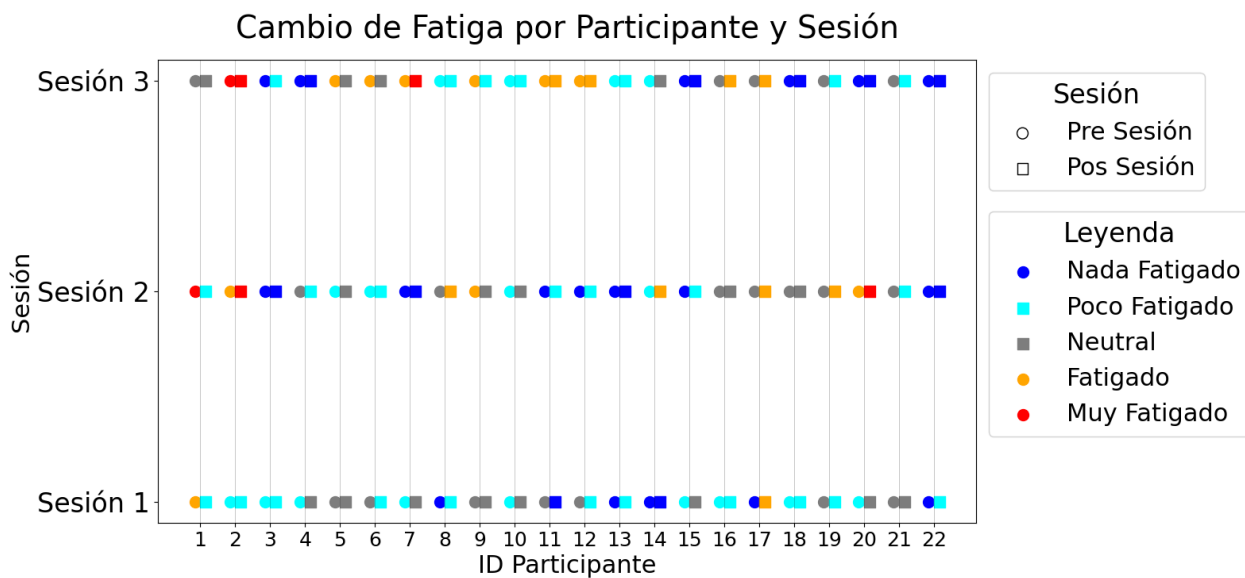


Figura 4.1.1.1: Cambio de fatiga por participante y por sesión.

Durante la primera sesión, la mayoría de los participantes partieron de estados de poco fatigado a nada fatigado y se observaron los siguientes patrones destacados:

Sesión 1: tendencia a la mejora leve.

- 11 de 22 participantes (50%) mostraron una mejoría en su nivel de fatiga tras la sesión. Por ejemplo, el participante 1 pasó de Fatigado a Poco Fatigado, y el 11 de Neutral a Nada Fatigado.
- 9 participantes (41%) permanecieron en el mismo nivel pre y pos.
- Solo 2 casos (9%) reflejaron un empeoramiento, como el ID 17 que pasó de Nada Fatigado a Fatigado.

Esto sugiere un efecto leve positivo general en la percepción de fatiga, con tendencia a estabilización o mejora, especialmente entre quienes comenzaron en niveles más altos.

Sesión 2: variabilidad alta y cambios bidireccionales.

- 9 personas mejoraron su nivel de fatiga. Ejemplo: el ID 1 pasó de Muy Fatigado a Poco Fatigado.
- 7 personas empeoraron, como el ID 8 (de Neutral a Fatigado) y el ID 20 (de Fatigado a Muy Fatigado).
- 6 participantes no mostraron cambios.

Este resultado indica que el efecto de la sesión fue más heterogéneo, posiblemente condicionado por el estado inicial. Algunos mejoraron drásticamente mientras que otros experimentaron mayor fatiga.

Sesión 3: disminución de la mejora.

- 6 participantes mejoraron, como el ID 5 y 6, que pasaron de Fatigado a Neutral.
- 5 empeoraron, como el ID 16 y 17, que pasaron de Neutral a Fatigado.
- 11 personas permanecieron sin cambios.

Esto sugiere que hacia la tercera sesión, el efecto de intervención podría estar disminuyendo en impacto o que la fatiga acumulada a lo largo de los días podría estar modulando los resultados.

A lo largo de las tres sesiones se observó una evolución distinta en la percepción de fatiga por parte de los participantes. En la primera sesión, la mayoría reportó una mejora en su nivel de fatiga, lo que sugiere un efecto positivo inicial de la actividad o del entorno experimental. Sin embargo, en la segunda sesión los resultados fueron más variados: algunos mejoraron, pero otros presentaron un aumento en la fatiga.

En la tercera sesión, se notó una tendencia a la estabilidad: varios participantes mantuvieron su nivel de fatiga y se encontraron menos casos de mejora. Esto podría indicar una adaptación al experimento o una posible acumulación de cansancio con el paso del tiempo.

En conjunto, los datos muestran que el nivel de fatiga es variable destacando la importancia de combinar medidas subjetivas con indicadores objetivos (como atención o estrés) para tener una visión más completa del estado mental de cada persona.

4.1.1.2. Niveles de Ánimo Pre sesión y Pos sesión

Los datos de la Fig. 4.1.1.2, muestran la evolución del estado de ánimo de los participantes en tres sesiones, con registros previos y posteriores a cada evaluación.

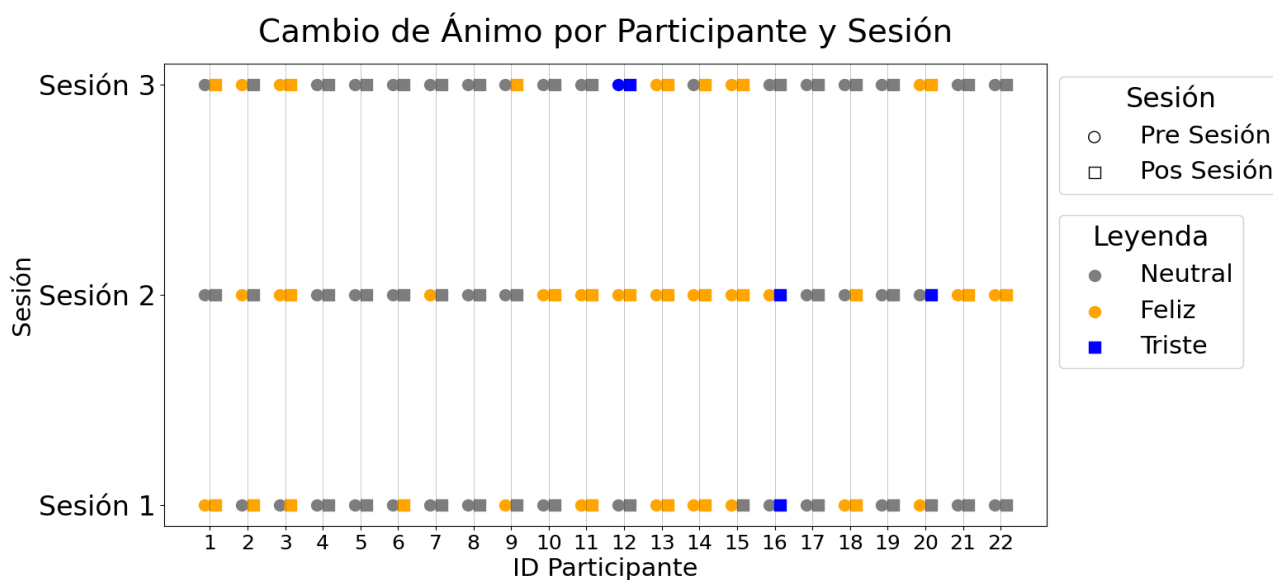


Figura 4.1.1.2: Cambio de ánimo por participante y por sesión.

En general, se observa una estabilidad emocional predominante, aunque con algunos cambios sutiles que vale la pena destacar:

Sesión 1: la mayoría de los participantes se mantuvo en estado neutral tanto antes como después de la sesión. No obstante, se observan varios casos de mejora, especialmente de neutral a feliz, lo cual podría estar asociado a una experiencia positiva durante la actividad. Solo un caso reportó un descenso (neutral a triste), indicando que en general la intervención no generó afectación emocional negativa.

Sesión 2: aquí se presenta una ligera tendencia hacia el descenso del ánimo, con algunos casos de cambio de feliz a neutral y dos casos aislados de neutral/feliz a triste. Esto podría estar relacionado con el cansancio acumulado o variaciones individuales, aunque la mayoría permaneció estable.

Sesión 3: en esta sesión reaparece un patrón más equilibrado ya que hay algunos casos de mejora (neutral a feliz), pero también se mantiene una fuerte proporción de estados neutros. Destaca que el único caso de estado “triste” (participante 12) se

mantuvo sin cambio, lo que indica persistencia de un bajo estado de ánimo en ese individuo.

En conjunto, los datos reflejan una tendencia a la estabilidad emocional con algunas fluctuaciones leves y mayormente positivas, además, las mejoras observadas (principalmente de neutral a feliz) podrían indicar que las actividades realizadas en las sesiones fueron percibidas de forma estimulante o placentera por varios participantes.

4.2. Dataset

Durante la ejecución de la tarea experimental de memoria de trabajo tipo Letterwheel, cada participante realizaba las tareas cognitivas mientras portaba el dispositivo de registro Neurosky Mindwave. En este contexto, se implementó un sistema de captura de datos en tiempo real que permitió recolectar simultáneamente información del desempeño en la tarea y métricas fisiológicas del estado mental del participante, específicamente las bandas de frecuencia alpha (α), beta (β) y theta (θ).

La información capturada fue almacenada de forma continua en archivos CSV estructurados, generando así el primer conjunto de datos *log*. Este archivo contenía, por cada intento, variables como: ID del participante, número de sesión, tamaño del conjunto (SetSize), posición objetivo (Target_index), posición ingresada por el participante (Input_position), posición reportada por el sistema (Output_position), ubicación espacial del estímulo, entre otras asociadas a la dinámica de presentación y respuesta del estímulo, incluyendo tiempos de reacción (RT) y selección textual o espacial (TextN, LocationN, IndexN).

Simultáneamente, por cada segundo de reloj, se registraron los valores numéricos de las señales EEG en las bandas α , β y θ . Esta sincronización temporal entre eventos cognitivos y actividad cerebral permitió asociar la dinámica neurofisiológica con el rendimiento conductual en la tarea.

A partir de este primer registro bruto, se generaron cuatro versiones derivadas del dataset:

1. Dataset sin señales EEG: se extrajeron únicamente las variables conductuales relacionadas con el rendimiento en la tarea (Letterwheel) para análisis exploratorios independientes del estado fisiológico.
2. Dataset con métricas derivadas de atención y estrés: empleando ecuaciones matemáticas previamente descritas en el Capítulo 3, se calcularon índices representativos de los niveles de atención y estrés a partir de las combinaciones de las bandas EEG registradas.
3. Dataset de promedios por intento: para facilitar los análisis agregados y reducir el ruido de variabilidad, se calcularon los promedios de atención y estrés por cada ejercicio realizado, permitiendo una representación más estable del estado cognitivo por unidad de tarea.
4. Dataset final normalizado: todas las variables fueron estandarizadas mediante normalización (Ec. 2.7.2), permitiendo la comparación homogénea entre sujetos y facilitando la implementación de modelos de análisis predictivo.

Este proceso de construcción progresiva del dataset no solo garantiza una trazabilidad clara desde los datos originales hasta las matrices de análisis, sino que también permite preservar tanto la riqueza temporal de las señales EEG como la precisión del comportamiento observable durante la tarea cognitiva.

4.2.1. Análisis Exploratorio de Datos

Para obtener una comprensión preliminar de las variables utilizadas en el modelo, se realizó un análisis exploratorio de datos (*EDA, por sus siglas en inglés*) centrado en los niveles promedios de atención y estrés registrados por los participantes. Esta etapa

permitió identificar patrones generales, evaluar la calidad de los datos, y tomar decisiones sobre transformaciones necesarias para el posterior análisis computacional.

Inicialmente, se graficaron las distribuciones (Fig. 4.2.1) de los valores promedios de atención y estrés, en ambos casos, se procedió a normalizar los datos con el objetivo de facilitar su comparación y controlar las diferencias de escala entre las variables.

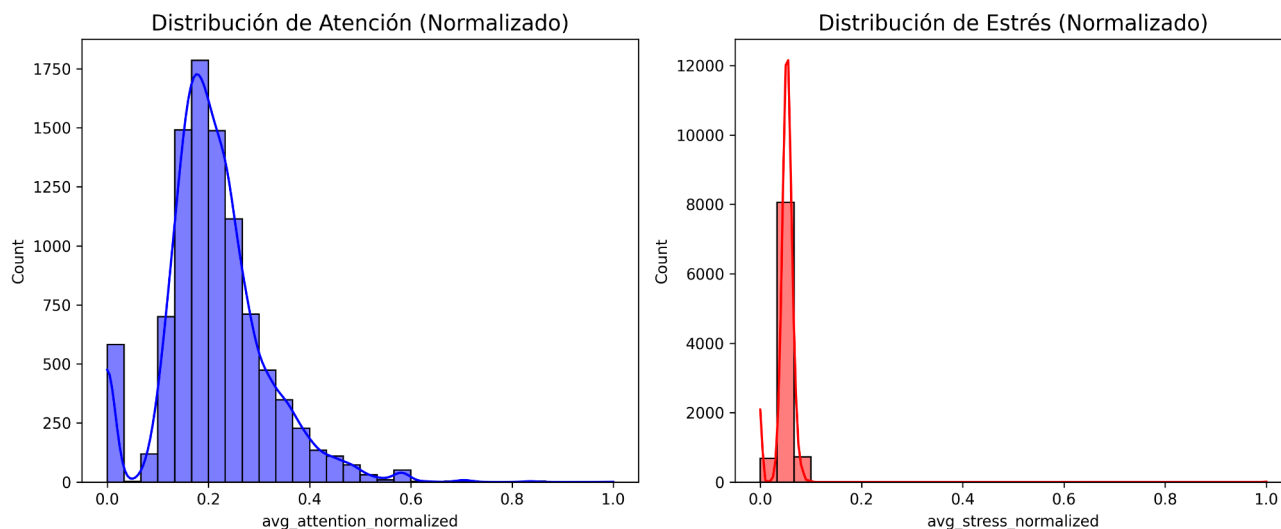


Figura 4.2.1: Distribuciones de las variables atención (izq) y estrés (der).

En la distribución de la atención normalizada, se observó una clara asimetría positiva (sesgo a la derecha), donde la mayor parte de los valores se concentraron entre 0.1 y 0.4, con un pico cercano a 0.2. Esta distribución indica que la mayoría de los participantes mantuvieron niveles bajos a moderados de atención durante las sesiones, con muy pocos casos alcanzando niveles cercanos al máximo. La forma de la curva sugiere variabilidad significativa, lo cual convierte a la atención en una variable potencialmente útil.

Dado este hallazgo, se justifica la incorporación de un modelo entrenado con solo la variable de atención (modelo neuro 1) para el análisis comparativo. Este modelo fue construido con el objetivo de evaluar el poder predictivo individual de la atención y contrastarlo frente al modelo original (modelo base) y al modelo con atención y estrés

(modelo neuro 2). Esta inclusión permite distinguir con mayor claridad el aporte único de la atención en la predicción del rendimiento, y verificar si esta señal, por sí sola, podría ofrecer resultados competitivos o incluso superiores en ciertos escenarios.

En contraste, la distribución del estrés normalizado presentó una concentración mucho más pronunciada en un rango reducido (entre 0.05 y 0.1), evidenciando una baja dispersión de los datos, esta configuración podría interpretarse como una homogeneidad en los niveles de estrés entre participantes. Esta concentración extrema puede reducir el poder explicativo del estrés como variable predictiva si no se aplican técnicas adicionales de transformación o modelado. En conjunto, los datos sugieren que, aunque casi todos los estudiantes tienen niveles similares de estrés, sus niveles de atención sí varían considerablemente.

De la misma manera, siguiendo el análisis en cuanto a la distribución de clases, se observó un desequilibrio moderado en las etiquetas de rendimiento utilizadas como variable objetivo. La clase correspondiente a la clase 0, obtuvo sobre representación en comparación con la clase 1, este desequilibrio tiene implicaciones directas sobre el entrenamiento de modelos predictivos, ya que puede sesgar la predicción hacia la clase mayoritaria y reducir la sensibilidad del modelo para detectar casos minoritarios. Para mitigar este efecto, se consideraron métricas de evaluación más robustas como la sensibilidad específica por clase y el F1-score, que permiten un análisis más equitativo del desempeño.

Por otro lado, el análisis de correlación entre variables predictoras reveló una relación positiva media entre los niveles de atención y la memoria de trabajo, mientras que la correlación entre estrés y rendimiento fue muy baja. Esta correlación directa de la atención sugiere que esta variable, aunque limitada, puede tener un valor moderado como predictor individual del desempeño en tareas de memoria de trabajo, lo cual se alinea con los resultados observados en el modelo entrenado exclusivamente con esta variable. En cambio, el estrés mostró una distribución homogénea y escasa variabilidad,

lo cual también se reflejó en su mínima relación estadística con la variable objetivo. Estos hallazgos apoyan la decisión de tratar a la atención como un modulador dinámico relevante dentro del modelo computacional, mientras que el estrés podría requerir una transformación no lineal o técnicas adicionales para capturar su posible influencia indirecta.

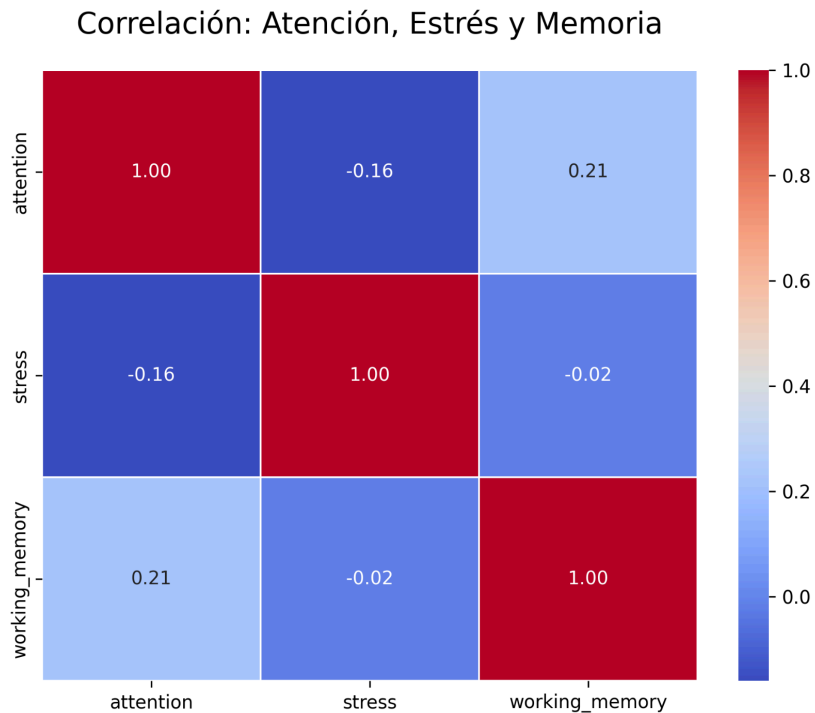


Figura 4.2.2 Correlación entre las variables de memoria, atención y estrés

4.3. Rendimiento

Para evaluar el rendimiento, se analizaron sus métricas utilizando los valores derivados de la matriz de confusión, tales como la exactitud, la precisión, la sensibilidad (recall), la especificidad y el F1 score. Estas métricas permiten no solo evaluar la eficacia general del modelo, sino también su comportamiento frente a clases desbalanceadas o errores tipo I y II.

4.3.1. Matriz de Confusión de Cada Modelo

A continuación, se presentan las matrices de confusión correspondientes a los modelos evaluados. Estas matrices permiten observar el desempeño de los clasificadores al mostrar la distribución de verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos. Cada matriz resume la capacidad del modelo para predecir correctamente las clases y detectar los errores cometidos en el proceso de clasificación facilitando una interpretación clara y comparativa del comportamiento de los modelos.

4.3.1.1. Modelo Base

A continuación, se presenta la matriz de confusión del Modelo Base:

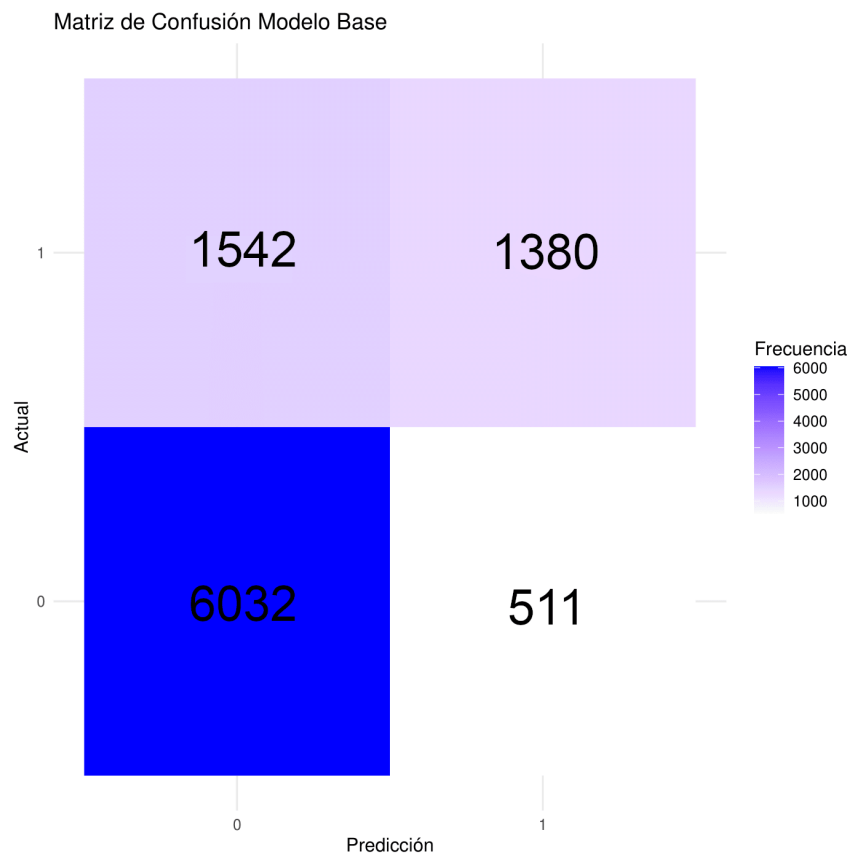


Figura 4.3.1.1: Matriz de confusión del Modelo Base.

Este modelo está muy inclinado a predecir que un intento será correcto (clase 0). Predice acertadamente 6032 intentos correctos, pero se equivoca en 1542 intentos que realmente fueron incorrectos (falsos positivos). Por otro lado, identifica correctamente 1380 intentos incorrectos. Su alta sensibilidad (92.19 %) y exactitud global (78.31 %) lo hacen una buena línea base. Sin embargo, su baja especificidad (47.23 %) sugiere una tendencia a sobreestimar la ocurrencia de intentos correctos, lo que puede ser problemático si es importante detectar con precisión cuándo ocurren errores.

4.3.1.2. Modelo Neuro 2

A continuación, se presenta gráficamente la matriz de confusión del Modelo Neuro 2 (modelo con las variables de atención y estrés):

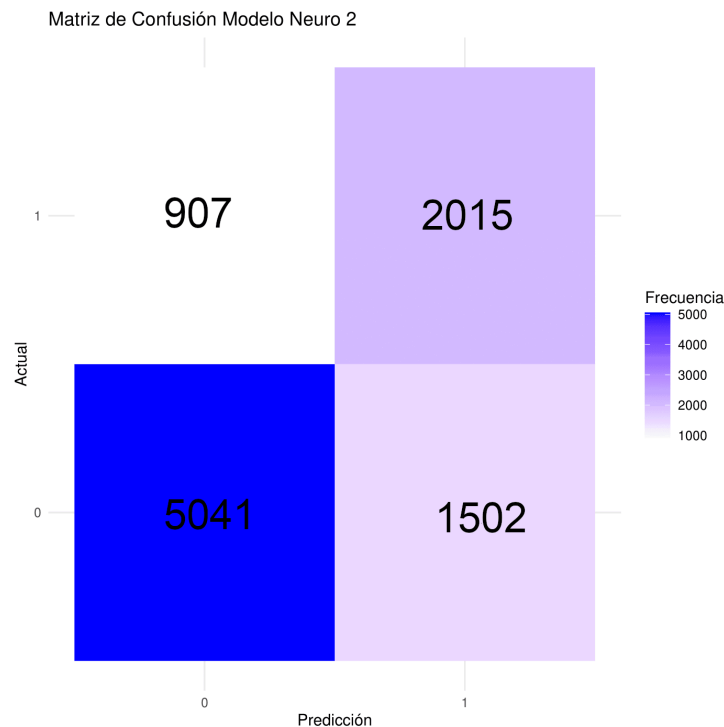


Figura 4.3.1.2: Matriz de confusión del Modelo Neuro 2.

Este modelo también acierta en 5041 intentos correctos, pero reduce significativamente los errores de sobreestimación: solo 907 intentos incorrectos fueron clasificados como correctos. Además, identifica mejor los intentos fallidos, con 2015 predicciones correctas de intentos incorrectos, lo que se refleja en su mejor especificidad (68.96 %). Este es un modelo más equilibrado, que mejora el reconocimiento de errores sin sacrificar demasiado la identificación de aciertos.

4.3.1.3. Modelo Neuro 1

A continuación, se presenta gráficamente la matriz de confusión del Modelo Neuro 1, con solo los niveles de atención como parámetro. Este modelo fue construido de forma adicional con el objetivo de enriquecer el análisis comparativo, permitiendo evaluar de manera más específica el impacto aislado de la atención sobre el rendimiento de la memoria de trabajo.

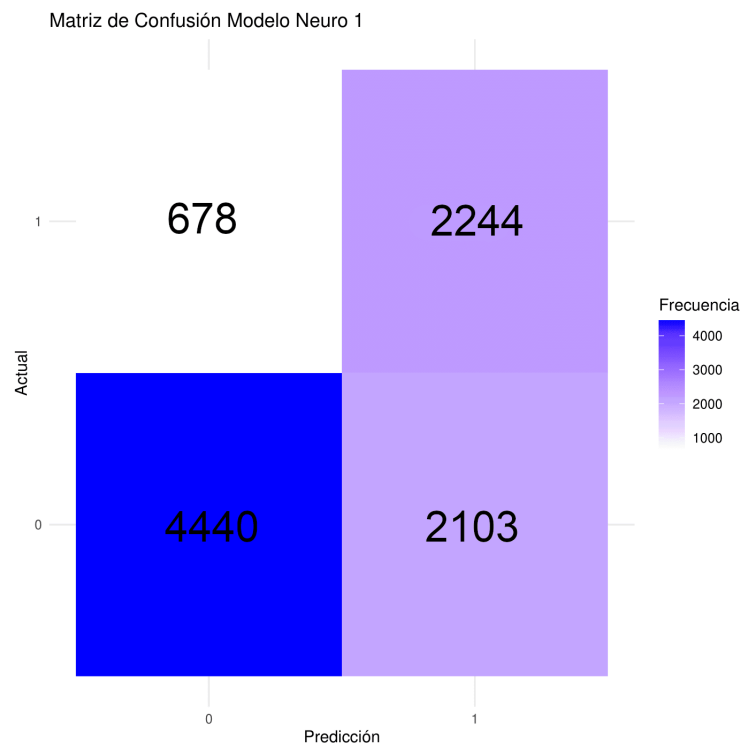


Figura 4.3.1.3: Matriz de confusión del Modelo Neuro 1.

Este modelo predice acertadamente 4440 intentos correctos, y logra reducir aún más los errores de sobreestimación: solo 678 intentos fallidos fueron clasificados incorrectamente como aciertos. Además, detecta correctamente 2244 intentos incorrectos, mostrando la mayor especificidad (76.80 %) de todos los modelos. Si bien sacrifica sensibilidad (67.86 %), es especialmente útil en contextos donde es más importante no asumir erróneamente que un intento será exitoso.

4.3.2. Tabla Comparativa de Modelos

En la Tabla 4.3.2.1 se comparan tres modelos que predicen el rendimiento de la memoria de trabajo según la correctitud de los intentos del Letterwheel basado en diferentes variables:

- Modelo Base: es el Modelo de Interferencia original.
- Modelo Neuro 2: modelo extendido que usa tanto las variables de atención como estrés.
- Modelo Neuro 1: modelo extendido que usa únicamente la variable de atención.

Métrica	Modelo Base	Modelo Neuro 2	Modelo Neuro 1	Descripción
Exactitud	78.31%	74.55%	70.62%	Proporción total de predicciones correctas.
Precision	79.64%	84.75%	86.75%	De todas las predicciones positivas, ¿Cuántas fueron realmente positivas?
Sensibilidad (recall)	92.19%	77.04%	67.86%	Capacidad para identificar correctamente los positivos reales (Clase 0).
Especificidad	47.23%	68.96%	76.80%	Capacidad para identificar correctamente los negativos reales (Clase 1)

F1 score	85.46%	80.71%	76.15%	Promedio armónico entre precisión y recall. Evalúa el equilibrio entre ambos.
----------	--------	--------	--------	-------------------------------------------------------------------------------

Tabla 4.3.2.1: Comparativa de las métricas derivadas de la matriz de confusión.

En la Tabla 4.3.2.2 se resaltan los aspectos más destacables:

Métrica	Modelo Base	Modelo Neuro 2	Modelo Neuro 1	Análisis
Especificidad	47.23%	68.96%	76.80%	Mayor capacidad para identificar correctamente la clase 1 (menos falsos positivos).
Reducción de Falsos Positivos (FP)	1542	907	678	Reducción significativa en errores de clasificación para la clase 0.
Verdaderos Negativos (VN)	1380	2015	2244	Mejora progresiva en la detección correcta de casos de clase 1.
Predicción de Clase 1	Más débil	Más equilibrada	Más precisa	El rendimiento sobre la clase menos representada mejora significativamente.

Tabla 4.3.2.2: Métricas resaltantes derivadas de la matriz de confusión.

De esta manera, aunque los tres modelos fueron evaluados con el mismo conjunto de datos, cada modelo presenta fortalezas distintas que pueden ser útiles según el contexto de aplicación, siendo así los siguientes análisis:

El modelo base, basado únicamente en el Modelo de Interferencia original (sin variables fisiológicas), logra identificar correctamente la mayoría de los intentos del Letterwheel con respuesta correcta (6032 casos). Sin embargo, también comete muchos errores clasificando como correctos a tales intentos que en realidad fueron incorrectos

(1542 errores de este tipo). Por tanto, este modelo es útil si el objetivo es no dejar por fuera la correctitud aunque se corra el riesgo de etiquetar intentos correctos cuando no es el caso.

El Modelo Neuro 1, que considera únicamente la variable de atención medida mediante EEG, es mucho más estricto. Comete menos errores al predecir como correctos intentos que en realidad fueron incorrectos (solo 678 casos mal clasificados), pero también falla en reconocer correctamente 2103 intentos que sí fueron correctos, clasificándolos erróneamente como incorrectos. Esto lo hace ideal en escenarios donde lo más importante es tener alta precisión al identificar intentos correctos, incluso si se sacrifica la capacidad de reconocerlos todos. Su alta especificidad (76.80 %) lo convierte en un modelo muy conservador, útil cuando se busca minimizar los falsos positivos.

Por otro lado, el Modelo Neuro 2, que combina los niveles de atención y estrés como variables predictoras, ofrece un balance más equilibrado. Predice correctamente 5041 intentos correctos, una cifra mayor que la del modelo neuro 1 (solo atención), aunque menor que la del modelo base. Además, reduce notablemente los errores al clasificar incorrectamente intentos que en realidad fueron fallidos (solo 907 falsos positivos, frente a 1542 del modelo base). Aunque no es tan conservador como el modelo neuro 1 ni tan inclinado a etiquetar como correctos como el modelo base, representa una opción más balanceada para contextos donde se busca tanto identificar los intentos correctos como evitar errores en las predicciones.

En resumen, la comparación de estos tres modelos permite observar cómo el uso de variables neurofisiológicas puede afinar las predicciones del rendimiento cognitivo. Según las prioridades del estudio (mayor cobertura, mayor precisión o equilibrio), uno u otro modelo puede ser más conveniente.

4.3.3. Gráfico de Barras Verticales: Exactitud por Sujeto

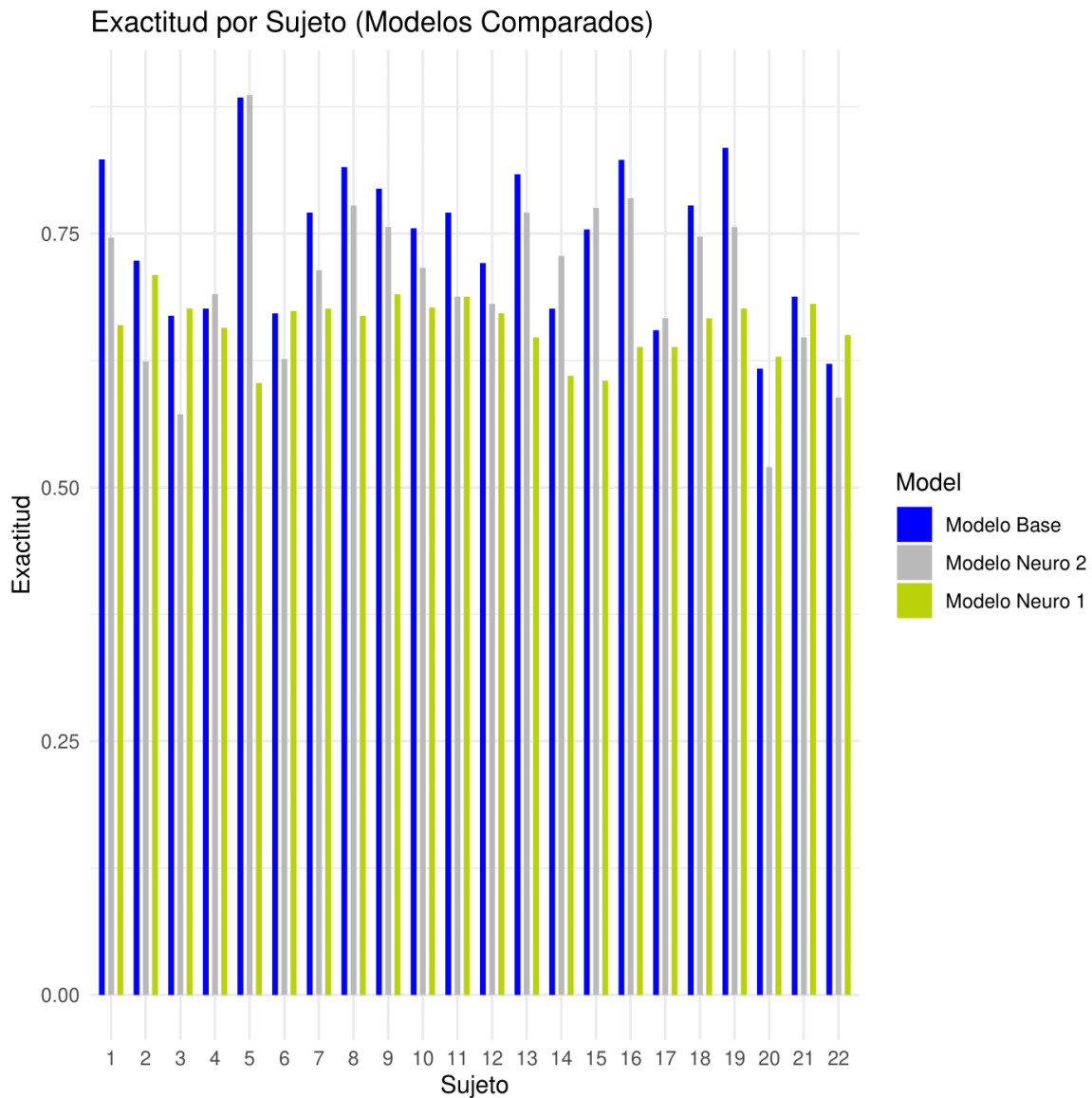


Figura 4.3.3: Gráfica de Barras Verticales: Exactitud por Sujeto.

Al observar la Fig. 4.3.3, exactitud por sujeto entre los tres modelos evaluados, se muestra que el modelo base alcanza consistentemente niveles más altos de exactitud en la mayoría de los participantes, por ejemplo, en el Sujeto 1, el modelo base obtiene una

exactitud cercana al 0.85, mientras que el modelo neuro 2 y el modelo neuro 1 se ubican alrededor de 0.74 y 0.68 respectivamente.

El modelo neuro 2, que incorpora información de los niveles de atención y estrés, presenta un desempeño intermedio, aunque no supera al modelo base en la mayoría de los casos, sí mejora con respecto al modelo que utiliza únicamente atención (modelo neuro 1). Un ejemplo de esto se observa en el Sujeto 13, donde el modelo neuro 2 logra una precisión de aproximadamente 0.80, por encima del modelo neuro 1 (alrededor de 0.70) y no muy lejos del modelo base (alrededor de 0.85).

Por otro lado, el modelo neuro 1, es el que muestra los valores más bajos de exactitud de forma sistemática, en sujetos como el Sujeto 5, su desempeño apenas alcanza los 0.68, mientras que el modelo base y el modelo neuro 2 se elevan por encima de 0.90. Este patrón sugiere que utilizar solo la variable de atención como predictor no es suficiente si se desea mayor exactitud.

Finalmente, en conjunto, estos hallazgos sugieren que, si bien las señales neurofisiológicas como la atención y el estrés aportan información valiosa, su impacto predictivo depende en gran medida de cómo se integran al modelo y del contexto cognitivo en el que se aplican.

Capítulo V - Conclusiones y Recomendaciones

El presente capítulo expone las conclusiones generales del estudio, las recomendaciones derivadas de los hallazgos obtenidos y posibles líneas de trabajos futuros. Se busca reflexionar sobre el alcance real del modelo propuesto, su aplicación en contextos educativos y experimentales, y los aportes que genera tanto a nivel metodológico como desde una perspectiva interdisciplinaria. Este apartado final resume los resultados obtenidos y su relación con los objetivos planteados al inicio de la investigación.

5.1. Conclusiones

La presente investigación logró diseñar, implementar y validar un modelo computacional capaz de predecir el rendimiento en tareas de memoria de trabajo en función de los niveles de atención y estrés mental, registrados a través de un dispositivo EEG de bajo costo (NeuroSky MindWave). Para ello, se integraron datos neurofisiológicos con parámetros conductuales durante la ejecución de la tarea Letterwheel, adaptada a condiciones experimentales controladas. La estructura del dataset, el proceso de normalización y el uso de métricas precisas permitieron desarrollar modelos de predicción que reflejan el vínculo entre estados cognitivos internos y el desempeño observado.

Entre los hallazgos más relevantes, se encontró que la atención mostró una mayor variabilidad entre los participantes y una mayor capacidad discriminativa en los modelos, mientras que el estrés presentó valores más homogéneos, lo que podría limitar su capacidad explicativa si no se incorporan transformaciones adicionales. El análisis exploratorio de datos (EDA) permitió visualizar estas tendencias y tomar decisiones informadas sobre la estructura final del dataset.

El modelo que incorporó las variables neurofisiológicas (Modelo Neuro 2: atención y estrés) alcanzó una precisión comparable al modelo base, pero mostró un mejor equilibrio entre sensibilidad y especificidad, lo cual es crucial para evitar errores sistemáticos en la predicción. Esto demuestra que el uso de señales EEG, incluso provenientes de dispositivos accesibles, puede mejorar la capacidad predictiva de modelos cognitivos computacionales.

Dada la evidencia observada durante el análisis exploratorio particularmente la distribución asimétrica de la atención normalizada y su mayor dispersión entre individuos, se consideró relevante construir un tercer modelo adicional (Modelo Neuro 1), basado únicamente en la variable de atención. Este modelo fue incorporado con el objetivo de comparar directamente el efecto aislado de esta variable sobre la predicción del rendimiento. Aunque su exactitud global fue ligeramente menor, destacó por su alta especificidad, es decir, por su capacidad para reducir los falsos positivos. Esto lo convierte en una alternativa interesante en contextos donde es más importante evitar sobreestimaciones del rendimiento que maximizar la sensibilidad.

La comparación entre los tres modelos permite identificar distintos enfoques según el objetivo del sistema predictivo: mientras que el Modelo Base prioriza una alta sensibilidad (identificando a la mayoría de los casos con buen rendimiento, aunque con mayor riesgo de errores), el Modelo Neuro 1 adopta una postura más conservadora, siendo preciso pero selectivo. El Modelo Neuro 2 se posiciona como una opción intermedia y balanceada, lo cual lo hace versátil para aplicaciones generales. En conjunto, este análisis aporta una comprensión más rica y matizada sobre el rol de los estados atencionales y emocionales en el desempeño cognitivo.

Asimismo, se respetaron los principios éticos fundamentales mediante la obtención de consentimiento informado, el anonimato de los datos y el diseño metodológico que evitó sesgos inducidos por expectativas. En conjunto, el estudio aportó una herramienta innovadora para la comprensión del rendimiento cognitivo en

estudiantes universitarios, alineando la teoría de la memoria de trabajo con técnicas contemporáneas de adquisición y modelado de datos.

Finalmente, esta investigación constituye un aporte interdisciplinario donde la Ingeniería de Sistemas, la psicología cognitiva y la ciencia de datos convergen para ofrecer un modelo predictivo. Los resultados invitan a repensar cómo evaluamos y adaptamos los procesos educativos según el estado mental de los estudiantes, proponiendo una educación más personalizada, más humana, y más basada en datos reales que emergen desde la cognición.

5.2. Recomendaciones

- **Ampliar y diversificar la muestra estudiantil**

Se recomienda replicar el estudio con una muestra más amplia y diversa que incluya estudiantes de distintas carreras, niveles académicos (pregrado y posgrado), e incluso de diferentes universidades o contextos culturales. Esta expansión permitiría evaluar la generalización del modelo más allá del grupo inicial, fortaleciendo su validez externa. Además, al incorporar perfiles cognitivos y emocionales más variados, sería posible observar cómo se comporta el modelo ante diferencias individuales significativas, enriqueciendo la comprensión del vínculo entre atención, estrés y rendimiento. Esto haría que los resultados sean más representativos y aplicables en escenarios reales de mayor diversidad educativa.

- **Optimizar el preprocesamiento de señales fisiológicas**

Aunque se aplicaron transformaciones básicas a los datos de atención y estrés, futuras investigaciones podrían explorar técnicas avanzadas de filtrado, reducción de ruido y extracción de características que capten mejor las variaciones temporales y no

lineales del EEG, lo cual podría mejorar significativamente el poder predictivo del modelo.

- **Profundizar en el análisis multivariado de las variables subjetivas**

Las encuestas aplicadas antes y después de la tarea experimental constituyen una fuente rica de datos subjetivos relacionados con la percepción interna del participante (nivel de fatiga, motivación, comprensión de la tarea, entre otros). Se recomienda aprovechar este recurso mediante análisis estadísticos multivariados más avanzados. Métodos como el análisis factorial exploratorio (AFE) [70] podrían identificar patrones latentes entre las respuestas, mientras que técnicas como el análisis de varianza multivariado (MANOVA) [71], permitirían evaluar el impacto conjunto de estas variables sobre los resultados de rendimiento y fisiológicos. Integrar estos análisis enriquecería la interpretación de los datos y podría revelar interacciones complejas entre factores subjetivos y objetivos que afectan la memoria de trabajo.

5.3. Trabajos Futuros

Dado el alcance y los hallazgos del presente estudio, se abren múltiples líneas de investigación prometedoras.

En primer lugar, se sugiere una profundización en el análisis de variables fisiológicas mediante la inclusión de nuevas métricas extraídas de las señales EEG crudas, como la conectividad funcional, coherencia o la descomposición en bandas de frecuencia mediante técnicas como Fast Fourier Transform (FFT) o análisis de wavelets [72]. También, en entornos educativos, las emociones son variables útiles [73] que pueden ser extraída de señales EEG. Esto permitiría obtener una representación más compleja del estado neurofisiológico del participante, superando las limitaciones de las métricas derivadas básicas como los índices de atención y estrés.

Segundo, se propone la expansión de la muestra hacia una población más diversa, tanto en términos de áreas académicas como de contextos socioculturales, lo cual permitiría evaluar la generalización del modelo propuesto y refinar su capacidad predictiva a diferentes perfiles estudiantiles. Adicionalmente, es propicio evaluar la utilidad del modelo en poblaciones con discapacidad cognitiva, considerando que algunos estudios muestran su variabilidad en el rendimiento sobre tareas de memoria [74].

Desde el punto de vista metodológico, sería conveniente aplicar técnicas estadísticas multivariadas avanzadas, como MANOVA o análisis factorial exploratorio, para evaluar interacciones entre variables latentes no observadas directamente, lo cual podría enriquecer la comprensión de los factores que modulan el rendimiento en tareas de memoria de trabajo.

En términos de modelado, por un lado, se plantea como línea futura el entrenamiento de modelos computacionales más complejos (como redes neuronales recurrentes o modelos basados en atención tipo Transformer), que puedan captar patrones no lineales entre la actividad cerebral y el desempeño cognitivo. Por otro lado, se pudieran agregar las variables (atención y estrés) a otros modelos de predicción, por ejemplo, un modelo basado en aprendizaje por refuerzo para predecir el rendimiento estudiantil [75].

Finalmente, se visualiza como una extensión natural del presente trabajo la integración del sistema en entornos educativos inteligentes, capaces de adaptar dinámicamente los contenidos o la dificultad de las tareas según los estados mentales del estudiante en tiempo real, acercándose así al ideal de una educación verdaderamente personalizada y neuro adaptativa.

Referencias Bibliográficas

[1] M. Silva-Barragán y C. Ramos-Galarza, “Modelos de Organización Cerebral: un recorrido neuropsicológico.”, *Rev. Ecuat. Neurol.*, vol. 29, n.º 3, pp. 74–83, 2020. <https://doi.org/10.46997/revecuatneurol29300074>

[2] T. Malim, “Introduction to Cognitive Processes”, en *Cognitive Processes*. London: Macmillan Educ. UK, 1994, pp. 1–10. https://doi.org/10.1007/978-1-349-13133-4_1

[3] M. K. Johnson y L. Hasher, “Human Learning and Memory”, *Annu. Rev. Psychol.*, vol. 38, n.º 1, pp. 631–668, enero de 1987. <https://doi.org/10.1146/annurev.ps.38.020187.003215>

[4] A. D. Baddeley, “The episodic buffer”, en *Exploring Working Memory*. Abingdon, Oxon ; New York, NY : Routledge, 2017. | Series: World library of psychologists: Routledge, 2017, pp. 297–311. <https://doi.org/10.4324/9781315111261-24>

[5] M. W. Eysenck, “Cognitive development”, en *Simply Psychology*, 5ª ed. London: Routledge, 2022, pp. 143–159. <https://doi.org/10.4324/9781003100492-10>

[6] E. Tulving, “Précis of Elements of episodic memory”, *Behav. Brain Sci.*, vol. 7, n.º 2, pp. 223–238, junio de 1984. <https://doi.org/10.1017/s0140525x0004440x>

[7] M. W. Eysenck (ed.), “Working Memory Model,” *Simply Psychology*, actualización del 3 de junio 2024. [En línea]. Disponible en: <https://www.simplypsychology.org/working-memory.html> [Último acceso: 14 junio 2025].

[8] F. Monteiro et al., “Optimizing working memory assessment: development of shortened versions of complex spans, updating, and binding tasks”, *Psycholog. Res.*, vol.

89, n.º 2, marzo de 2025. Accedido el 14 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.1007/s00426-025-02083-7>

[9] N. Parra-Bolaños, “Atención y Memoria en estudiantes con bajo rendimiento académico. Un estudio exploratorio”, *ReiDoCrea*, 2017. Accedido el 14 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.30827/digibug.45029>

[10] C. Ramos-Galarza, P. Acosta-Rodas, M. Bolaños-Pasquel y N. Lepe-Martínez, “The role of executive functions in academic performance and behaviour of university students”, *J. Appl. Res. Higher Educ.*, vol. 12, n.º 3, pp. 444–455, octubre de 2019. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.1108/jarhe-10-2018-0221>

[11] C. Sandi, “Stress and cognition”, *Wiley Interdisciplinary Rev.: Cogn. Sci.*, vol. 4, n.º 3, pp. 245–261, enero de 2013. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.1002/wcs.1222>

[12] L. Li, M. Sun, M. Qi, Y. Li y D. Li, “Neural correlates of emotional working memory predict depression and anxiety”, *Frontiers Neurosci.*, vol. 19, mayo de 2025. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.3389/fnins.2025.1574901>

[13] J. Shmueli, “To Explain or to Predict?,” *Stat. Sci.*, vol. 25, no. 3, pp. 289–310, 2010.

[14] T. Hastie, R. Tibshirani y J. Friedman, “Unsupervised Learning”, en *The Elements of Statistical Learning*. New York, NY: Springer N. Y., 2008, pp. 1–101. https://doi.org/10.1007/b94608_14

[15] P. M. C. Harrison, R. Bianco, M. Chait y M. T. Pearce, “Correction: PPM-Decay: A computational model of auditory prediction with memory decay”, *PLOS Comput. Biol.*, vol. 17, n.º 5, mayo de 2021, art. n.º e1008995. <https://doi.org/10.1371/journal.pcbi.1008995>

[16] L. M. Reder, A. Nhouyvanisvong, C. D. Schunn, M. S. Ayers, P. Angstadt y K. Hiraki, “A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm.”, *J. Exp. Psychol.: Learn., Memory, Cognition*, vol. 26, n.º 2, pp. 294–320, 2000. <https://doi.org/10.1037/0278-7393.26.2.294>

[17] K. Oberauer y H.-Y. Lin, “An interference model of visual working memory.”, *Psycholog. Rev.*, vol. 124, n.º 1, pp. 21–59, 2017. <https://doi.org/10.1037/rev0000044>

[18] Chang-Hwan. Im, Ed., *Computational EEG Analysis*. Singapore: Springer Singap., 2018. <https://doi.org/10.1007/978-981-13-0908-3>

[19] D. Ni, S. Wang y G. Liu, “The EEG-Based Attention Analysis in Multimedia m-Learning”, *Comput. Math. Methods Medicine*, vol. 2020, pp. 1–10, junio de 2020. <https://doi.org/10.1155/2020/4837291>

[20] S. E. Gathercole, S. J. Pickering, C. Knight y Z. Stegmann, “Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age”, *Appl. Cogn. Psychol.*, vol. 18, n.º 1, pp. 1–16, diciembre de 2003. <https://doi.org/10.1002/acp.934>

[21] R. W. Engle, S. W. Tuholski, J. E. Laughlin y A. R. A. Conway, “Working memory, short-term memory, and general fluid intelligence: A latent-variable approach.”, *J. Exp. Psychol.: General*, vol. 128, n.º 3, pp. 309–331, septiembre de 1999. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.1037/0096-3445.128.3.309>

[22] V. R. LeBlanc, “The Effects of Acute Stress on Performance: Implications for Health Professions Education”, *Acad. Med.*, vol. 84, Supplement, pp. S25—S33, octubre de 2009. <https://doi.org/10.1097/acm.0b013e3181b37b8f>

[23] G. Vos, M. Ebrahimpour, L. van Eijk, Z. Sarnyai y M. R. Azghadi, “Stress monitoring using low-cost electroencephalogram devices: A systematic literature review”, *Int. J. Med. Inform.*, p. 105859, marzo de 2025. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.1016/j.ijmedinf.2025.105859>

[24] H.-M. Süß, K. Oberauer, W. W. Wittmann, O. Wilhelm y R. Schulze, “Working-memory capacity explains reasoning ability—and a little bit more”, *Intelligence*, vol. 30, n.º 3, pp. 261–288, mayo de 2002. [https://doi.org/10.1016/s0160-2896\(01\)00100-3](https://doi.org/10.1016/s0160-2896(01)00100-3)

[25] B. Cárcamo, “Modelos de la memoria de trabajo de Baddeley y Cowan: una revisión bibliográfica comparativa,” *Revista Chilena de Neuro-Psiquiatría*, vol. 13, 2018.

[26] A. Miyake. (Editor) y P. Shah. (Editor), Eds., *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge Univ. Press, 2001.

[27] A. Baddeley y G. J. Hitch, “Working memory,” en *Psychology of Learning and Motivation*, vol. 8, Academic Press, 1974, pp. 47–89.

[28] N. Cowan, “The magical number 4 in short-term memory,” *Behav. Brain Sci.*, vol. 24, no. 1, pp. 87–185, 2001.

[29] K. Oberauer, S. Lewandowsky, S. Farrell, C. Jarrold y M. Greaves, “Modeling working memory: An interference model of complex span”, *Psychonomic Bull. & Rev.*, vol. 19, n.º 5, pp. 779–819, junio de 2012. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.3758/s13423-012-0272-4>

[30] K. Oberauer y R. Kliegl, “Beyond resources: Formal models of complexity effects and age differences in working memory”, *Eur. J. Cogn. Psychol.*, vol. 13, n.º 1-2, pp. 187–215, marzo de 2001. <https://doi.org/10.1080/09541440042000278>

[31] K. Oberauer and H.-Y. Lin, “An interference model of visual working memory,” *Psychol. Rev.*, vol. 124, no. 1, pp. 21–59, 2016, doi: 10.1037/rev0000044.

[32] O. Wilhelm, A. Hildebrandt y K. Oberauer, “What is working memory capacity, and how can we measure it?”, *Frontiers Psychol.*, vol. 4, 2013. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.3389/fpsyg.2013.00433>

[33] C. Artuso y P. Palladino, “Binding and content updating in working memory tasks”, *Brit. J. Psychol.*, vol. 105, n.º 2, pp. 226–242, febrero de 2013. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.1111/bjop.12024>

[34] R. D. Astuti, B. Suhardi, P. W. Laksono y N. Susanto, “Investigating the Relationship between Noise Exposure and Human Cognitive Performance: Attention, Stress, and Mental Workload Based on EEG Signals Using Power Spectrum Density”, *Appl. Sci.*, vol. 14, n.º 7, p. 2699, marzo de 2024. <https://doi.org/10.3390/app14072699>

[35] J.-R. Wang y S. Hsieh, “Neurofeedback training improves attention and working memory performance”, *Clin. Neurophysiol.*, vol. 124, n.º 12, pp. 2406–2420, diciembre de 2013. <https://doi.org/10.1016/j.clinph.2013.05.020>

[36] M. I. Posner y S. E. Petersen, “The attention system of the human brain,” *Annu. Rev. Neurosci.*, vol. 13, pp. 25–42, mar. 1990, doi: 10.1146/[annurev.ne.13.030190.000325](https://doi.org/10.1146/annurev.ne.13.030190.000325).

[37] M. M. Sohlberg y C. A. Mateer, “Effectiveness of an attention training program,” *J. Clin. Exp. Neuropsychol.*, vol. 9, no. 2, pp. 117–130, abr. 1987, doi: 10.1080/01688638708405352.

[38] K. M. Lukasik, O. Waris, A. Soveri, M. Lehtonen y M. Laine, “The Relationship of Anxiety and Stress With Working Memory Performance in a Large Non-depressed Sample”, *Frontiers Psychol.*, vol. 10, enero de 2019. Accedido el 15 de junio de 2025. [En línea]. Disponible: <https://doi.org/10.3389/fpsyg.2019.00004>

[39] R. M. Yerkes y J. D. Dodson, “The relation of strength of stimulus to rapidity of habit-formation,” *J. Comp. Neurol. Psychol.*, vol. 18, no. 5, pp. 459–482, 1908, doi: 10.1002/cne.920180503.

[40] A. Gevins, M. E. Smith, L. K. McEvoy, H. Leong y J. Le, “Electroencephalographic imaging of higher brain function”, *Philos. Trans. Roy. Soc. London. Ser. B: Biol. Sci.*, vol. 354, n.º 1387, pp. 1125–1134, julio de 1999. <https://doi.org/10.1098/rstb.1999.0468>

[41] W. Ray y H. Cole, “EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes”, *Science*, vol. 228, n.º 4700, pp. 750–752, mayo de 1985. <https://doi.org/10.1126/science.3992243>

[42] B. Pitchford y K. M. Arnell, “Resting EEG in alpha and beta bands predicts individual differences in attentional breadth”, *Conscious. Cogn.*, vol. 75, p. 102803, octubre de 2019. <https://doi.org/10.1016/j.concog.2019.102803>

[43] P. Putman, B. Verkuil, E. Arias-Garcia, I. Pantazi y C. van Schie, “EEG theta/beta ratio as a potential biomarker for attentional control and resilience against deleterious effects of stress on attention”, *Cogn., Affect., & Behav. Neurosci.*, vol. 14, n.º 2, pp. 782–791, diciembre de 2013. <https://doi.org/10.3758/s13415-013-0238-7>

[44] M. Arns, C. K. Conners y H. C. Kraemer, “A decade of EEG Theta/Beta Ratio Research in ADHD: a meta-analysis,” *J. Atten. Disord.*, vol. 17, no. 5, pp. 374–383, jul. 2013, doi: 10.1177/1087054712460087.

[45] A. Bluschke et al., “The neuronal mechanisms underlying improvement of impulsivity in ADHD by theta/beta neurofeedback,” *Sci. Rep.*, vol. 6, art. 31178, 2016, doi: 10.1038/srep31178.

[46] J. A. Putman, J. P. van Peer, M. J. Maimari y S. van der Werff, “The theta/beta ratio as a biomarker for attentional control and resilience against the effects of

stress on attention,” *Psychophysiology*, vol. 51, no. 9, pp. e13, sep. 2014, doi:10.1111/psyp.12161.

[47] A. Zavala, A. D. Jang, E. Youghee y T. W. Cavanagh, “Cognitive control involves theta power within trials and beta power across trials from medial prefrontal cortex and subthalamic nucleus,” *J. Neurosci.*, vol. 36, no. 37, pp. 960–973, sep. 2016, doi:10.1523/JNEUROSCI.1222-16.2016.

[48] W. Klimesch, “Alpha-band oscillations, attention, and controlled access to stored information,” *Trends Cogn. Sci.*, vol. 16, no. 12, pp. 606–617, dic. 2012, doi: 10.1016/j.tics.2012.10.007.

[49] NeuroSky, “ThinkGear ASIC Module (TGAM) EEG sensor biosensor specifications,” NeuroSky Biosensors, 2025. [En línea]. Disponible: <https://neurosky.com/biosensors/eeg-sensor/biosensors/> [Último acceso: 14 jun. 2025]

[50] F. Ayubi, S. Maslovat y T. Jafarzadeh, “Low-Cost EEG Multi-Subject Recording Platform for Student Attention Assessment,” *Sensors*, vol. 23, no. 1, art. 560, ene. 2023, doi:10.3390/s23010560.

[51] SparkFun Electronics, “MindWave Mobile 2 Transition Doc,” SparkFun, 2018. [En línea]. Disponible: <https://cdn.sparkfun.com/assets/0/6/a/8/9/MWMobile2TransitionDoc.pdf> [Último acceso: 14 jun. 2025]

[52] R. Maskeliūnas, R. Damasevičius, I. Martišius y M. Vasiljevas, “Consumer-grade EEG devices: Are they usable for control tasks?,” *PeerJ*, vol. 4, e1746, 2016, doi:10.7717/peerj.1746.

[53] M. Mansoor, B. Naz, R. Jaffari y A. Ansari, “Brain controlled wheelchair with NeuroSky MindWave,” *Sindh Univ. Res. J. (Sci. Ser.)*, vol. 51, no. 4, pp. 709–712, 2019, doi:10.26692/surj/2019.12.112.

[54] A. Sezer, Y. İnel, A. Ç. Seçkin y U. Uluçınar, “An Investigation of University Students’ Attention Levels in Real Classroom Settings with NeuroSky’s MindWave Mobile EEG Device,” en Proc. IETC 2015, Estambul, Turquía, may. 2015.

[55] G. Rebolledo-Mendez et al., “Assessing NeuroSky’s usability to detect attention levels in an assessment exercise,” en User Model. User-Adapt. Interact, vol. 21, pp. 65–86, 2011, doi:10.1007/s11257-010-9083-6.

[56] M. A. Boudewyn, C. Luck y E. A. Molholm, “Managing EEG studies: how to prepare and what to do once data collection begins,” Psychophysiology, vol. 60, no. 11, e14481, nov. 2023, doi:10.1111/psyp.14481.

[57] M. W. Berry and M. Browne, Understanding Search Engines: Mathematical Modeling and Text Retrieval, SIAM, 1999.

[58] F. Peng, “Reproducible research in computational science: Principles and practices,” Frontiers in Neuroscience, vol. 13, p. 262, 2019, doi: 10.3389/fnins.2019.00262.

[59] J. Illes, S. M. Racine, K. Kirschen, W. W. Gabriel, L. A. Nyberg, and J. R. Andrews, “Ethical challenges of neuroimaging technologies: Controversies, questions, and consensus,” *Nature Reviews Neuroscience*, vol. 7, no. 1, pp. 61–67, Jan. 2006, doi: 10.1038/nrn1823.

[60] A. GitHub Contributors, “NeuroPy: A simple interface for Neurosky Mindwave EEG headset,” GitHub, 2016. [Online]. Disponible: <https://github.com/lihas/neuroPy>

[61] C. H. Zelada, Modelo Lineal Generalizado (GLM), RPubS, 2017. [Online]. Disponible: <https://rpubs.com/chzelada/275494>

[62] F. G. Arias, El proyecto de investigación: introducción a la metodología científica, 5.ª ed., Bogotá: Editorial XYZ, 2006, cap. 7, pp. 84–85.

[63] V. Buchholz, J. Spence y B. Roberts, “Age-related changes in ERP components during working memory tasks,” *Biol. Psychol.*, vol. 52, no. 3, pp. 207–232, 2000.

[64] E. Illes et al., “Ethical challenges in neuroimaging research,” *Nature Rev. Neurosci.*, vol. 7, no. 1, pp. 61–67, 2006, doi:10.1038/nrn1823.

[65] M. A. Siepman, W. Kirch, et al., “Effects of caffeine on topographic quantitative EEG,” *Neuropsychobiology*, vol. 45, no. 3, pp. 161–166, 2002, doi:10.1159/000054958.

[66] P. F. Schyns, J. T. Megraw y C. J. Wruck, “Effects of uncorrected vision on participation in visual cognitive tasks: implications for experimental design,” *Front. Hum. Neurosci.*, vol. 14, art. 882557, 2022, doi: 10.3389/fnhum.2022.882557.

[67] C. R. Hepner y N. Nozari, “Resource allocation in phonological working memory: Same or different principles from vision?,” *J. Mem. Lang.*, vol. 106, pp. 172–188, jun. 2019, doi:10.1016/j.jml.2019.03.003.

[68] OSF, “Repositorio de datos,” Open Science Framework [En línea]. Disponible en: <https://osf.io/znqmw/files/osfstorage>. [Accedido: 15-jun-2025].

[69] L. E. McMorris, “Effects of acute exercise on cognitive performance: A meta-analytic review,” *Brain Sciences*, vol. 7, no. 9, p. 110, Sep. 2017, doi: 10.3390/brainsci7090110.

[70] H. F. Kaiser, "The application of electronic computers to factor analysis," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 141–151, 1960. doi:10.1177/001316446002000116.

[71] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed., Wiley, 2003.

[72] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 4th ed., Upper Saddle River, NJ, USA: Prentice Hall, 2006, ch. 4.

[73] J. Pérez, E. Dapena and J. Aguilar. Emotions as implicit feedback for adapting difficulty in tutoring systems based on reinforcement learning. *Education and Information Technologies*, 2024. doi: 10.1007/s10639-024-12699-8.

[74] J. Pérez, M. Azuaje, C. León and O. Pedroza, "Effects of Social Robotics on Episodic Memory in Children With Intellectual Disabilities," in *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, vol. 16, no. 4, pp. 393-399, Nov. 2021, doi: 10.1109/RITA.2021.3125899.

[75] J. Pérez, E. Dapena, J. Aguilar and G. Carrillo, "Reinforcement Learning for Estimating Student Proficiency in Math Word Problems," 2022 XVII Latin American Conference on Learning Technologies (LACLO), Armenia, Colombia, 2022, pp. 01-06, doi: 10.1109/LACLO56648.2022.10013399.

Anexos

Anexo 1: Carta de Consentimiento Informado



Proyecto: Análisis de la Relación entre la Memoria de Trabajo, la atención y el estrés mental en estudiantes de la Escuela de Ingeniería de Sistemas de la Universidad de Los Andes

Investigadora responsable: Daniella Zambrano – Estudiante de Ingeniería de Sistemas

Estimado/a _____ C.I: _____

Le invito a participar en el presente estudio, cuyo objetivo es analizar cómo se relacionan la memoria de trabajo junto con los niveles de atención y niveles de estrés mental en estudiantes universitarios, en este caso, específicamente con estudiantes de la Escuela de Ingeniería de Sistemas, de la Universidad de Los Andes. Esto con el fin de contribuir a mi Proyecto Final de Grado. Los datos obtenidos serán utilizados únicamente con fines académicos y de investigación.

Procedimientos:

Durante el estudio, se le solicitará:

- Realizar una tarea de memoria donde deberá recordar y repetir palabras aleatorias sin un significado gramatical.
- Usar un dispositivo llamado **MindWave**, que mide la actividad cerebral para conocer su nivel de atención y estrés mental mientras realiza la tarea.
- Responder una breve encuesta tanto inicial como final, donde se le pedirán algunos datos como nivel de fatiga y estado emocional.

La actividad durará aproximadamente **40 a 50 minutos**, en un ambiente controlado.

Riesgos e inconvenientes:

Este estudio **no representa riesgos físicos o psicológicos directos**. Sin embargo, algunas tareas podrían ocasionar cansancio.

Derechos, responsabilidades y beneficios:

- Su participación es **completamente voluntaria**.
- Tiene **derecho a retirarse en cualquier momento**, sin dar explicaciones y sin que esto le genere ninguna consecuencia negativa.
- Puede **hacer preguntas** y recibir respuestas claras antes, durante y después del estudio.
- Todos los datos serán tratados de manera **confidencial y anónima**. Su nombre no aparecerá en ningún informe, publicación o resultado.
- La información recopilada **no será compartida con terceros** y se almacenará de forma segura.
- Aunque no recibirá compensación económica, su participación es **muy valiosa para el desarrollo académico de este proyecto** y puede beneficiar a futuras investigaciones.

Confidencialidad y manejo de la información:

Toda la información recopilada será protegida, su identidad no será divulgada y los datos serán usados únicamente con fines investigativos.

Declaración de consentimiento:

He leído y entendido la información proporcionada. Comprendo los objetivos del estudio, los procedimientos, los posibles riesgos y mis derechos como participante. Acepto participar de manera voluntaria y sé que puedo retirarme en cualquier momento.

Nombre completo: _____

Firma: _____

Fecha: ___ / ___ / _____

Datos de contacto:

Para cualquier pregunta o información adicional, puede comunicarse al correo:
daniellamzr@gmail.com

Gracias por considerar participar en este estudio.

Anexo 2: Encuesta Inicial y Final

Escala tipo Likert:

Para el nivel de Cansancio:

1. **Nada fatigado** – Me siento completamente descansado/a y alerta.
2. **Poco fatigado** – Siento una ligera fatiga, pero puedo concentrarme bien.
3. **Neutral** – Estoy algo cansado/a, pero aún así puedo realizar la tarea sin mucha dificultad.
4. **Fatigado** – Me cuesta concentrarme y siento la necesidad de descansar.
5. **Muy fatigado** – Estoy extremadamente cansado/a y me resulta difícil continuar.

Para el estado de ánimo:

1. **Feliz** – Me siento bien, relajado/a y con una actitud positiva hacia la actividad.
2. **Neutral** – No tengo una emoción predominante en este momento, me siento en equilibrio.
3. **Triste** – Me siento desmotivado/a o con baja energía emocional.
4. **Enojado** – Estoy frustrado/a ó molesto/a por algo, lo que podría afectar mi concentración.
5. **Asustado** – Siento ansiedad o nerviosismo respecto a la tarea o mi desempeño.