



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Habilidades científicas en Física Experimental:

Una metodología automatizada para el análisis de
informes de laboratorio.

Lorena Stephanie Pagella Martínez

Programa de Posgrado en Física de PEDECIBA
Facultad de Ciencias, Facultad de Ingeniería
Universidad de la República

Montevideo – Uruguay
Mayo de 2026



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Habilidades científicas en Física Experimental:

Una metodología automatizada para el análisis de
informes de laboratorio.

Lorena Stephanie Pagella Martínez

Tesis de Maestría presentada al Programa de
Posgrado en Física de PEDECIBA, Facultad de
Ciencias de la Universidad de la República, como
parte de los requisitos necesarios para la obtención
del título de Magíster en Física de PEDECIBA.

Director:

Dra. Cecilia Stari

Codirector:

Dra. Aiala Rosá

Montevideo – Uruguay

Mayo de 2026

Pagella Martínez, Lorena Stephanie

Habilidades científicas en Física Experimental: / Lorena Stephanie Pagella Martínez. - Montevideo: Universidad de la República, Facultad de Ciencias, Facultad de Ingeniería, 2026.

XVI, 235 p.: il.; 29, 7cm.

Director:

Cecilia Stari

Codirector:

Aiala Rosá

Tesis de Maestría – Universidad de la República, Programa en Física de PEDECIBA, 2026.

Referencias bibliográficas: p. 125 – 138.

1. Física Experimental, 2. Taxonomía de Bloom, 3. Procesamiento Lenguaje Natural, 4. Habilidades científicas, 5. Investigación en Enseñanza de la Física. I. Stari, Cecilia, *et al.* II. Universidad de la República, Programa de Posgrado en Física de PEDECIBA. III. Título.

INTEGRANTES DEL TRIBUNAL DE DEFENSA DE TESIS

Dra. Marcela Peláez

Dra. Lucía Amy

Dr. Caracé Gutiérrez

Dr. Guillermo Moncecchi

Dra. Ximena Otegui

Montevideo – Uruguay

Mayo de 2026

Agradecimientos

Quiero expresar mi agradecimiento a todas las personas e instituciones que hicieron posible la realización de esta tesis.

En primer lugar, a mis tutoras, Cecilia y Aiala, por su orientación, acompañamiento y lectura atenta durante las distintas etapas de este trabajo.

Al equipo docente de Física Experimental 1 de Facultad de Ingeniería y a los estudiantes, por compartir sus perspectivas y contribuir al desarrollo de esta investigación.

A *los anotadores*, Rodrigo y Martín, quienes dedicaron generosamente su tiempo a colaborar con este trabajo. Gracias por su apoyo y por el tiempo dedicado, su colaboración fue fundamental para llevar este trabajo adelante.

A la Agencia Nacional de Investigación e Innovación y al Consejo de Formación en Educación, por la beca otorgada, cuyo apoyo resultó fundamental para avanzar en el desarrollo de esta tesis.

A mis compañeras y compañeros con quienes transité los distintos cursos y compartí instancias de trabajo a lo largo de este proceso. En especial, a Patricia, por animarme a comenzar la Maestría.

Finalmente, a mi familia y a mis amigas, por el apoyo, la paciencia ante mis ausencias y el aliento durante este proceso.

RESUMEN

El laboratorio ocupa un lugar central en la enseñanza y el aprendizaje de la física. En este contexto, los informes de laboratorio constituyen una fuente relevante de información, ya que permiten analizar algunas de las habilidades científicas desarrolladas en los cursos y evidenciadas en las producciones escritas del estudiantado. Esta tesis analiza las habilidades científicas presentes en informes elaborados por estudiantes de Física Experimental 1 de la Facultad de Ingeniería de la Universidad de la República, mediante técnicas de Inteligencia Artificial, especialmente de Procesamiento del Lenguaje Natural (PLN). El enfoque teórico articula recomendaciones derivadas de la investigación en educación en física sobre los laboratorios universitarios con la taxonomía revisada de Bloom. La metodología incluyó la definición de clases para la clasificación de textos, la elaboración de una guía de anotación y la construcción de un corpus anotado. Luego se evaluaron distintos algoritmos de aprendizaje automático y se analizaron sus errores. También se examinaron los objetivos del curso, las guías de laboratorio y las percepciones de docentes y estudiantes. Finalmente, se analizaron las habilidades científicas identificadas mediante la clasificación automática de los informes de un semestre. Los resultados muestran que es posible desarrollar una metodología automatizada para analizar textos producidos en cursos de laboratorio y evaluar, desde una perspectiva global, algunas de las habilidades científicas evidenciadas en la escritura estudiantil. Además, permiten observar que el diseño de las guías de laboratorio incide en la escritura de los informes y en las habilidades identificables en ellos. En particular, se encontró que las consignas más abiertas no siempre aseguran la presencia de habilidades asociadas a procesos cognitivos de orden superior, y que existen diferencias entre docentes y estudiantes respecto de las habilidades consideradas prioritarias. En conjunto, la tesis contribuye a caracterizar las habilidades científicas en un curso introductorio de Física Experimental y aporta una metodología basada en PLN para el análisis sistemático de producciones escritas en la enseñanza de la física.

Palabras claves:

Física Experimental, Taxonomía de Bloom, Procesamiento Lenguaje Natural, Habilidades científicas, Investigación en Enseñanza de la Física.

ABSTRACT

Laboratory work is central to the teaching and learning of physics. In this context, laboratory reports constitute a relevant source of information, as they allow for the analysis of some of the scientific skills developed in courses and evidenced in students' written productions. This thesis analyzes the scientific skills present in reports written by students enrolled in Experimental Physics 1 at the School of Engineering of the Universidad de la República, using Artificial Intelligence techniques, especially Natural Language Processing (NLP). The theoretical framework combines recommendations from physics education research on university laboratories with Bloom's revised taxonomy. The methodology included the definition of classes for text classification, the development of an annotation guide, and the construction of an annotated corpus. Different machine learning methods were then evaluated, and their errors were analyzed. The course objectives, laboratory guides, and instructors' and students' perceptions were also examined. Finally, the scientific skills identified through the automatic classification of reports from one semester were analyzed. The results show that it is possible to develop an automated methodology for analyzing texts produced in laboratory courses and for assessing, from a global perspective, some of the scientific skills evidenced in students' writing. They also indicate that the design of laboratory guides influences the writing of reports and the skills that can be identified in them. In particular, it was observed that more open-ended prompts do not always ensure the presence of skills associated with higher-order cognitive processes, and that there are differences between instructors and students regarding the skills considered to be priorities. Overall, this thesis contributes to characterizing the scientific skills developed in an introductory Experimental Physics course and offers an NLP-based methodology for the systematic analysis of written productions in physics education.

Keywords:

Experimental Physics, Bloom's Taxonomy, Natural Language Processing, Scientific Skills, Physics Education Research.

Lista de figuras

2.1	Jerarquía de la categorías en la Taxonomía de Bloom.	15
2.2	Procesos cognitivos de acuerdo a la Taxonomía de Bloom revisada por Anderson y Krathwohl (2001)	17
2.3	Tabla de las dimensiones dela TBA.	20
3.1	Enfoques para la definición de la inteligencia artificial	23
3.2	Ciclo MATTER adaptado de Pustejovsky y Stubbs (2014) . . .	27
3.3	Matriz de confusión con etiquetas reales y predichas.	30
3.4	Red neuronal de dos capas adaptado de Jurafsky y Martin (2025c)	33
4.1	Representación gráfica de la metodología de investigación utilizada.	41
4.2	Diagrama con la agrupación de las habilidades para los cursos de laboratorio de acuerdo a la Kozminski et al. (2014)	47
5.1	Ejemplo de un enunciado y su versión lematizada.	59
5.2	Representación en dos dimensiones de los <i>sentence embeddings</i> para las clases APL y EDE sobre el corpus de entrenamiento. .	60
5.3	Ejemplo con los 3 vecinos más cercanos a un enunciado de la clase EDE.	62
5.4	Comparación del <i>Macro F1-score</i> para la clase APL sobre el corpus de desarrollo.	63
5.5	Estructura general del <i>prompt</i> utilizado para la tarea de clasificación.	67
5.6	Estructura general del <i>prompt</i> utilizado para la clasificación binaria.	68
5.7	<i>Macro F1-score</i> para cada una de las clases y modelo en el corpus de evaluación.	69

5.8	<i>Macro F1-score</i> para cada clase luego de la votación por mayoría simple.	72
6.1	Clasificación de los objetivos del curso Física Experimental 1 en la tabla de la Taxonomía de Bloom revisada por Anderson y Krathwohl (2001)	82
6.2	Frecuencia de los objetivos más importantes para el curso según los docentes.	84
6.3	Comparación entre docentes y estudiantes acerca de la prioridad en las habilidades a desarrollar en el curso Física Experimental 1.	87
6.4	Distribución de respuestas de docentes a la pregunta sobre la incorporación de herramientas adicionales de evaluación en el curso de Física Experimental 1.	88
6.5	Distribución de respuestas sobre en qué medida se consideran distintas habilidades en la evaluación del curso. Pregunta de selección múltiple.	89
7.1	Comparación de la distribución porcentual de enunciados por clase y por práctica (indicadas con P1, P2, P3 y P4).	93
7.2	Nubes de palabras de las cuatro prácticas analizadas, con las 40 palabras más frecuentes según su lema.	95
7.3	Esquema de verificación cruzada sobre habilidades vinculadas a la experimentación y medición en laboratorio.	108
7.4	Distribución de enunciados pertenecientes o no a la clase EDE por clasificador en la Práctica N°1.	110
7.5	Distribución de la cantidad de enunciados de la clase EDE por informe en la primera y última práctica del semestre. Se asignó valor cero a los informes en los que no se identificaron enunciados.	111
7.6	Esquema de verificación cruzada sobre habilidades vinculadas al diseño experimental.	116
D.1	Desglose de objetivos dentro de cada una de las áreas clave en los cursos de laboratorio según las recomendaciones de la AAPT y códigos asociados.	153

Lista de tablas

2.1	Áreas de enfque en los cursos de laboratorio según la Kozminski et al., 2014	14
2.2	Verbos relacionados a cada categoría de la TB. Tabla adaptada de Ferraz y Belhot (2010 , p.426)	16
3.1	Interpretación de los valores de la Kappa de Cohen de acuerdo con Landis y Koch (1977)	29
4.1	Descripción de las dimensiones del proceso cognitivo de la TBA.	46
4.2	Guía de anotación con las habilidades científicas según los procesos cognitivos de la TBA.	48
4.3	Clases, códigos y criterios de inclusión definidos en la guía de anotación.	53
4.4	Nivel de acuerdo entre evaluadores humanos según el estadístico κ de Cohen para cada clase interpretado de acuerdo a Landis y Koch (1977)	55
4.5	Distribución de enunciados anotados por clase antes y después de ampliar el corpus.	55
4.6	Distribución de la cantidad de enunciados pertenecientes a cada clase y por corpus.	56
5.1	Pruebas con distintas técnicas de clasificación realizadas sobre el corpus de desarrollo.	64
5.2	Comparativa de resultados de la métrica <i>Macro F1-score</i> entre MLP y LR sobre el corpus de desarrollo para las diferentes clases.	66
5.3	Enunciados FP y FN en la clasificación del corpus de evaluación, por modelo, para la clase APL.	74

5.4	Total de enunciados clasificados incorrectamente sobre el corpus de evaluación, por modelo, para la clase EDE.	77
6.1	Objetivos del curso Física Experimental 1 extraídos del Programa del Curso (véase Anexo 4).	81
6.2	Objetivos del curso agrupados por proceso cognitivo (TBA) y clase.	83
7.1	Cantidad de enunciados extraídos para el análisis de las secciones <i>Metodología</i> , <i>Análisis de datos</i> y <i>Conclusiones</i> correspondientes a informes del primer semestre de 2024.	92
7.2	Cantidad de enunciados clasificados con AA por práctica y por clase.	93
7.3	Justificaciones identificadas en los enunciados sobre la necesidad de realizar múltiples mediciones.	99
A.1	Palabras asociadas a las categorías de la TBA utilizadas en los informes de laboratorio. Fuente: Elaboración propia.	144
G.1	Lista de <i>stopwords</i> en español	166
1.1	Descripción de las categorías de la TB. Elaboración propia. Fuente: Bloom (1990).	184

Lista de siglas

Lista de siglas

AA Aprendizaje Automático

AAPT American Association of Physics Teachers

AIC Acuerdo Interanotador o Intercodificador

BoW Bag of Words

FCI Force Concept Inventory

FN Falso Negativo

FP Falso Positivo

FS Few-Shot

IA Inteligencia Artificial

ISLE Interactive Science Learning Environment

LLM Large Language Model

LR Logistic Regression

MLP Multilayer Perceptron

MNB Multinomial Naive Bayes

PCA Análisis de componentes principales

PER Physics Education Research

PLN Procesamiento del Lenguaje Natural

PSSC Physical Science Study Committee

RTP Real Time Physics

SDLs Socratic Dialogue Labs

SVM Support Vector Machine

TB Taxonomía de Bloom

TBA Taxonomía de Bloom revisada por Anderson y Krathwohl

ZS Zero-Shot

k-NN k-Nearest Neighbors

Tabla de contenidos

Lista de figuras	VIII
Lista de tablas	X
Lista de siglas	XIII
1 Introducción.	1
2 Enseñanza de la Física Experimental: entre objetivos y habilidades.	5
2.1 Investigaciones en Educación en Física Experimental	6
2.1.1 Objetivos de los cursos de laboratorio y las habilidades científicas	11
2.1.2 Objetivos y la Taxonomía de Bloom	15
3 La Inteligencia Artificial y la investigación en Educación en Física.	22
3.1 El Procesamiento del Lenguaje Natural.	25
3.2 Métodos de clasificación basados en aprendizaje automático supervisado y su evaluación.	26
3.3 Modelos de Clasificación.	31
3.3.1 Representación de palabras y texto.	34
3.3.2 Modelos de lenguaje neuronales y clasificación de textos.	36
3.4 Antecedentes sobre el uso de IA en la enseñanza de la física experimental.	38
3.5 Síntesis del capítulo.	39
4 Diseño de la investigación y construcción del corpus.	41
4.1 Modelado y construcción de la guía de anotación.	43

4.2	Construcción del corpus anotado.	54
5	Clasificación de enunciados: experimentos y resultados.	57
5.1	Clasificación con técnicas de PLN.	57
5.2	Configuración y evaluación de los algoritmos de AA utilizados. .	65
5.3	Análisis de errores en la clasificación automática.	73
5.3.1	Análisis de errores sobre el corpus de evaluación para la clase APL.	73
5.3.2	Análisis de errores sobre el corpus de evaluación para la clase EDE.	76
5.4	Síntesis del capítulo.	79
6	Objetivos y habilidades: entre el diseño y la percepción.	80
7	Habilidades identificadas en el curso de física experimental.	91
7.1	Los informes del curso: descripción y clasificación automática de enunciados.	91
7.2	Habilidades prácticas propias de la experimentación y medición en laboratorio (clase APL).	94
7.3	Habilidad para evaluar el diseño y procedimiento experimental (clase EDE).	109
8	Conclusiones y consideraciones finales.	118
8.1	Las habilidades científicas en física experimental.	119
8.2	El desarrollo de una metodología automatizada.	121
8.3	Consideraciones finales.	123
	Referencias	125
	Publicaciones y presentaciones en congresos	139
	Glosario	141
	Apéndices	142
	Apéndice A Lisado de palabras pertenecientes a la categorías de la TBA en Física.	143
	Apéndice B Palabras y secuencias de palabras asociadas con las diferentes categorías de la TBA.	145

Apéndice C Descripción de las categorías y procesos cognitivos de la TBA aplicado a la Física Experimental.	147
Apéndice D Desglose de las áreas clave en los cursos de laboratorio y códigos asociados.. . . .	152
Apéndice E Guía de anotación (segunda versión).	154
Apéndice F Guía de anotación, versión final.	161
Apéndice G Lista de <i>Stop Words</i> en español.	165
Apéndice H Encuesta electrónica realizada a los docentes.	167
Apéndice I Encuesta electrónica realizada a los estudiantes.. . . .	177
Anexos	182
Anexo 1 Categorías de la Taxonomía de Bloom	183
Anexo 2 Guía general para la elaboración de informes	185
Anexo 3 Guías de Laboratorio	190
3.1 Guía de trabajo de la Práctica 1 - Determinación de la aceleración de la gravedad local.	190
3.2 Guía y pautas de corrección de informe: Péndulo Simple y Estadística.	194
3.3 Guía de trabajo de la Práctica 2 - Caracterización de dispositivos de medición de temperatura.	197
3.4 Pautas de corrección de informe: Caracterización de dispositivos de medición de temperatura.	203
3.5 Guía de trabajo de la Práctica 3 - Oscilaciones amortiguadas.	206
3.6 Guías de trabajo de la Práctica 4 - Determinación de la velocidad del sonido, Determinación del calor específico de un metal y Movimiento en un plano inclinado.	214
Anexo 4 Programa del curso Física Experimental 1	228

Capítulo 1

Introducción.

El laboratorio ocupa un lugar central en la enseñanza y el aprendizaje de la física (Hofstein & Lunetta, 2004) y constituye un componente relevante en la formación de futuros ingenieros. En Uruguay, la aproximación a los cursos de física experimental comienza en la educación media superior (APFU, 2002). Su importancia radica en que el trabajo experimental en el laboratorio involucra el desarrollo de habilidades propias de la formación en ciencias y, en particular, de la física (Kozminski et al., 2014). Estos espacios exigen a los estudiantes sostener, revisar y, en ocasiones, reformular sus procedimientos e interpretaciones frente a obstáculos, resultados imprevistos o discrepancias entre la evidencia experimental y las predicciones teóricas. Este proceso posee un valor formativo, en tanto promueve otros objetivos, como el modelado experimental, el pensamiento crítico y la comprensión de la naturaleza de la ciencia (Pollard et al., 2021).

En este contexto, la investigación en Educación en Física (*Physics Education Research*, PER) ha mostrado, desde sus inicios, un marcado interés por comprender cómo se enseña y aprende física. Como en toda actividad científica, la investigación en educación en física se configura como un campo empírico en el que se desarrollan estudios, se analizan datos y se elaboran conclusiones. Su desarrollo se vincula, en gran medida, con la insatisfacción frente a los resultados de aprendizaje obtenidos en las aulas tradicionales de física. En particular, se observó que muchos estudiantes resolvían problemas correctamente en términos numéricos, pero no lograban explicar los conceptos involucrados y, en algunos casos, mantenían interpretaciones conceptuales erróneas (Mestre & Docktor, 2021). En este campo, se ha consolidado además un área de interés

vinculada con la enseñanza y el aprendizaje de la física experimental en cursos de laboratorio, que se ha ocupado, entre otros aspectos, de analizar cómo los estudiantes desarrollan habilidades propias del trabajo experimental (May, 2023; Wei et al., 2026).

En PER, los cursos de laboratorio han recibido especial atención, ya que exigen que los estudiantes sinteticen información procedente de diversas fuentes y articulen habilidades matemáticas, computacionales y prácticas, integrando múltiples saberes. Al mismo tiempo, sus objetivos se han transformado históricamente en diálogo con los avances de la investigación sobre la enseñanza y el aprendizaje de la Física. En esta línea, la Asociación Americana de Profesores de Física (*American Association of Physics Teachers*, AAPT) publicó en 2014 el informe *Recommendations for the Undergraduate Physics Laboratory Curriculum*, en el que se proponen orientaciones destinadas a promover el desarrollo de habilidades y competencias propias del siglo XXI. Entre sus principales aportes, el informe reconoce seis áreas clave para los cursos introductorios de Física Experimental: la construcción del conocimiento, el modelado, el diseño de experimentos, el desarrollo de habilidades técnicas y prácticas de laboratorio, el análisis y la visualización de datos, y la comunicación en física (Kozminski et al., 2014).

El desarrollo de PER ha estado acompañado por la incorporación de nuevas herramientas orientadas a la investigación y a la enseñanza de la física. En particular, la Inteligencia Artificial (IA) ha ampliado las posibilidades de este campo, al tiempo que ha introducido nuevos interrogantes y desafíos. En la literatura reciente, sus aplicaciones incluyen el uso de técnicas de Procesamiento de Lenguaje Natural (PLN) para analizar grandes volúmenes de datos provenientes de producciones estudiantiles con fines de investigación. También se identifican implementaciones orientadas a la enseñanza, entre ellas la evaluación y calificación, la retroalimentación y el apoyo a las actividades de laboratorio (Küchemann & Kuhn, 2026).

En la enseñanza de la Física Experimental, la incorporación de IA abre nuevas posibilidades para analizar las producciones elaboradas por los estudiantes en los cursos de laboratorio. En ese marco, los informes elaborados en estos cursos ocupan un lugar central, en tanto constituyen una fuente valiosa de información sobre distintos aspectos del trabajo experimental y son utilizados como instrumento de evaluación tanto en la educación media como en la universitaria.

Esta investigación surge de una inquietud vinculada a las particularidades de la enseñanza de la física experimental. Si bien los informes permiten conocer el trabajo realizado por los estudiantes, suelen evaluarse de manera aislada, en relación con cada práctica, sin que a partir de ellos se construya necesariamente una mirada retrospectiva sobre la evolución del proceso de enseñanza y aprendizaje a lo largo del curso. Esto dificulta reconocer, desde una perspectiva global, qué habilidades se promueven, cuáles adquieren mayor presencia en las producciones estudiantiles y cuáles permanecen menos visibles. A esto se suma la falta de una visión general sobre las habilidades que docentes y estudiantes consideran más relevantes o identifican como efectivamente priorizadas durante el curso, y sobre la relación de estas habilidades con los objetivos del curso y con las guías de trabajo experimental proporcionadas a los estudiantes.

Este trabajo se propone responder una pregunta central: ¿cuáles son las habilidades científicas que se promueven en un curso introductorio universitario de Física Experimental y cómo se evidencian en los informes de laboratorio producidos por los estudiantes? De esta pregunta se desprenden otros interrogantes que orientan el análisis: ¿cómo se vinculan dichas habilidades con los objetivos del curso y con el diseño de las guías de trabajo práctico?, ¿cómo varía a lo largo del curso el grado en que estas habilidades se evidencian en los informes de laboratorio?, ¿qué percepciones tienen docentes y estudiantes acerca de los objetivos y habilidades que consideran relevantes en el curso?, y ¿cuáles de esos objetivos y habilidades perciben ambos grupos como efectivamente priorizados en el desarrollo del curso?

Con este propósito, se diseñó una metodología para automatizar el análisis de textos provenientes de informes de laboratorio, mediante el uso de técnicas de Procesamiento del Lenguaje Natural (PLN) aplicadas a la clasificación de texto. Para ello, se construyó y validó una guía de anotación sustentada en literatura específica, se seleccionó y anotó un conjunto de enunciados extraídos de diferentes secciones de los informes de laboratorio, denominado corpus, y se entrenaron y evaluaron distintos algoritmos de clasificación. Finalmente, los resultados obtenidos se contrastaron con los objetivos del curso, las características de las guías de laboratorio y las percepciones de docentes y estudiantes, con el fin de aportar una comprensión más amplia sobre las habilidades científicas promovidas en el curso y aquellas que efectivamente logran evidenciarse en las producciones escritas de los estudiantes.

Luego de esta introducción, el trabajo se organiza en siete capítulos. El

Capítulo 2 desarrolla el marco conceptual vinculado con la enseñanza de la física experimental, atendiendo a las investigaciones sobre los cursos de laboratorio, los objetivos que los orientan, las habilidades científicas que en ellos se promueven y la Taxonomía de Bloom revisada como marco teórico para la clasificación de objetivos. El Capítulo 3 incorpora el segundo eje teórico de la investigación, referido a la Inteligencia Artificial, con particular énfasis en el Procesamiento del Lenguaje Natural y en los métodos de clasificación de texto. En el Capítulo 4 se presenta el diseño de la investigación y el proceso de construcción del corpus anotado que sustenta el análisis. El Capítulo 5 trata sobre los métodos de clasificación automática utilizados, su evaluación y el correspondiente análisis de errores. A continuación, el Capítulo 6 analiza la relación entre objetivos y habilidades a partir del contraste entre el diseño del curso y las percepciones de docentes y estudiantes. El Capítulo 7 examina las habilidades identificadas en un curso de Física Experimental mediante el análisis automatizado de enunciados de informes de laboratorio, con foco en las habilidades prácticas de laboratorio y en la evaluación del diseño o del procedimiento experimental. Por último, el Capítulo 8 sintetiza los principales aportes de este estudio y plantea algunas perspectivas a futuro.

Capítulo 2

Enseñanza de la Física

Experimental: entre objetivos y habilidades.

El surgimiento del campo de Investigación en Educación en Física (PER) estuvo asociado, entre otros factores, a una creciente preocupación por las limitaciones de la enseñanza tradicional para favorecer una comprensión profunda de los conceptos físicos. En particular, diversas investigaciones mostraron que muchos estudiantes podían obtener resultados correctos en la resolución de problemas sin lograr explicar adecuadamente los conceptos involucrados (Mestre & Docktor, 2021). En este contexto, la enseñanza de la física experimental en cursos de laboratorio comenzó a ser objeto de revisión, especialmente en relación con los objetivos que orientan su diseño, las habilidades que se busca promover y las formas de evaluar los aprendizajes.

En este trabajo, se entiende por *objetivos de aprendizaje* a aquellos enunciados que explicitan lo que se espera que los estudiantes logren, comprendan o sean capaces de demostrar al finalizar un proceso de aprendizaje (Anderson & Krathwohl, 2001; Gogus, 2012). A su vez, el término *habilidad* se emplea para designar la facultad de una persona para desplegar los procesos cognitivos necesarios para la construcción de conocimiento y la resolución de problemas (Di Mauro et al., 2015; Etkina et al., 2006). Desde esta perspectiva, analizar los objetivos de los cursos de laboratorio permite reconocer qué dimensiones de la formación experimental se priorizan y cómo estas se integran con el desarrollo de habilidades científicas.

Con este propósito, este capítulo ofrece un panorama de investigaciones en educación en física experimental, revisa aportes de la literatura sobre los objetivos de los cursos de laboratorio y su relación con las habilidades científicas, e introduce la Taxonomía de Bloom revisada como marco para analizar y clasificar objetivos de aprendizaje.

2.1. Investigaciones en Educación en Física Experimental

La visión de docentes e investigadores sobre la enseñanza y el aprendizaje de la física experimental ha cambiado a lo largo de la historia. De acuerdo con la revisión de May (2023), en los primeros cursos de física experimental en Estados Unidos, desde principios del siglo XIX hasta la Revolución Industrial, predominaban las prácticas demostrativas dirigidas por el docente, en las que el estudiante adoptaba un rol principalmente pasivo como observador. Posteriormente, la enseñanza se desplazó hacia prácticas con estructura confirmatoria, conocidas comúnmente como "laboratorios de libro de cocina". Este último enfoque ha sido objeto de debate y cuestionamiento en la comunidad de enseñanza de la física, por parte de docentes, investigadores y científicos que han señalado que el objetivo central de las prácticas experimentales debería ser el desarrollo de la capacidad de indagación de los estudiantes. En la actualidad, los objetivos de los cursos de física experimental continúan siendo objeto de discusión, priorizándose el desarrollo de habilidades experimentales tales como el modelado científico, el diseño experimental y la comunicación, por encima de la reafirmación de contenidos teóricos (May, 2023).

A fines del siglo XX se inició un período de innovación y reforma en los cursos de física experimental en el ámbito académico principalmente estadounidense, impulsado por el diseño basado en la evidencia, y un creciente énfasis en los objetivos educativos. Varios grupos de PER comenzaron a aplicar técnicas de investigación para evaluar y mejorar tanto la enseñanza como el aprendizaje en los cursos de laboratorio (May, 2023). Algunos ejemplos destacados de las investigaciones y reformas en laboratorios de física introductoria son los siguientes:

- *Laboratorios de diálogo socrático* (Socratic Dialogue Labs, SDLs), desarrollados por Richard Hake en la década del 90. Estos laboratorios busca-

ban la construcción mental de conceptos en los estudiantes a través del conflicto conceptual, el análisis extenso con múltiples representaciones (verbal, diagramática, matemática) de experimentos newtonianos simples, la exposición repetida a estos, la discusión colaborativa entre pares y el diálogo socrático con instructores (Hake, 1992). Esta investigación mostró que los SDLs lograron mayores ganancias en el test Force Concept Inventory (FCI) (PhysPort, 2025) en comparación con los laboratorios tradicionales. Los SDLs presentaban limitaciones para el desarrollo de habilidades experimentales debido a su formato prescriptivo y al uso limitado de computadoras y sensores (May, 2023).

- *Física en tiempo real* (*Real Time Physics*, RTP) (Sokoloff et al., 2007) fue desarrollado en la década de 1990 como respuesta a una demanda creciente por utilizar los laboratorios para mejorar el aprendizaje conceptual e integrar computadoras y sensores digitales en la experimentación. Los objetivos de este proyecto incluían la comprensión de un conjunto de conceptos físicos relacionados, la experimentación con el mundo físico mediante microcomputadoras para la recopilación, visualización y análisis de datos en tiempo real, y el desarrollo de habilidades tradicionales de laboratorio mientras se abordaban los temas propios de las clases tradicionales, combinando actividades conceptuales y experimentos cuantitativos.
- *Entorno de Aprendizaje de la Ciencia por Investigación* (Investigative Science Learning Environment, ISLE) desarrollado por Eugenia Etkina y otros colaboradores con dos objetivos principales: ayudar a los estudiantes a aprender física mediante su participación en procesos que reflejen la práctica científica y mejorar su bienestar durante el aprendizaje (Etkina et al., 2025). Este enfoque didáctico promueve que los estudiantes, trabajando en grupos, construyan sus propias ideas mediante un ciclo en el que observan fenómenos y patrones, desarrollan explicaciones, hacen predicciones sobre el resultado de experimentos de prueba, evalúan si los resultados coinciden con sus predicciones y revisan sus explicaciones si es necesario. Las investigaciones muestran que los estudiantes que participan de laboratorios ISLE son más cuidadosos con las mediciones, consideran las suposiciones en el procedimiento matemático, evalúan los resultados y se comunican mejor (Etkina et al., 2010; Mestre & Docktor,

2021).

Aunque las reformas mencionadas se consolidaron hacia fines del siglo XX en el ámbito académico norteamericano, existen antecedentes previos en el proceso de renovación de la enseñanza introductoria de la física. En particular, el *Physical Science Study Committee* (PSSC) impulsó, en la década de 1950, una propuesta en la que el trabajo práctico de laboratorio ocupó un lugar central como espacio de exploración experimental y construcción conceptual, más allá de la mera verificación de contenidos (American Association of Physics Teachers, [s. f.](#) Editorial Reverté, [s. f.](#)). Su influencia se extendió rápidamente a otros países y contribuyó a instalar la necesidad de renovar la enseñanza de las ciencias también en América Latina. En Brasil, la traducción, adaptación y circulación de proyectos como el del PSSC contribuyeron a la formación de grupos universitarios dedicados a la enseñanza de la física, la formación docente y la producción de materiales didácticos. Estos procesos también favorecieron una mayor atención al papel de las actividades práctico-experimentales y a la consolidación de un campo específico de investigación (Pereira & Moreira, [2017](#); Villani et al., [2010](#)).

En América Latina, en países como Brasil (Ribeiro et al., [1997](#)), Colombia (Castiblanco & Vizcaíno, [2008](#)), Venezuela (Sebastià, [2006](#)) y Argentina (Becerra Labra & Silva Arias, [2024](#)), se han llevado a cabo trabajos en los que se analiza la necesidad de reformar los laboratorios de ciencias, en particular los de física, así como de modificar el uso de prácticas tipo “receta”. Para superar este problema, estos estudios aportan propuestas orientadas a transformar las rutinas mecánicas en experiencias auténticas, capaces de articular la teoría con la práctica y de fomentar tanto la adquisición de conocimientos científicos como una comprensión más profunda de la naturaleza de la ciencia. En esta misma línea, otros trabajos se han enfocado en explorar las percepciones de docentes y estudiantes sobre los trabajos prácticos de laboratorio y en analizar la introducción de cambios en las metodologías de enseñanza (Antúnez et al., [2008](#); Ortigoza et al., [2016](#); Petrucci et al., [2006](#); Sepúlveda Navarro & Carrasco Maturana, [2024](#)).

Con el transcurso del tiempo la comunidad de PER fue reconocida como un campo de investigación relevante dentro de la propia comunidad de físicos. En este proceso, numerosas universidades estadounidenses con departamentos de Física crearon institutos orientados al estudio de los procesos de enseñanza y

aprendizaje de la disciplina (Riveros, 2017). En la actualidad, la investigación en Educación en Física abarca los distintos niveles educativos y cuenta con revistas especializadas dedicadas a este campo. En el marco de esta consolidación, en 1998 la Asociación Americana de Profesores de Física (American Association of Physics Teachers, AAPT) publicó recomendaciones para el plan de estudios de laboratorio de física de pregrado (AAPT, 1998). Este documento introdujo una serie de objetivos para estos cursos, orientados al desarrollo de habilidades y competencias en los estudiantes como futuros físicos. Posteriormente, de acuerdo con May (2023), en el siglo XXI se impulsaron nuevas transformaciones en los cursos de laboratorio de física, lo que condujo a una mayor diversidad y a la aparición de temas clave, entre ellos:

- **El modelado**, orientado a que los estudiantes elaboraran modelos conceptuales de fenómenos físicos y los revisaran a partir de la experimentación.
- **La comunicación científica**, mediante la incorporación de actividades de redacción de informes y de registro en cuadernos de laboratorio, con el propósito de favorecer la alfabetización científica.
- **El pensamiento crítico, computacional y el análisis de datos**, orientados a que los estudiantes realizaran un análisis riguroso de los datos experimentales para extraer conclusiones y repensar las implicancias de sus hallazgos.
- **La experimentación colaborativa**, centrada en el trabajo en equipo y en la preparación para la labor en entornos de investigación científica.
- **El pensamiento interdisciplinario**, enfocado en establecer conexiones entre la física y otros campos del conocimiento.

Actualmente, las investigaciones sobre la enseñanza de la física en laboratorios evidencian una ampliación de las líneas de análisis orientadas a evaluar críticamente la eficacia de diversas estrategias de enseñanza. Según May (2023) existen varios temas destacados dentro de los que señala:

- La equidad y particularmente la relación con el género o factores demográficos.

- Estudio de enfoques y técnicas que inciden en las actitudes de los estudiantes hacia la física experimental.
- Uso de simulaciones para mejorar el aprendizaje conceptual.
- Uso de laboratorios remotos, muy útiles durante la pandemia COVID-19 y que lograron aproximadamente los mismos resultados en cuanto a las actitudes experimentales de estudiantes en comparación con los cursos presenciales.
- Mejora de la comprensión de los estudiantes sobre el error experimental y la incertidumbre de la medición.
- Habilidades de los estudiantes en el modelado científico, el diseño experimental, el análisis de datos y la interpretación, habilidades que contribuyen a "pensar como un físico".

A estos temas destacados hay que agregar las recientes investigaciones acerca del uso de INTELIGENCIA ARTIFICIAL (IA) en educación en física. La revisión de Bralin et al. (2024) identifica seis áreas temáticas en las que los investigadores de PER han empleado IA o APRENDIZAJE AUTOMÁTICO (AA)¹:

1. Desarrollo y análisis de evaluaciones. Es el área de mayor tamaño que puede subdividirse en caracterización de las respuestas de los estudiantes y evaluación en física con IA generativa, y análisis de evaluaciones estandarizadas.
2. Estudios relacionados al éxito estudiantil. Desarrollo de modelos predictivos del desempeño estudiantil, comparaciones entre clases tradicionales y remotas, exploración de la elección de los cursos de los estudiantes, entre otros.
3. Participación estudiantil. Exploración de la participación dentro y fuera de entornos académicos, en foros académicos, interacciones en los laboratorios, roles de los estudiantes en los grupos de trabajo en laboratorios de física, entre otros.

¹Estos conceptos serán desarrollados con mayor detalle en el [Capítulo 3](#)

4. Descripción general teórica y metodológica. Estudios que explican cómo realizar afirmaciones confiables al utilizar la IA para la codificación automática, evaluar modelos estadísticos, realizar análisis de redes sociales y analizar preguntas de múltiple opción.
5. Estudios que han utilizado IA para predecir factores que influyen en las admisiones universitarias.
6. Otros usos de la IA, menos explorados en este campo, incluyen la síntesis de literatura y la evaluación de la efectividad de intervenciones en los cursos, entre otros.

2.1.1. Objetivos de los cursos de laboratorio y las habilidades científicas

Las actividades experimentales de laboratorio se han considerado históricamente esenciales para el aprendizaje de la física y, a su vez, los objetivos de los cursos de laboratorio han evolucionado a lo largo del tiempo, en especial a partir de los avances en la investigación sobre la enseñanza y el aprendizaje de la física. Hasta 1970, según Trumper (2003), los propósitos de los trabajos de laboratorio podían agruparse en cuatro categorías: (i) *habilidades* de uso y manipulación de instrumentos, de indagación, comunicación, pensamiento crítico y resolución de problemas; (ii) aplicación de *conceptos* aprendidos y descubrimiento de nuevos conceptos; (iii) comprensión de *la naturaleza de la ciencia*, su desarrollo y el trabajo de los científicos; y (iv) el desarrollo de *actitudes* como curiosidad, objetividad, precisión y cooperación para trabajar en equipo.

A partir de la influencia de los aportes constructivistas de Vygotsky (1978) y del reconocimiento de la importancia de la interacción social en los procesos de aprendizaje, el laboratorio comenzó a concebirse en Estados Unidos como un espacio propicio para favorecer dicha interacción y promover el desarrollo cognitivo de los estudiantes (Trumper, 2003). Desde esta perspectiva, en 1998, la AAPT desarrolló un conjunto de objetivos comunes para la enseñanza en el laboratorio, los cuales han servido como referencia para orientar y evaluar programas introductorios de laboratorio. Los objetivos identificados en este documento son *El Arte de la Experimentación*, *Habilidades técnicas y analíticas*, *Aprendizaje Conceptual*, *Comprensión de las bases del conocimiento en físi-*

ca y Desarrollo de habilidades de aprendizaje colaborativo. En conjunto, estos objetivos apuntan a involucrar a los estudiantes en experiencias significativas de investigación, favorecer el desarrollo de herramientas y habilidades básicas de la física experimental, fortalecer el dominio de los conceptos fundamentales, promover la distinción entre inferencias basadas en teorías y resultados experimentales, y desarrollar capacidades de trabajo colaborativo necesarias para su formación. Además, el documento sugiere que los logros del trabajo de laboratorio se evidencien mediante la comunicación de resultados, a través de la elaboración de informes escritos y la realización de presentaciones orales en diversos formatos.

En relación con lo anterior, otros autores también han propuesto clasificaciones para los objetivos de los cursos de laboratorio. De acuerdo con Wilcox y Lewandowski (2017) existen al menos tres categorías: reforzar los conceptos que se enseñan en los cursos teóricos, desarrollar habilidades prácticas de laboratorio, o fomentar creencias expertas sobre la naturaleza y la importancia de la física experimental. Estos objetivos no son excluyentes, aunque el diseño de actividades y evaluaciones para determinar el éxito de estos objetivos varía según la prioridad que se le otorgue a cada uno.

En el contexto latinoamericano Borges (2002) ha señalado que los objetivos del laboratorio de ciencias suelen permanecer implícitos y son escasamente discutidos en la planificación docente, lo que favorece que los estudiantes asocien las actividades prácticas principalmente con la verificación y la demostración. Entre los propósitos tradicionalmente atribuidos al laboratorio, el autor menciona la verificación de teorías, la enseñanza del método científico, la facilitación del aprendizaje conceptual y el desarrollo de habilidades prácticas y técnicas, aunque advierte cierta ambigüedad en relación a cuáles serían efectivamente estas últimas. En una línea similar, Werner da Rosa y Becker da Rosa (2007) encuentran que, en cursos de física para ingeniería, los docentes priorizan como objetivos del laboratorio el manejo de equipos e instrumentos de medición, la observación e interpretación de fenómenos, la articulación entre teoría y práctica, y la motivación por el estudio de la física. Complementariamente, Carrascosa et al. (2006) cuestionan las prácticas centradas en la mera verificación y proponen orientar el laboratorio como una actividad de investigación, articulada con la formulación de problemas, la construcción de hipótesis, el diseño de estrategias de resolución, la interpretación de resultados y la comunicación de estos.

En 2014, la AAPT presentó un nuevo documento titulado *Recommendations for the Undergraduate Physics Laboratory Curriculum*, en el que incluyó una serie de recomendaciones orientadas a promover el desarrollo de habilidades y competencias del siglo XXI. Estas sugerencias se fundamentan en una revisión del estado del currículo de los laboratorios de física de pregrado y en investigaciones vinculadas con la enseñanza de esta disciplina (Kozminski et al., 2014). De acuerdo con este documento, los objetivos de los cursos introductorios de laboratorio se sustentan en la *Construcción del conocimiento* y abarcan propósitos más generales organizados en cinco áreas clave: *Modelado*, *Diseño de Experimentos*, *Desarrollo de habilidades técnicas y prácticas de laboratorio*, *Análisis y visualización de datos* y *Comunicación de la física* (Tabla 2.1). Estas áreas agrupan recomendaciones orientadas a que los estudiantes de física y de otras ciencias aprendan a pensar como físicos y a desarrollar investigaciones experimentales acordes con la práctica científica.

Las recomendaciones de la AAPT también resultan útiles a los profesores, ya que brindan orientaciones para el desarrollo de habilidades de laboratorio a la hora de diseñar sus cursos. Estas recomendaciones también promueven que estudiantes de otras disciplinas se inicien en la realización de investigaciones experimentales desde una perspectiva propia de la física, y que amplíen su comprensión de la naturaleza de la ciencia.

Ante la ausencia de una definición única de *habilidad*, en este trabajo la *habilidad científica* se entiende como la facultad de una persona para aplicar procesos cognitivos relacionados con la construcción del conocimiento científico (Di Mauro et al., 2015; Etkina et al., 2006). Se utiliza el término *habilidad* para describir algunos de los procedimientos, procesos y métodos necesarios que los científicos emplean para la construcción del conocimiento y resolución de problemas experimentales. Para favorecer el desarrollo de estas habilidades, es necesario involucrar a los estudiantes en actividades de aprendizaje apropiadas y asegurar que comprendan con claridad los objetivos, las habilidades que se espera que desarrollen y los criterios que definen un desempeño de calidad. Además, ofrecer una retroalimentación oportuna e incorporar espacios de autoevaluación basados en criterios previamente establecidos permite fortalecer el proceso de desarrollo de estas habilidades (Etkina et al., 2006).

Tabla 2.1: Áreas de enfoque en los cursos de laboratorio introductorios adaptadas de Kozminski et al., 2014. Fuente: Elaboración propia.

Área	Descripción
Construcción de conocimiento	Los estudiantes deben adquirir conciencia de que son capaces de hacer ciencia, recopilando información, analizando e interpretando datos reales de manera ética. Se fomenta que la construcción del conocimiento no dependa de la autoridad externa.
Modelado	Un modelo implica un vínculo entre la teoría y el experimento. Los estudiantes deben ser capaces de desarrollar modelos para representar sistemas físicos, interpretar y predecir resultados en el contexto del experimento de laboratorio. Además deben reconocer las limitaciones del modelo, de los dispositivos de medición y sus incertidumbres.
Diseño de experimentos	Los estudiantes deben ser capaces de crear preguntas científicas, diseñar y desarrollar experimentos para responder a las mismas. Resolver desafíos en la ejecución de experimentos, con el fin de evaluar modelos e hipótesis, teniendo en cuenta restricciones como presupuesto, tiempo, seguridad y recursos disponibles.
Desarrollo de habilidades técnicas y prácticas de laboratorio	Los estudiantes deben alcanzar un manejo competente de los instrumentos comúnmente utilizados en mediciones estándar de laboratorio. Deben comprender las limitaciones de sus dispositivos de medición y cómo elegir el equipo adecuado para mediciones específicas.
Análisis y visualización de datos	Los estudiantes deben ser capaces de procesar y presentar datos mediante herramientas estadísticas, interpretando de manera crítica su validez, sus limitaciones y las incertidumbres asociadas. Elegir métodos gráficos apropiados para representar datos, ajustarlos y extraer cantidades físicas de los parámetros del ajuste. También deben poder cuantificar las incertidumbres y propagarlas realizando los cálculos correspondientes, así como comparar sus resultados experimentales con modelos matemáticos, computacionales o simulaciones.

Continúa en la siguiente página

Área	Descripción
Comunicación en física	Los estudiantes deben aprender a presentar argumentos respaldados por evidencia experimental, utilizando formatos escritos u orales adecuados. Estos argumentos deben incluir elementos como gráficos, tablas, resultados numéricos con incertidumbres y diagramas. Deben realizar presentaciones propias de la disciplina, como informes técnicos, artículos estilo revista científica y presentaciones orales o en póster. Las habilidades de comunicación interpersonal también deben desarrollarse en el laboratorio mediante el trabajo en equipo y la colaboración.

2.1.2. Objetivos y la Taxonomía de Bloom

Las taxonomías constituyen herramientas de clasificación, y la Taxonomía de Bloom es una de las más reconocidas en el ámbito educativo. De acuerdo con Bloom (1990, p. 13), el propósito principal de elaborar una taxonomía de objetivos educativos es facilitar la comunicación, favoreciendo el intercambio de ideas y materiales entre investigadores en educación y responsables del desarrollo curricular. En este marco, la Taxonomía de Bloom (TB) permite clasificar los resultados de aprendizaje esperados, entendidos como los modos en que se espera que el estudiantado actúe, piense o sienta como consecuencia de su participación en una secuencia de enseñanza y aprendizaje (Bloom, 1990, p. 14).



Figura 2.1: Jerarquía de las categorías de la Taxonomía de Bloom, 1990. Fuente: Elaboración propia.

La TB contiene seis categorías: *Conocimiento*, *Comprensión*, *Aplicación*, *Análisis*, *Síntesis* y *Evaluación*, organizadas de manera jerárquica, tal como se

muestra en la [Figura 2.1](#). En consecuencia, los objetivos ubicados en una categoría determinada suponen el dominio de procesos asociados a las categorías anteriores. En el Anexo 1 se puede consultar la [Tabla 1.1](#) donde se describen cada una de las categorías.

Cada categoría del dominio cognitivo de la TB está asociada a verbos utilizados en el diseño de objetivos de aprendizaje. En la [Tabla 2.2](#) se muestran algunos de los verbos comprendidos en cada categoría de acuerdo con Ferraz y Belhot (2010).

Tabla 2.2: Verbos relacionados a cada categoría de la TB. Tabla adaptada de Ferraz y Belhot (2010, p.426)

Conocimiento	Comprensión	Aplicación	Análisis	Síntesis	Evaluación
enumerar	convertir	aplicar	analizar	categorizar	evaluar
definir	decodificar	programar	reducir	combinar	escoger
describir	definir	emplear	clasificar	compilar	concluir
identificar	describir	manipular	comparar	crear	contrastar
denominar	distinguir	operar	contrastar	construir	criticar
nombrar	discriminar	resolver	deducir	elaborar	decidir
combinar	estimar	construir	diagramar	formular	defender
resaltar	explicar	modificar	distinguir	generalizar	justificar
recordar	generalizar	producir	inferir	inventar	validar
relacionar	ejemplificar	esbozar	diferenciar	planear	juzgar
reproducir	ilustrar	practicar	relacionar	proponer	argumentar
declarar	reformular	usar	separar	reorganizar	recomendar
etiquetar	seleccionar	ejecutar	subdividir	reescribir	
memorizar	traducir		discriminar	estructurar	
ordenar	clasificar		esquematizar	proyectar	
reconocer	resumir		examinar		
listar	interpretar		testear		
	parafrasear		cuestionar		

Cuarenta y cinco años más tarde, la Taxonomía de Bloom fue revisada por Anderson y Krathwohl (2001) (TBA), refinando las categorías con el objetivo de reflejar mejor la complejidad del pensamiento humano y el aprendizaje. Esta versión actualizada puede servir para múltiples propósitos, entre ellos, proporcionar un lenguaje común para los objetivos de aprendizaje, que facilite la comunicación entre personas, materias y niveles de grado, y actuar como un mecanismo para determinar la congruencia entre los objetivos educativos (orientados al aprendizaje de los estudiantes), las actividades y las evaluaciones dentro de una unidad, curso, asignatura o plan de estudios (Gamboa Solano et al., 2023).

La TBA no indica lo que los estudiantes deben aprender, pero facilita la escritura de los estándares de aprendizaje a un lenguaje centrado en logros. De

esta forma ofrece una perspectiva útil para orientar las decisiones curriculares. Esta taxonomía se organiza en una estructura bidimensional, compuesta por una *dimensión* referida al *conocimiento* y otra a los *procesos cognitivos*. La intersección de ambas dimensiones da lugar a una matriz que permite clasificar los objetivos de aprendizaje.

La dimensión del **proceso cognitivo** comprende seis categorías: *Recordar*, *Comprender*, *Aplicar*, *Analizar*, *Evaluar* y *Crear*. En estas categorías subyace un supuesto de continuidad en la complejidad cognitiva. *Comprender* se considera más complejo que *Recordar* y así sucesivamente con el resto de las categorías.

La dimensión del **conocimiento**, por su parte, se divide en cuatro categorías: *Fáctico*, *Conceptual*, *Procedimental* y *Metacognitivo*. Estas se organizan en un continuo que va desde lo más concreto hasta lo más abstracto. Si bien existe un aumento general de abstracción, puede producirse superposición entre categorías, por ejemplo, cuando ciertos conocimientos procedimentales resultan más concretos que algunos conocimientos conceptuales.

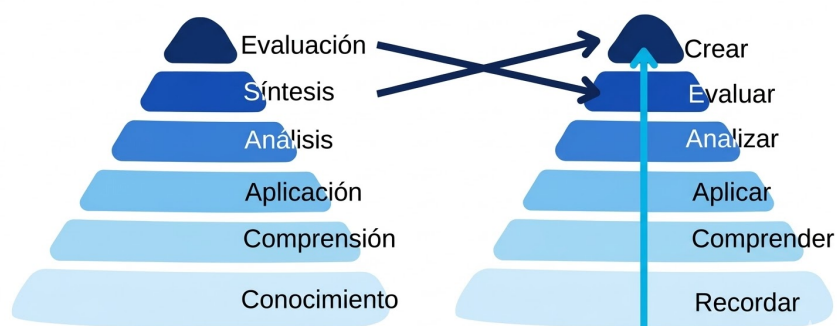


Figura 2.2: Modificaciones en las dimensiones del proceso cognitivo de la TBA (Anderson & Krathwohl, 2001). Fuente: Elaboración propia.

La separación explícita entre conocimiento y proceso cognitivo constituye uno de los cambios fundamentales respecto a la taxonomía original de Bloom. Además, se realizó la modificación de las denominaciones de las categorías a formas verbales para enfatizar la acción cognitiva: *Conocimiento* se cambió por *Recordar*, y las categorías *Evaluación* y *Síntesis* fueron reemplazadas y reorganizadas como *Evaluar* y *Crear* (Ferraz & Belhot, 2010). Estas modificaciones se ilustran en la [Figura 2.2](#).

Para clasificar un objetivo de aprendizaje en la tabla de la TBA es ne-

cesario identificar dos componentes: el proceso cognitivo que se espera que el estudiante realice y el tipo de conocimiento sobre el cual opera dicho proceso. A continuación se describen las categorías de la dimensión del proceso cognitivo y posteriormente, las categorías de la dimensión del conocimiento.

De acuerdo con Anderson y Krathwohl (2001) cuando los objetivos promueven la retención del material presentado tal como se enseñó, la categoría del **proceso cognitivo** relevante es *Recordar*. Recordar implica recuperar conocimiento relevante de la memoria a largo plazo. Esta recuperación constituye un proceso esencial para el aprendizaje, ya que interviene en la resolución de problemas más complejos.

La siguiente categoría dentro de los procesos cognitivos es *Comprender*. Los estudiantes demuestran comprensión cuando construyen significado a partir de distintos mensajes, utilizando diversas formas de comunicación. Esta comprensión se manifiesta cuando logran establecer conexiones entre el conocimiento previo y el conocimiento “nuevo” (Anderson & Krathwohl, 2001). *Comprender* implica una variedad de habilidades cognitivas, tales como interpretar información al transformarla de una forma de representación a otra, ejemplificar mediante casos específicos, clasificar a partir de la pertenencia a una categoría determinada, resumir mediante la elaboración de una interpretación general y la selección de ideas principales, inferir patrones en una serie de ejemplos o instancias, comparar similitudes y diferencias entre objetos, eventos, ideas, problemas o situaciones, y explicar mediante modelos de causa y efecto.

La categoría *Aplicar* involucra usar procedimientos para resolver ejercicios y problemas y puede subdividirse en: ejecutar e implementar. Ejecutar está asociado al uso de técnicas y métodos, a una secuencia de pasos ordenada para resolver un ejercicio (una situación que se considera familiar). Implementar implica que el estudiante seleccione y use un procedimiento para realizar una tarea desconocida (Anderson & Krathwohl, 2001).

Analizar implica dividir el material en sus partes constituyentes y determinar cómo se relacionan las partes entre sí y con una estructura general. Esta categoría incluye procesos cognitivos como diferenciar, organizar y atribuir.

Emitir juicios basados en criterios y estándares define a la categoría *Evaluar*. Uno de los procesos cognitivos asociados a esta categoría es comprobar. El estudiante comprueba si existen inconsistencias internas en una operación o producto, si los datos apoyan o refutan una hipótesis, o si el material presenta-

do tiene partes que se contradicen entre sí. Otro proceso cognitivo asociado a evaluar es criticar o juzgar. Un estudiante critica cuando identifica las características positivas o negativas de un producto y emite un juicio basado al menos parcialmente en esas características. Este proceso también suele denominarse pensamiento crítico (Anderson & Krathwohl, 2001).

La última categoría de los procesos cognitivos es **Crear**. Según Anderson y Krathwohl (2001) *Crear* supone unir elementos para formar un todo coherente o funcional. Los objetivos que se encuentran dentro de esta categoría implican que los estudiantes creen un nuevo producto.

En la **dimensión del conocimiento** se encuentran las categorías Fáctico, Conceptual, Procedimental y Metacognitivo. El conocimiento **Fáctico** es el conocimiento de contenidos discretos, aislados, de fragmentos de información, lo que incluye el conocimiento de terminología, detalles y lenguaje específico. El conocimiento **Conceptual** implica el conocimiento de formas más complejas y organizadas como clasificaciones, categorías, principios, modelos, estructuras. El conocimiento **Procedimental** está asociado a saber “cómo hacer algo”, incluye el conocimiento de técnicas, algoritmos, métodos y criterios para decidir “cuándo hacer qué” en un contexto disciplinar específico. Finalmente, el conocimiento **Metacognitivo** implica la conciencia y el conocimiento de la propia cognición, lo que se relaciona con que los estudiantes sean más conscientes y responsables de su propio aprendizaje y de sus procesos de pensamiento.

La [Figura 2.3](#) muestra la tabla de la taxonomía que puede ser utilizada por los docentes para visualizar con mayor claridad la variedad de posibles objetivos de aprendizaje. Al analizar los objetivos de un curso mediante esta tabla, se pueden identificar si existen dimensiones con numerosas entradas en filas y columnas, así como áreas de ciertas dimensiones sin entradas, lo que podría indicar que ciertos objetivos no han sido considerados.

Una vez que los objetivos se han colocado en una celda particular de la Tabla de la [Figura 2.3](#) se abordan las estrategias de enseñanza para ayudar a los estudiantes a lograr los objetivos. Diferentes objetivos requieren diferentes estrategias de enseñanza, actividades de aprendizaje, materiales curriculares y formas de evaluar. Es necesario entonces lograr cierto grado de alineación o correspondencia entre los objetivos, las estrategias de enseñanza y la evaluación.

La enseñanza y el aprendizaje de la física requiere del desarrollo de habilidades de pensamiento de orden superior, como la capacidad de analizar

TABLA DE LA TBA						
DIMENSIÓN DEL CONOCIMIENTO	DIMENSIÓN DEL PROCESO COGNITIVO					
	1. RECORDAR	2. COMPRENDER	3. APLICAR	4. ANALIZAR	5. EVALUAR	6. CREAR
A. CONOCIMIENTO FÁCTICO						
B. CONOCIMIENTO CONCEPTUAL						
C. CONOCIMIENTO PROCEDIMENTAL						
D. CONOCIMIENTO METACOGNITIVO						

Figura 2.3: Tabla con las dimensiones de la TBA traducida al español, tomada y adaptada de Anderson y Krathwohl (2001, p. 28)

situaciones, formular hipótesis y evaluar resultados experimentales. Estas habilidades están estrechamente relacionadas con los tres niveles de procesos cognitivos superiores de la TBA que requieren que los estudiantes piensen en forma crítica y creativa (Bhakti et al., 2025). La literatura en PER muestra que esta forma de clasificación se ha utilizado con diversos fines, algunos de ellos son: el uso de la TBA para el desarrollo de evaluaciones (Poluakan et al., 2019; Qadar et al., 2025), el análisis de preguntas de exámenes a partir de los verbos pertenecientes a las diferentes categorías para identificar los dominios a los que pertenecen (Damayanti et al., 2020; Dias et al., 2021; Ding, 2014; Nurlailiyah et al., 2019; Silva & Martins, 2014; Thacker, 2023), el diseño de tareas que aborden diferentes niveles de la TBA (Bandyopadhyay et al., 2023; Wijesinghe et al., 2025), evaluar la calidad de preguntas generadas por estudiantes (Bates et al., 2014), evaluar los aprendizajes de los estudiantes al realizar reformas en las metodologías de enseñanza como el uso de laboratorios virtuales (Gunawan et al., 2018) o predecir el rendimiento académico de estudiantes (Suleimeno-va et al., 2025). Además de estos trabajos, se encuentran otros que utilizan la TBA para analizar la enseñanza y el aprendizaje de la física, e incorporan

herramientas de IA. Algunas de las temáticas encontradas son: el uso de IA para el diseño de lecciones con objetivos de aprendizaje basados en la TBA (Liu, [2025](#)), el desarrollo de robots interactivos para brindar asistencia a los laboratorios de física de los estudiantes (Latif et al., [2024](#)) o la mejora de la calidad y relevancia de las preguntas generadas automáticamente adaptadas a las necesidades educativas (Luo et al., [2024](#); Omopekunola & Kardanova, [2026](#)).

Capítulo 3

La Inteligencia Artificial y la investigación en Educación en Física.

La denominación *Homo Sapiens* –el hombre sabio– da cuenta del lugar central que ocupa nuestra inteligencia en la definición de nuestra especie. A lo largo de la historia, los seres humanos han buscado comprender los procesos implicados en el pensamiento y el actuar. Por su parte, el campo de la Inteligencia Artificial (IA) se ocupa tanto de la comprensión de la inteligencia como de la construcción de entidades inteligentes, entendidas como máquinas capaces de actuar de manera eficaz y segura ante una amplia variedad de situaciones novedosas (Russell & Norvig, 2010). La definición de IA puede formularse en relación con los procesos de pensamiento y razonamiento, o bien en términos de comportamiento. A su vez, estas concepciones se articulan como muestra la [Figura 3.1](#) con dos criterios para valorar su desempeño: la fidelidad al comportamiento humano y la racionalidad, entendida como la capacidad de actuar correctamente (Russell & Norvig, 2010).

Desde esta perspectiva, Russell y Norvig (2021) plantean que, para pensar, una máquina debería contar con ciertas capacidades. Entre ellas se encuentran cuatro capacidades principales: el *Procesamiento del Lenguaje Natural* (PLN), que le permitiría comunicarse eficazmente en un lenguaje humano, el *conocimiento*, entendido como la capacidad de almacenar lo que sabe o percibe, la *automatización*, que es el razonamiento necesario para responder preguntas y extraer nuevas conclusiones, y el *aprendizaje automático*, que le permitiría

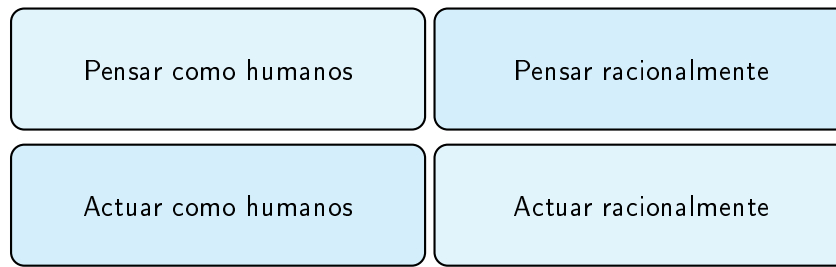


Figura 3.1: Enfoques para la definición de la inteligencia artificial. Adaptado de Russell y Norvig (2010, p. 2).

adaptarse a nuevas circunstancias, detectar patrones y extrapolarlos.

La IA también puede definirse, de acuerdo con Russell y Norvig (2021, p.7), “como el estudio de agentes que reciben percepciones del entorno y realizan acciones. Cada uno de estos agentes implementa una función que asigna secuencias de percepciones a acciones, y abarca diferentes maneras de representar estas funciones, como agentes reactivos, planificadores en tiempo real, sistemas de teoría de decisiones y sistemas de aprendizaje profundo”¹. Esta definición concibe a la IA como el campo dedicado al estudio, diseño y desarrollo de sistemas o agentes capaces de percibir información del entorno, aprender o adaptar su comportamiento y actuar de manera orientada al logro de determinados objetivos (Bartneck et al., 2021). En particular, en este trabajo las entradas del sistema corresponden a enunciados extraídos de informes de laboratorio, mientras que su salida consiste en la predicción de las etiquetas asociadas a las clases definidas para identificar distintas habilidades científicas utilizando técnicas de PLN.

La integración de la IA en la enseñanza en general, y en la enseñanza de la física en particular, ofrece numerosas posibilidades. Entre ellas, se destaca el desarrollo de sistemas de tutoría inteligente orientados a emular la interacción con un tutor humano mediante retroalimentación y orientación personalizadas, ajustadas al desempeño del estudiante y a sus diversas formas de aprender. El uso de IA también posibilita la generación automática de preguntas y problemas, lo que amplía y diversifica los recursos disponibles para la evaluación. Además, los sistemas de evaluación automatizada permiten analizar el rendi-

¹“We define AI as the study of agents that receive percepts from the environment and perform actions. Each such agent implements a function that maps percept sequences to actions, and we cover different ways to represent these functions, such as reactive agents, real-time planners, decision-theoretic systems, and deep learning systems.” Russell y Norvig (2021, p. 7).

miento de los estudiantes, monitorear su progreso e identificar áreas de mejora, lo que facilita intervenciones oportunas y apoyos específicos por parte de los docentes. La IA también contribuye al desarrollo de simulaciones virtuales que favorecen la visualización y comprensión de conceptos físicos abstractos, al tiempo que fortalecen las habilidades de resolución de problemas (Cordero, 2024; Santos et al., 2025).

El Aprendizaje Automático (AA) constituye un área centrada en el desarrollo de sistemas capaces de construir modelos a partir de datos y mejorar su desempeño mediante la experiencia (Russell & Norvig, 2021). En el ámbito educativo, ello permite analizar grandes volúmenes de datos estudiantiles para identificar patrones y tendencias que podrían no resultar evidentes a simple vista, así como examinar respuestas a problemas abiertos no solo con el objetivo de corregirlas, sino también de caracterizar la calidad del razonamiento y el grado de comprensión conceptual evidenciado (Cordero, 2024; Santos et al., 2025).

La implementación de la IA en la enseñanza, y en particular de la enseñanza de la física, también plantea desafíos, entre ellos la disponibilidad de infraestructura técnica adecuada, la capacitación y el acompañamiento a los docentes, la privacidad y seguridad de los datos, y la incorporación explícita de criterios éticos para el uso de estas tecnologías. La literatura señala además, líneas de trabajo para la investigación futura, en particular la necesidad de estudios longitudinales que permitan evaluar efectos sostenidos del uso de IA tanto sobre los aprendizajes como en las prácticas docentes (Mahligawati et al., 2023; Santos et al., 2025). También se indica que debe priorizarse el análisis de estrategias para mitigar sesgos y posibles efectos discriminatorios asociados a estos sistemas, con el fin de resguardar la equidad en el acceso y en las oportunidades de aprendizaje. De manera complementaria, comprender las percepciones, actitudes y niveles de aceptación de estudiantes y docentes frente al uso de IA en la enseñanza y el aprendizaje de la física constituye un insumo clave para orientar decisiones pedagógicas y de implementación.

Este trabajo tiene como objetivo analizar las producciones escritas de estudiantes, en particular los enunciados presentes en informes de laboratorio. Por este motivo, se enmarca en una disciplina específica de la IA, el *Procesamiento del Lenguaje Natural* (PLN). En el apartado siguiente se introduce el PLN como enfoque para el análisis de textos.

3.1. El Procesamiento del Lenguaje Natural.

El PROCESAMIENTO DEL LENGUAJE NATURAL (PLN) es una subdisciplina de la IA que intenta resolver con computadoras tareas vinculadas al lenguaje humano, permitiendo la comunicación entre el humano y la computadora a través del lenguaje natural o resolviendo tareas que implican algún tipo de procesamiento de texto o habla¹. Sus orígenes suelen situarse a comienzos de la década de 1950, a partir del test de Turing ²(Jurafsky & Martin, 2009).

El PLN permite a las computadoras comprender, analizar e interactuar con el lenguaje humano. Para ello, clásicamente se distinguen distintas etapas de análisis que permiten estudiar componentes del lenguaje y favorecen la comprensión de su estructura y significado. Estas etapas son:

- Fonética y Fonología: estudio de los sonidos lingüísticos.
- Morfología: estudio de la estructura interna de las palabras.
- Sintaxis: estudio de la estructuración (orden y agrupamiento) de las palabras en unidades mayores.
- Semántica: estudio del significado.
- Discurso: estudio de las unidades mayores a la oración.
- Pragmática: estudio de cómo el lenguaje se utiliza para cumplir objetivos.

Muchos de los problemas abordados por el PLN pueden formularse como problemas de *clasificación*. La clasificación consiste en asignar a una entidad una o más etiquetas de un conjunto previamente definido, a partir de la información contenida en un conjunto de datos. Por ejemplo, para identificar correos electrónicos no deseados, es posible entrenar un algoritmo utilizando un conjunto previamente etiquetado, en el que cada correo recibe una etiqueta [+spam] o [-spam]. En este caso, la detección de correo no deseado se apoya en la presencia de ciertas palabras o frases específicas en los correos, las cuales

¹“The goal of this new field is to get computers to perform useful tasks involving human language, tasks like enabling human-machine communication, improving human-human communication, or simply doing useful processing of text or speech.” (Jurafsky & Martin, 2009)

²Propuesto por Alan Turing en 1950, el test consiste en evaluar si una máquina puede exhibir un comportamiento indistinguible del de un ser humano, considerándose que lo supera cuando un evaluador humano no logra determinar, a partir de un intercambio escrito, si las respuestas provienen de una persona o de una computadora.

son aprovechadas por el clasificador para identificar casos positivos o negativos (es decir, spam o no spam) (Pustejovsky & Stubbs, 2014, p.22). Este ejemplo corresponde a un enfoque de aprendizaje automático supervisado, aunque la detección de correo no deseado también puede abordarse mediante otros enfoques. La clasificación en PLN comprende, además, otras tareas, como el análisis de sentimiento y la clasificación por temas predefinidos.

El AA ha sido uno de los enfoques predominantes en el desarrollo reciente del PLN. En este ámbito, su aplicación a conjuntos de datos suele estructurarse en torno a dos grandes enfoques: el aprendizaje supervisado y el aprendizaje no supervisado. El *aprendizaje supervisado* comprende métodos que permiten generar una función capaz de relacionar datos de entrada con un conjunto predefinido de etiquetas. Estas etiquetas representan los resultados esperados y suelen ser asignadas por personas que anotan manualmente los datos utilizados para entrenar el modelo. El *aprendizaje no supervisado*, por su parte, incluye métodos destinados a identificar estructuras o patrones en conjuntos de datos sin etiquetar (Pustejovsky & Stubbs, 2014, p.21). Más recientemente, se ha desarrollado el *aprendizaje autosupervisado*, en el que el algoritmo aprende características discriminativas a partir de grandes cantidades de instancias sin etiquetar, sin depender de anotaciones humanas (Gui et al., 2024).

3.2. Métodos de clasificación basados en aprendizaje automático supervisado y su evaluación.

Para que un algoritmo de clasificación logre capturar adecuadamente el comportamiento deseado, los datos empleados en su entrenamiento deben ser lo suficientemente variados y representativos. En la [Figura 3.2](#) se muestra la metodología que Pustejovsky y Stubbs (2014, p.23) denominan MATTER, cuyo acrónimo describe las etapas del proceso: *Modelado, Anotación, Entrenamiento, Testeo, Evaluación y Revisión*. En esta metodología, las clases o categorías empleadas en la anotación generalmente provienen de un modelado teórico previo. La descripción de cada clase constituye la base para la etapa de anotación y se integran en un ciclo iterativo de desarrollo, en el cual se entrenan y evalúan algoritmos destinados para anotar o etiquetar automáticamente textos.



Figura 3.2: Ciclo MATTER adaptado de Pustejovsky y Stubbs (2014)

Modelar el fenómeno es el primer paso. Según Pustejovsky y Stubbs (2014), un *modelo* es una caracterización de un determinado fenómeno expresada en términos más generales que los elementos en el dominio que se está modelando. Luego de contar con un modelo inicial para los fenómenos asociados con la tarea a resolver, se obtiene una primera especificación de clases o etiquetas para la anotación. Este documento, a partir del cual se planifica la anotación de acuerdo con las características del modelo, se denomina GUÍA DE ANOTACIÓN (Pustejovsky & Stubbs, 2014, p.27).

Los conjuntos de datos se denominan CORPUS. Cuando los datos se etiquetan como miembro de alguna de estas clases, entonces se ha realizado la ANOTACIÓN. En este trabajo, el esquema de anotación admitió etiquetado múltiple, es decir, a cada enunciado se le asignaron tantas etiquetas como clases se identificaron en él. Un conjunto de datos anotado con la misma especificación se denomina CORPUS ANOTADO. Es necesario que todos los anotadores humanos sean capacitados para que marquen el conjunto de datos según las clases que se han definido. Esto es bastante complejo, ya que requiere múltiples iteraciones entre el modelado y la anotación. La *guía de anotación* orienta a los anotadores para reconocer en el texto los elementos y características que deben asociarse a las distintas etiquetas.

Para evaluar la concordancia entre anotadores en una tarea de anotación, se utiliza el *Acuerdo Interanotador* (AIC, también denominado acuerdo intercodificador), que permite comparar las decisiones tomadas por distintos anota-

dores. Si distintos anotadores producen resultados consistentemente similares, puede inferirse que aplican de manera semejante las directrices establecidas en la guía de anotación (Artstein & Poesio, 2008). Lograr un buen AIC indica que la tarea está adecuadamente definida, que otros anotadores podrían continuar con el trabajo y que las instrucciones son interpretadas de manera similar (Pustejovsky & Stubbs, 2014, p.27), lo que aporta evidencia sobre la reproducibilidad de la tarea. Por otra parte, un bajo acuerdo entre anotadores también puede reflejar que la tarea presenta un alto grado de complejidad o subjetividad, sin que ello implique necesariamente que esté mal definida.

Calcular el AIC no consiste en contar cuántas etiquetas hay en el conjunto de datos y calcular cuántas veces coinciden los anotadores, porque el uso de estos porcentajes directos no tiene en cuenta los acuerdos aleatorios que probablemente ocurran al anotar textos. Para crear cifras que puedan compararse entre estudios se han desarrollado métricas estadísticas. En PLN, entre las métricas más utilizadas se encuentran la *Kappa de Cohen* y algunas de sus variantes, como la *Kappa de Fleiss*, así como el *Alfa de Krippendorff*.

La *Kappa de Cohen* (κ) (Cohen, 1960), utilizada en este trabajo, mide la concordancia entre dos anotadores, considerando la posibilidad de concordancia aleatoria. La misma se calcula de acuerdo con la ecuación:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (3.1)$$

donde:

- P_o es la proporción de acuerdos observados entre dos anotadores, al realizar la anotación en forma independiente. Se determina como el número total de ítems en los que los anotadores coinciden, dividido por el total de los ítems.
- P_e es la proporción de acuerdos esperados por azar. El coeficiente κ asume que cada anotador tiene su propia distribución de probabilidad, por lo que el cálculo se realiza en tres pasos. Primero se calcula la probabilidad de que cada anotador utilice cada categoría, a partir de la frecuencia con que la usó sobre el total de ítems. Luego, para cada categoría, se calcula el acuerdo por azar multiplicando la probabilidad de que el anotador 1 la utilice por la probabilidad de que el anotador 2 utilice esa misma categoría. Por último se suman las probabilidades de azar de todas las categorías (Artstein & Poesio, 2008).

- $P_o - P_e$ representa la proporción de los casos en los que se produjo un acuerdo más allá del azar y constituye el numerador del coeficiente.

Por lo tanto la prueba de acuerdo se plantea con respecto a $1 - P_e$, la proporción de acuerdos para los cuales podría haber desacuerdo entre los anotadores.

El coeficiente κ de Cohen es simplemente la proporción de desacuerdos esperados por azar que no ocurren o, de forma alternativa, la proporción de acuerdo una vez que se ha eliminado el acuerdo debido al azar (Cohen, 1960).

La interpretación de los valores del coeficiente κ según Pustejovsky y Stubbs (2014, p.132-133) depende de la complejidad de la tarea de anotación, por lo que no existe una regla que defina estrictamente cuáles puntuaciones son buenas o no. Por ejemplo, se espera que en las tareas de etiquetado gramatical la puntuación de κ sea cercana a 1.0, en cambio, las tareas que requieren una mayor interpretación del texto no se someten a un estándar tan alto. Landis y Koch (1977) proporcionan una referencia para interpretar κ que se muestra en la Tabla 3.1. Artstein y Poesio (2008) en cambio, identifican una falta de consenso sobre cómo interpretar los valores de los coeficientes de acuerdo en la práctica actual considerando que valores por encima de 0.8 aseguran una calidad de anotación razonable cuando se necesita establecer un umbral, pero tienen dudas de que este punto de corte sea apropiado para todos los propósitos, ya que existen estudios en PLN que han obtenidos corpus útiles con un nivel de 0.7, recomendando a los investigadores siempre informar la metodología utilizada.

Tabla 3.1: Interpretación de los valores de la Kappa de Cohen de acuerdo con Landis y Koch (1977)

Rango de κ	Nivel de concordancia
Menor que 0	Sin concordancia
0,01 – 0,20	Leve
0,21 – 0,40	Regular
0,41 – 0,60	Moderado
0,61 – 0,80	Sustancial
0,81 – 1,00	Perfecto

Una vez definido el corpus, es momento de entrenar y probar los algoritmos. Una forma común de hacerlo es dividir el corpus en tres partes: el *corpus de*

entrenamiento (train corpus), el *corpus de desarrollo (development corpus)* y el *corpus de evaluación (test corpus)*. El *conjunto de entrenamiento* se emplea para entrenar el algoritmo que se utilizará para la tarea. De acuerdo con el diagrama de la [Figura 3.2](#), el *conjunto de desarrollo* se utiliza para el análisis de errores durante la etapa de testeo. Una vez entrenado el algoritmo, este se ejecuta sobre el conjunto de desarrollo y se revisan las fallas que presenta al etiquetar correctamente el corpus, además de ajustarse sus parámetros. Posteriormente, el algoritmo se vuelve a entrenar y a probar nuevamente las veces que sea necesario, hasta obtener resultados que se consideren satisfactorios. Una vez completada la fase de entrenamiento y ajuste, el algoritmo se ejecuta con el corpus de evaluación reservado, que no se ha utilizado hasta el momento ni en el entrenamiento ni en las pruebas de desarrollo (Pustejovsky & Stubbs, 2014, p.29).

Tras aplicar algoritmos de AA, se necesitan herramientas para evaluar su rendimiento. Para determinar estas métricas lo primero que se realiza es una *matriz de confusión* que permite identificar las predicciones del algoritmo que coinciden o no con las etiquetas anotadas por humanos (etiqueta real), tal como se muestra en la [Figura 3.3](#). En el caso de una clasificación binaria, esta matriz distingue entre *verdaderos positivos* (VP), cuando el modelo predice correctamente la clase positiva, *verdaderos negativos* (VN), cuando predice correctamente la clase negativa, *falsos positivos* (FP), cuando asigna erróneamente la clase positiva y *falsos negativos* (FN), cuando no logra identificar un caso que sí pertenece a la clase positiva (Jurafsky & Martin, 2025b).

		Predicción Algoritmo	
		Positivo	Negativo
Etiqueta Real	Positivo	Verdadero Positivo (VP)	Falso Negativo (FN)
	Negativo	Falso Positivo (FP)	Verdadero Negativo (VN)

Figura 3.3: Matriz de confusión con etiquetas reales y predichas.

Una de las métricas utilizadas para evaluar el rendimiento de un algoritmo es determinar la PRECISION (*precisión*) con la que éste etiqueta el conjunto de datos, es decir cuántas de las instancias que han sido clasificadas como positivas

realmente son ejemplos de la clase positiva (Jurafsky & Martin, 2025b). Esta métrica se determina mediante la [Ecuación 3.2](#).

$$Precision = \frac{VP}{VP + FP} \quad (3.2)$$

Otra de las métricas utilizadas es el RECALL (*cobertura*), que indica qué proporción de las etiquetas realmente positivas ha sido correctamente predicha por el algoritmo (Jurafsky & Martin, 2025b). De acuerdo con la [Ecuación 3.3](#), se calcula como la razón entre el número de verdaderos positivos y el número total de etiquetas positivas .

$$Recall = \frac{VP}{VP + FN} \quad (3.3)$$

Los valores de *Precision* y *Recall* suelen competir entre sí. Dado que es deseable que el algoritmo alcance simultáneamente valores elevados en ambas métricas, estas se combinan en una métrica denominada F1-SCORE, definida como la media armónica de ambas y calculada según la [Ecuación 3.4](#) (Jurafsky & Martin, 2025b).

$$F_1\text{-Score} = 2 \times \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}} \quad (3.4)$$

Por último, ACCURACY (*Exactitud*) posiblemente sea la métrica más utilizada y la primera opción a la hora de evaluar el rendimiento de un algoritmo. Se calcula de acuerdo a la [Ecuación 3.5](#) y se define como la proporción de etiquetas correctas sobre el total de las etiquetas posibles, dando una visión de la proporción de acierto para ambas clases (Jurafsky & Martin, 2025b). A pesar de su amplia usabilidad, esta métrica no es la más apropiada en las situaciones en que las clases del corpus están desbalanceadas.

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN} \quad (3.5)$$

3.3. Modelos de Clasificación.

El objetivo de un problema de clasificación es, a partir de una observación (datos de entrada), extraer un conjunto de características que permitan asignarla a una de varias clases discretas. La clasificación puede implementarse de diversas maneras, por ejemplo, mediante reglas definidas manualmente que

permiten etiquetar un texto como perteneciente a una clase específica cuando en él aparece determinada palabra o expresión. La forma más común de hacer clasificación es usar el *AA supervisado*, aunque también pueden utilizarse los *Grandes Modelos de Lenguajes* (LLM, por su sigla en inglés) mediante la técnica de *prompting*, para asignar una clase a un texto. Existen muchos modelos de clasificación en PLN, a continuación se detallan los usados en este trabajo.

El modelo ***Naïve Bayes*** es un modelo probabilístico que se basa en el teorema de Bayes. Este aprende una probabilidad de que cada texto pertenezca a una clase y luego elige como clasificación a la clase con la probabilidad mayor. Esta forma de clasificación se basa en la regla de Bayes, que permite estimar la probabilidad de que un enunciado x pertenezca a una clase C :

$$P(C \mid x) = \frac{P(x \mid C) P(C)}{P(x)}, \quad (3.6)$$

donde $P(C \mid x)$ representa, en el contexto de la clasificación, la probabilidad de que un enunciado x pertenezca a la clase C . Por su parte, $P(x \mid C)$ corresponde a la probabilidad de observar ese enunciado si efectivamente perteneciera a dicha clase. $P(C)$ indica la probabilidad previa de la clase, mientras que $P(x)$ expresa la probabilidad total de observar el enunciado en el corpus (Jurafsky & Martin, 2026c).

Support Vector Machine (SVM) es un algoritmo de decisión, no proporciona probabilidades. El algoritmo SVM determina un hiperplano que separa de la mejor manera posible las muestras de datos. Este hiperplano es denominado *frontera de decisión*. Todas las muestras que caen de un lado de la frontera pertenecen a la clase A y las del otro lado, a la clase B (Bishop, 2009).

Regresión logística (LR, por su sigla en inglés) es, en su forma más simple, un modelo de clasificación binaria (aunque puede generalizarse al caso *multinomial*). A partir de una combinación lineal de las características de entrada y mediante una función de activación (habitualmente una función sigmoide), estima la probabilidad de que una observación pertenezca a una de las dos clases posibles. Luego, se define un umbral de decisión sobre dicha probabilidad para asignar la clase correspondiente, por ejemplo, si la probabilidad es mayor que 0.5 se asigna la clase B y, en caso contrario, la clase A (Jurafsky & Martin, 2025b).

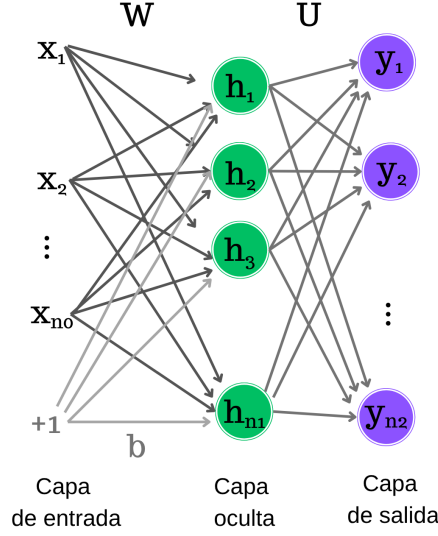


Figura 3.4: Diagrama adaptado de Jurafsky y Martin (2025c) de una red neuronal *feedforward* de dos capas, una oculta y una de salida.

Las **redes neuronales** son modelos computacionales inspirados en el cerebro humano y se utilizan para reconocer patrones y hacer predicciones. Estas redes consisten en capas de nodos o neuronas interconectadas, que procesan y transforman la información. Si bien comparten gran parte de la misma base matemática que la regresión logística, son clasificadores más potentes, y una red neuronal mínima de una sola capa puede aprender cualquier función (Gallardo Fernández et al., 2024; Jurafsky & Martin, 2025c).

Las redes multicapa de propagación directa, conocidas como redes *feedforward* o **Multi-layer Perceptrons** (MLP), se caracterizan porque la información fluye desde la capa de entrada hacia las capas sucesivas hasta la capa de salida, de modo que las salidas de cada capa se transmiten únicamente a las unidades de la capa siguiente, sin conexiones de retroalimentación hacia capas anteriores. Las redes *feedforward* simples tienen tres tipos de nodos: unidades de entrada, unidades ocultas y unidades de salida como se muestra en el esquema de la Figura 3.4. La capa de entrada $\mathbf{x}$ es un vector de valores escalares. El núcleo de la red neuronal es la capa oculta $\mathbf{h}$, formada por h_i unidades. Cada una de ellas es una unidad neuronal que toma la suma ponderada de sus entradas y luego aplica una función no lineal. En la arquitectura estándar cada capa está totalmente conectada, es decir, que cada unidad toma como entrada todas las salidas de las unidades de la capa anterior. La operación que realiza

cada capa oculta consiste en multiplicar la matriz de pesos $\mathbf{W}$ por el vector de entrada $\mathbf{x}$, sumar el vector de sesgos $\mathbf{b}$ y aplicar una función de activación. Luego, la capa de salida también tiene una matriz de pesos $\mathbf{U}$ que se multiplica por su vector de entrada $\mathbf{h}$. Sin embargo, el resultado de esta multiplicación no puede constituir directamente la salida del clasificador, ya que es un vector de números reales, cuando lo que se necesita es un vector de probabilidades. Por ello, se aplica una función de normalización que convierte ese vector en una distribución de probabilidad. (Jurafsky & Martin, 2025c).

Además de los modelos que construyen una función de decisión durante la etapa de entrenamiento, el algoritmo de los k -vecinos más cercanos (***k-Nearest Neighbors***, *k-NN*) constituye uno de los métodos más representativos de los enfoques de clasificación supervisada basados en la similitud entre ejemplos. Este método pertenece al aprendizaje basado en instancias, en el que no se construye previamente una función general de clasificación a partir de los datos de entrenamiento. En cambio, ante una nueva consulta, el algoritmo compara la instancia con los ejemplos almacenados y realiza la predicción según su grado de similitud con ellos. Para ello, representa cada instancia como un punto en un espacio de atributos e identifica los k ejemplos de entrenamiento más cercanos. En problemas de clasificación, la clase asignada corresponde a la más frecuente entre esos vecinos (Mitchell, 1997).

3.3.1. Representación de palabras y texto.

En PLN se trabaja principalmente con texto. La mayoría de los clasificadores utilizan valores numéricos como datos de entrada, por lo que es necesario una representación numérica de los textos. Para esta representación de texto se utiliza un vector de atributos.

Una de las técnicas más básicas es la denominada *Bag of Words* (BoW, bolsa de palabras), que consiste en construir un vocabulario con las palabras que aparecen en el texto proporcionado. Luego, cada texto se representa mediante un vector de atributos de dimensión igual al tamaño del vocabulario, en el que cada componente indica la presencia o frecuencia de una palabra (Jurafsky & Martin, 2026c). En este proceso, es posible aplicar algunas variantes, como la eliminación de *stop words* (palabras vacías), la *lematización* o *TF-IDF*. Eliminar las palabras vacías consiste en suprimir algunas palabras extremadamente comunes, con poco valor para tareas de clasificación. El uso de una lista de

palabras vacías reduce la cantidad de entradas que un sistema debe almacenar (Manning et al., 2008, p. 25). Por otra parte, la lematización, tiene por objetivo eliminar únicamente las terminaciones flexivas y recuperar la forma base (lema o forma de diccionario) de la palabra, considerando su uso, por ejemplo, como verbo o como sustantivo (Manning et al., 2008, p. 31). En cuanto a TF-IDF, la frecuencia de término (*TF*) asigna mayor peso a las palabras que aparecen más veces en un documento, bajo la suposición de que están más relacionadas con su contenido. A su vez, la frecuencia inversa de documento (*IDF*) reduce el peso de las palabras que aparecen en muchos documentos del corpus. La combinación de ambas medidas da lugar a *TF-IDF*, que otorga mayor peso a los términos frecuentes en un documento pero poco comunes en el conjunto de documentos, ya que son los que mejor permiten discriminar su contenido (Manning et al., 2008, p. 109).

Otra forma de representar el texto consiste en partir de representaciones vectoriales de las palabras, lo que da lugar a métodos como el *one-hot encoding* y los *word embeddings*. En el método *one-hot encoding* se utiliza un vector del tamaño del vocabulario, que vale cero para todas las dimensiones excepto para la palabra que se quiere representar. Por otra parte, los *word embeddings* (vectores de palabra) capturan los datos semánticos y contextuales de las palabras en un espacio vectorial de múltiples dimensiones que le permite a las máquinas comprender las palabras en función de sus relaciones con otras palabras. En esta técnica, cada dimensión del vector representa un aspecto o característica diferente de la palabra y se espera que palabras con significados o usos similares tiendan a tener vectores que están cerca unos de otros en el espacio vectorial. Los *word embeddings* permiten medir la similitud o distancia entre palabras basándose en las propiedades regulares del espacio vectorial, lo que permite identificar palabras semánticamente similares y son útiles en PLN porque permiten a las máquinas comprender las palabras en su contexto. Existen diferencias a señalar entre los *word embeddings* y las representaciones de tipo *one-hot*. Mientras que los *word embeddings* corresponden a vectores densos, compuestos por valores reales diferentes de cero, las representaciones *one-hot* son vectores dispersos, en los que casi todos sus componentes son cero. Además, los *word embeddings* constituyen representaciones compactas, ya que suelen estar formados por algunos cientos de dimensiones, mientras que los vectores *one-hot* tienen una dimensionalidad igual al tamaño del vocabulario, que habitualmente es del orden de decenas de miles. Entre los métodos más

conocidos para la construcción de *embeddings* se encuentra *word2vec*, en el que cada palabra se representa mediante los pesos que el modelo aprende durante el entrenamiento. Estos pesos se ajustan a partir de la predicción de relaciones de coocurrencia entre palabras en grandes corpus de texto sin anotar. En este enfoque auto-supervisado, la propia distribución de las palabras en el corpus se utiliza durante el aprendizaje del modelo, y los vectores resultantes reflejan los contextos en los que las palabras suelen aparecer (Jurafsky & Martin, 2026a).

El surgimiento de la arquitectura *Transformer*, que se abordará en la siguiente subsección, junto con el desarrollo de los LLM, dio lugar a nuevas formas de representación vectorial del lenguaje. Entre ellas se encuentran los *embeddings contextuales*, en los que una misma palabra puede representarse mediante vectores diferentes según el contexto en el que aparece, y los *sentence embeddings*, que consisten en representaciones vectoriales de fragmentos de texto, como frases, oraciones o párrafos. Los *sentence embeddings* permiten capturar información semántica y contextual, lo que resulta de utilidad en tareas como la medición de similitud entre oraciones, el análisis de sentimientos y la clasificación de texto (Cheng et al., 2025).

3.3.2. Modelos de lenguaje neuronales y clasificación de textos.

Los modelos de lenguaje son modelos que asignan una probabilidad a las posibles palabras siguientes a una secuencia de texto (Jurafsky & Martin, 2026b). En particular, los LLM basados en la arquitectura *Transformer* se caracterizan por contar con una gran cantidad de parámetros y por ser entrenados sobre grandes corpus de texto. Estas características les permiten alcanzar un alto desempeño en diversas tareas de comprensión de textos (Alaniz Ferreira et al., 2025).

En el campo del PLN, los LLM como Gemini (Gemini Team, 2025; Google DeepMind, 2025), Llama (Touvron et al., 2023) o GPT (OpenAI et al., 2025; Radford et al., 2018) pueden emplearse en tareas de clasificación de texto. Si bien estos modelos se entrenan originalmente para predecir la siguiente palabra a partir del contexto, sus capacidades resultan aplicables a una amplia variedad de tareas, entre ellas la clasificación de texto. Además, su capacidad para generar texto a partir de una entrada los convierte en la base de lo que se conoce como *IA generativa* (Jurafsky & Martin, 2025a).

Los modelos de lenguaje ajustados mediante instrucciones (*prompts*) muestran un buen desempeño al seguir indicaciones, responder preguntas y sostener conversaciones. Un PROMPT (indicación o instrucción) es una cadena de texto que un usuario introduce en un modelo de lenguaje para conseguir la respuesta o tarea que se espera que realice. El proceso de encontrar *prompts* efectivos para una tarea se conoce como ingeniería de *prompts*. Los *prompts* pueden mejorarse especificando las respuestas permitidas, especificándose el rol del modelo de lenguaje como un asistente para determinada tarea e incluyendo ejemplos. La tarea de usar *prompts* con pocos ejemplos suele denominarse *few-shot prompting*, cuando el *prompt* no tiene ejemplos se denomina *zero-shot prompting* (Jurafsky & Martin, 2025a). Los ejemplos utilizados suelen obtenerse a partir del corpus de entrenamiento y pueden seleccionarse manualmente o mediante procedimientos automáticos, con el objetivo de mejorar el rendimiento del *prompt*. Por ejemplo, inspirándose en el método k-NN, es frecuente seleccionar ejemplos cercanos al enunciado que se busca clasificar dentro de un conjunto previamente anotado.

La *tokenización* es un tipo de segmentación de documentos. Los documentos pueden segmentarse en párrafos, enunciados, oraciones, frases y, en un nivel de análisis más fino, en *tokens* (palabras, unidades menores que la palabra y signos de puntuación). Al conjunto de *tokens* presentes en un corpus se le denomina *vocabulario* (Lane & Dyshel, 2025, p. 44-45). Esta segmentación resulta fundamental en los modelos actuales de procesamiento del lenguaje, ya que los *tokens* constituyen la unidad básica sobre la que operan arquitecturas como *Transformer*, que sirven de base para los LLM.

A partir del procesamiento de *tokens*, la arquitectura *Transformer* se apoya en mecanismos de autoatención (*self-attention*) que permiten construir representaciones contextualizadas de los *tokens* de una secuencia de entrada. En este proceso, la representación de cada *token* se obtiene mediante la integración de información proveniente de los demás *tokens* del contexto. De este modo, los modelos basados en *Transformer* generan representaciones contextuales del significado de las palabras, conocidas como *contextual embeddings*, fundamentales para tareas de comprensión y generación de texto (Alaniz Ferreira et al., 2025; Jurafsky & Martin, 2025a).

3.4. Antecedentes sobre el uso de IA en la enseñanza de la física experimental.

Los avances recientes en el desarrollo de la IA, junto con las técnicas de PLN abordadas en este capítulo para la clasificación automática de textos, se han incorporado progresivamente en la investigación en educación en física, y en particular en el campo de la enseñanza de la física experimental.

Al examinar las investigaciones que utilizan IA en cursos de física experimental, se advierte que estas se han desarrollado desde perspectivas diversas. Existen trabajos que utilizaron IA para el desarrollo de un Chatbot específico de física (Lazarte et al., 2024), y otras que buscan el desarrollo de robots de aprendizaje interactivo que incluye un chatbot para brindar apoyo educativo (Latif et al., 2024). También se encontraron varias investigaciones relacionadas al uso de la IA como el ChatGPT para facilitar las etapas de indagación en física como formulación de preguntas e hipótesis, diseño experimental, análisis de datos, identificación de errores, brindar retroalimentación, entre otros (Fadillah et al., 2024; Kilde-Westberg et al., 2025; Kotsis, 2025; Mills et al., 2025; Samuelsson et al., 2025). Otras investigaciones se relacionan con la *clasificación de respuestas o texto* generado por los estudiantes, como lo es el trabajo de Campbell et al. (2024) que utiliza el PLN, específicamente el sistema Watson de IBM, para clasificar automáticamente 479 respuestas escritas cortas de estudiantes relacionadas con ciertas conclusiones de un experimento hipotético. El trabajo de Fussell et al. (2022) investiga el efecto del tamaño del conjunto de datos de entrenamiento en la clasificación automática mediante redes neuronales. Los resultados indican que se necesita un conjunto de alrededor de 600 respuestas en total, con aproximadamente entre 80 y 100 respuestas positivas por cada código, para lograr un acuerdo sustancial con los evaluadores humanos. Esto significa que el clasificador automático alcanza un nivel de acuerdo comparable al de un evaluador humano. En otro trabajo, Bralin et al. (2023) utilizaron 683 ensayos escritos por estudiantes dada una tarea de laboratorio que consistía en encontrar un problema del mundo real relacionado a los contenidos del curso y debían describir cómo resolverlo. Utilizaron dos algoritmos de AA para identificar temas o patrones en los textos, lo que les permitió reconocer tópicos recurrentes antes de una lectura detallada, ahorrar tiempo y familiarizarse previamente con los trabajos. Un trabajo más actual es el de Fussell et al. (2025), en este se comparan diferentes modelos de lenguaje

para el análisis de texto buscando detectar habilidades experimentales específicas en 873 archivos de notas de laboratorio (digitales y en inglés) mediante el etiquetado a nivel de oración, incorporando a su trabajo el análisis de la incertidumbre estadística y sistemática.

En Uruguay existe un importante camino recorrido en investigaciones sobre la enseñanza de la física experimental. Existen múltiples trabajos que proponen experiencias que integran el uso de smartphones (Monteiro et al., 2019; Monteiro et al., 2014; Monteiro et al., 2023) y el diseño de experimentos de bajo costo con TIC (Calderón et al., 2015; Di Laccio, 2024). Estas propuestas surgen como respuesta a algunos desafíos que enfrenta la enseñanza y el aprendizaje de la física en el país, tales como la enseñanza a través de “recetas de cocina”, escasa actualización de propuestas experimentales, espacios de trabajo poco inclusivos, la primacía de la formación teórica sobre la experimental y las interpretaciones rígidas de los programas de los cursos (Di Laccio, 2017). En cambio, la investigación sobre el uso de IA en la enseñanza de la física, este es aún incipiente. Sin embargo, comienzan a surgir algunos aportes, como el de Suárez y Martí (2024) que explora el uso de ChatGPT como herramienta para generar diferentes tipos de problemas que pueden utilizarse en el aula junto con rúbricas para su evaluación. También se encuentra el trabajo de de Abreu et al. (2026) que explora el potencial de ChatGPT como asistente para la evaluación de informes de laboratorio en física experimental, proponiendo su uso como apoyo al proceso evaluativo, siempre bajo la supervisión y validación del docente. Además, existen otros trabajos vinculados al uso de IA, aunque no centrados específicamente en el trabajo de laboratorio. Entre ellos se encuentra el trabajo de Ballesta y Sguilla (2025), que indaga las percepciones sobre la IA en ámbitos educativos de enseñanza media en Uruguay. Por otra parte, el trabajo de Corrales y Meneses (2025) analiza las distintas interpretaciones de la primera ley de Newton presentes en textos clásicos de física utilizados en la formación docente e incorpora IA como herramienta para organizar y clasificar dichas interpretaciones.

3.5. Síntesis del capítulo.

Los elementos desarrollados en este capítulo ofrecen el marco conceptual y metodológico necesario para abordar el análisis automático de texto. La revisión realizada permitió presentar al PLN como un área de la IA que busca

resolver tareas vinculadas al lenguaje humano, las formas de representación de texto, los métodos de clasificación de textos abordados en este trabajo y las métricas empleadas para evaluar el desempeño de los modelos. La aplicación de estas herramientas exige la construcción de un corpus anotado que refleje las categorías teóricas definidas para el análisis del problema. Por ello, el capítulo siguiente se dedica a describir el proceso de construcción del corpus anotado.

Capítulo 4

Diseño de la investigación y construcción del corpus.

El diseño metodológico de la presente investigación se centra en la selección, el ajuste, la validación y la evaluación de algoritmos de AA capaces de automatizar el análisis de textos provenientes de informes de laboratorio elaborados por estudiantes, con el fin de identificar las habilidades científicas que se evidencian en ellos. Estas habilidades se contrastan con los objetivos formulados en el programa del curso de Física Experimental 1 y en las guías de trabajo proporcionadas a los estudiantes. Además, los resultados se comparan con las percepciones de docentes y estudiantes respecto al curso y al desarrollo de estas habilidades. La metodología de la investigación se organiza entonces en cuatro fases:

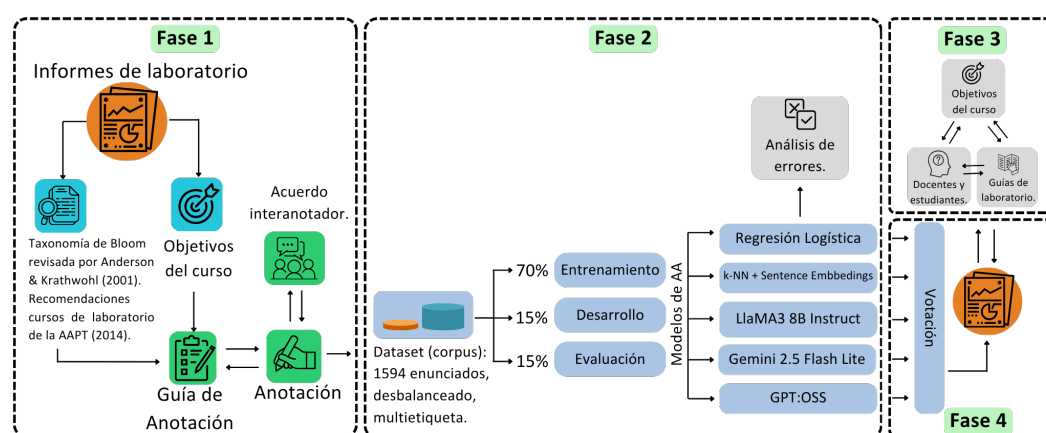


Figura 4.1: Representación gráfica de la metodología de investigación. Fuente: elaboración propia.

1. Definición del conjunto de clases para la clasificación, junto con la construcción y validación de una guía de anotación fundamentada en la literatura especializada, a partir de la cual se elaboró el corpus anotado.
2. Entrenamiento y evaluación de algoritmos de AA para la clasificación de textos, junto con el correspondiente análisis de errores.
3. Análisis de los objetivos del curso, de las guías de laboratorio y de las percepciones de docentes y estudiantes.
4. Análisis de los resultados obtenidos mediante la clasificación automática de los informes del semestre, en relación con el desarrollo de habilidades científicas.

De acuerdo con estas etapas y con el método de investigación que se muestra en la [Figura 4.1](#), se comenzó seleccionando para el análisis los informes de laboratorio correspondientes al primer semestre de 2024 de la unidad curricular Física Experimental 1 de la Facultad de Ingeniería de la Udelar. Una vez obtenido el acceso a estos informes, se procedió a su anonimización, con el fin de resguardar la identidad de los involucrados. A partir de este conjunto, se definió una muestra de informes que fue analizada por profesores de física con el propósito de desarrollar la guía de anotación y validar el proceso.

En esta primera etapa, se elaboró una guía de anotación que permitió clasificar los enunciados en cinco clases. Estas clases fueron definidas en función de las habilidades científicas que se espera promover en los cursos introductorios de laboratorio de física, tomando como referencia las recomendaciones de la AAPT y enmarcándolas en la Taxonomía Revisada de Bloom, particularmente en relación con los distintos procesos cognitivos involucrados.

Para realizar este análisis, los informes fueron segmentados manualmente en enunciados. La segmentación se efectuó sobre las secciones *Metodología*, *Análisis de resultados* y *Conclusiones* de los informes de laboratorio, considerando unidades con sentido dentro del texto y excluyendo aquellos fragmentos constituidos por desarrollo de cálculos, resultados numéricos, figuras, pies de figura o notas al pie. Durante el proceso de anotación, los enunciados fueron leídos por los anotadores respetando el orden en que aparecían en cada informe, y cada anotador podía asignar a cada uno de ellos una, varias o ninguna de las clases definidas.

Inicialmente el trabajo se llevó a cabo con la participación de dos anotado-

res, ambos profesores de física, quienes colaboraron en el diseño y ajuste de la guía de anotación. Este proceso buscó construir un instrumento claro, con el menor grado posible de ambigüedad, que contribuyera a reducir discrepancias en la aplicación de los criterios de clasificación.

Una vez alcanzada una versión satisfactoria de la guía, el investigador y uno de los anotadores expertos realizaron la anotación individual de enunciados extraídos de una muestra de informes. A partir de estas anotaciones, se evaluó el grado de acuerdo interevaluador (AIC). Cuando los resultados obtenidos fueron considerados satisfactorios para los propósitos del estudio, se avanzó hacia la etapa de ajuste y evaluación de los algoritmos de clasificación automática de textos, realizando las pruebas correspondientes sobre el corpus de desarrollo.

Una vez definidos los cinco algoritmos utilizados para la clasificación de texto, se evaluó el desempeño de cada uno sobre el corpus de evaluación. A partir de las predicciones obtenidas, se realizó un análisis de errores para cada modelo individual. Posteriormente, las predicciones generadas por los cinco modelos se integraron mediante un esquema de votación por mayoría simple, a partir del cual se obtuvo la clasificación final de cada enunciado.

Finalmente, los algoritmos de clasificación automática se utilizaron sobre el conjunto completo de informes correspondientes a las prácticas del semestre. A partir de ello, se seleccionaron dos clases con el propósito de analizar su comportamiento a lo largo del curso y contrastar los resultados obtenidos con los objetivos del curso, las características de las guías de laboratorio y las percepciones de docentes y estudiantes.

4.1. Modelado y construcción de la guía de anotación.

El desarrollo de la guía de anotación se realizó con base en la Taxonomía de Bloom revisada por Anderson y Krathwohl (2001) (TBA) desarrollada en la [Subsección 2.1.2](#) y en las recomendaciones para los cursos de laboratorio de la AAPT (Kozminski et al., 2014).

La TBA ha sido utilizada en la enseñanza de la física en diversos estudios. Por ejemplo, en el trabajo de Bates et al. (2014) se emplean los procesos cognitivos de la TBA para categorizar y evaluar la calidad de las preguntas

elaboradas por estudiantes de cursos introductorios de física. Esta taxonomía, también ha sido utilizada por Damayanti et al. (2020) y Silva Dias et al. (2021) para identificar los objetivos educativos priorizados en exámenes de física. Otros trabajos, como los de Baeza et al. (2023) y Buick (2011), toman la TBA como punto de partida para diseñar una taxonomía específica para la enseñanza y evaluación en física. En Uruguay, Olivera (2023) ha realizado un análisis preliminar cualitativo de informes de laboratorio basado en categorías de la TBA. En ese trabajo se examina la relación entre los niveles de la taxonomía presentes en las guías de laboratorio y en los informes estudiantiles a partir de la búsqueda, mediante RStudio, de palabras asociadas a las distintas categorías.

A partir de estos antecedentes, en este trabajo se llevó a cabo una primera aproximación metodológica sobre informes del año 2023. Esta etapa consistió en identificar los distintos niveles de los procesos cognitivos de la TBA mediante la búsqueda de palabras, principalmente verbos, en los textos escritos por los estudiantes, asociándolas con las categorías correspondientes. Si bien se tomó como punto de partida el conjunto de verbos que la taxonomía propone para cada nivel, la lectura de los informes de laboratorio permitió identificar el vocabulario efectivamente utilizado por el estudiantado. A partir de ello, las categorías fueron ampliadas mediante la incorporación de términos más específicos, tal como se presenta en la [Tabla A.1](#), incluida en el [Apéndice A](#). Sin embargo, la búsqueda de palabras y el análisis de su frecuencia no lograban capturar el significado de los enunciados. Se observó que, en algunos casos, ciertos términos se utilizaban como sinónimos o para aludir a acciones que se realizarían posteriormente, sin que ello implicara necesariamente el logro del proceso cognitivo correspondiente. Por ejemplo, se identificaron enunciados como el siguiente:

“Para analizarlo mejor calculamos el promedio y la desviación estándar, según las ecuaciones:”

Aunque en este enunciado se utiliza la palabra *analizar*, su contenido no da cuenta de un proceso de análisis propiamente dicho.

Para captar de mejor manera los procesos cognitivos desarrollados por los estudiantes, se optó por etiquetar segmentos de texto en los que se evidenciara alguno de los niveles de la TBA, considerando no solo palabras aisladas, sino conjuntos de palabras con sentido dentro del enunciado, como por ejemplo:

*“A diferencia de las mediciones repetidas, no fue posible tomar 50 mediciones consecutivas ya que el péndulo se desvió y empezó a colisionar con la fotopuerta, **pero creemos que** suficientes medidas para lograr un análisis **satisfactorio** fueron tomadas.”*

En este segmento se utiliza “*pero creemos que*” para realizar una valoración “*satisfactoria*” del análisis realizado, en otras palabras, evaluar el procedimiento utilizado para la toma de medidas. Algunas secuencias de palabras identificadas y asociadas a las diferentes categorías de la TBA pueden consultarse en el [Apéndice B](#).

La metodología basada en el conteo de frecuencias de términos, ya sean palabras individuales o secuencias, abrió paso al cuestionamiento de la validez de una relación directa entre la frecuencia de aparición y la categoría asignada. La co-ocurrencia de múltiples términos pertenecientes a una misma categoría en un mismo enunciado puede distorsionar la ponderación de dicha categoría, llevando a una sobreestimación de su relevancia en el texto analizado. A modo de ejemplo, el siguiente enunciado que debería ser incluido dentro de la categoría *Aplicar* incluye tres términos (marcados en negrita) asociados a la categoría.

*“**Utilizando** nuevamente las ecuaciones (10) y (11) **calcularemos** el **promedio** y la desviación estándar, respectivamente, de las mediciones sin datos atípicos.”*

Además de esta limitación, se constató que una misma palabra podía ser utilizada con propósitos semánticos distintos dependiendo del contexto. Por ejemplo, el término *elegir* podría asociarse a la categoría *Aplicar* en cierto contexto (ej. elegir cierto instrumento de medida), pero también podría ser vinculado a la categoría *Evaluar* si se selecciona un instrumento y se realiza una valoración de las características por las que fue elegido. Por estas razones se determinó necesario considerar el significado del enunciado completo, adoptando métodos que contemplen el contexto en el que se utilizan las palabras. Para ello primeramente se retomaron las definiciones de los diferentes procesos cognitivos de la TBA, tal como se muestran en la [Tabla 4.1](#), a partir de las cuales se construyó una descripción de los procesos cognitivos aplicados a Física Experimental. Este nuevo esquema de clasificación puede consultarse en el [Apéndice C](#).

Dimensiones del proceso cognitivo de la TBA		
Recordar Recupera conocimientos relevantes de la memoria a largo plazo		
Reconocer	Identificando	Localiza conocimiento en la memoria a largo plazo que sea consistente con el material presentado.
Recordar	Recuperando	Recupera conocimientos importantes de la memoria a largo plazo.
Comprender Construye significado a partir de instrucciones, incluida la comunicación, oral, escrita o gráfica.		
Interpretar	Aclarando, parafraseando, representando, traduciendo	Cambiar de una forma de representación a otra.
Ejemplificar	Ilustrar, dar un ejemplo de algo	Encontrar un ejemplo o ilustración específica de un concepto o principio.
Clasificar	Categorizar, incluir algo dentro de un grupo más grande	Reconoce que algo pertenece a cierta categoría. Implica detectar características o patrones relevantes que "se ajustan" a un caso específico, concepto o principio.
Resumir	Abstraer, generalizar	Resumir un tema o seleccionar puntos importantes.
Inferir	Concluyendo, prediciendo, extrapolando	Sacar una conclusión lógica a partir de la información presentada.
Comparar	Contrastar, mapear, emparejar	Detectar correspondencias entre dos ideas, objetos y similares. Detectar similitudes y diferencias con un conocimiento familiar.
Explicar	Construyendo modelos.	Construir un modelo de causa y efecto de un sistema.
Aplicar Realizar o usar un procedimiento en una determinada situación		
Ejecutar	Llevando a cabo	Aplicar un procedimiento a una tarea familiar, ej. realizar operaciones matemáticas.
Implementar	Usando	Aplicar un procedimiento a una tarea desconocida, ej. usar la 2da Ley de Newton en situaciones en las que sea apropiado.
Analizar Dividir el material en sus partes constituyentes y determinar cómo se relacionan las partes entre sí y con una estructura o propósito general.		
Diferenciar	Discriminando, distinguiendo, seleccionando, enfocar	Distinguir las partes relevantes de las irrelevantes o distinguir las partes importantes del material presentado.
Organizar	búsqueda y hallazgos, coherencia, integrando, esquematizando, análisis sintáctico, estructurando	Determinar cuáles elementos encajan, ajustan o funcionan dentro de una estructura.
Atribuir	Deconstruyendo	Determinar un punto de vista, sesgo, valores o intención subyacente en el material presentado.
Evaluar Emitir juicios basados en criterios y estándares		
Comprobando	Coordinando, detectando, monitoreando, realizando pruebas	Detectando inconsistencias o falacias dentro de un proceso o producto; determinar si un producto tiene consistencia interna; detectar la efectividad de un procedimiento a medida que se implementa. Ej. determinar si las conclusiones de un científico se derivan de los datos observados.
Criticando	Juzgar	Detectar inconsistencias entre un producto y criterios externos, determinando si un producto tiene consistencia externa. Detectar lo apropiado de un procedimiento para un problema dado, por ej. cuál de los métodos es mejor para resolver un problema dado.
Crear Juntar elementos para formar un todo coherente o funcional, reorganizar elementos en un nuevo patrón o estructura.		
Generar	Formular hipótesis	Proponer hipótesis alternativas basadas en un criterio, por ejemplo: generar hipótesis para dar cuenta de un fenómeno observado
Planificar	Diseñar	Diseñar un procedimiento para realizar alguna tarea, por ejemplo: planificar un trabajo de investigación de un fenómeno o tema.
Producir	Construir	Inventar un producto

Tabla 4.1: Descripción de las dimensiones del proceso cognitivo de la TBA.

Este modelo teórico de anotación requirió, además, revisar la *Guía para la elaboración de informes de laboratorio* proporcionada a los estudiantes, así como las guías de trabajo (también denominadas protocolos) entregadas para la realización de cada una de las cuatro prácticas experimentales desarrolladas durante el semestre (véase el Anexo 2).

El uso exclusivo de la TBA para la anotación permitía identificar los diferentes niveles de procesos cognitivos alcanzados por los estudiantes, pero no posibilitaba determinar qué habilidades científicas desarrollaban. Por este motivo, se buscó mejorar el modelo que sustenta la anotación integrando las áreas clave para los cursos de laboratorio recomendadas por la AAPT (Kozminski et al., 2014) con la TBA (Anderson & Krathwohl, 2001). Para establecer este vínculo se procedió a desglosar cada una de las seis áreas clave indicadas por la AAPT en habilidades esperadas en los cursos de laboratorio de física introductorios (véase Apéndice D). Luego, cada una de estas habilidades se reagrupó según los procesos cognitivos de la TBA.

En la Figura 4.2 se observa el desglose de las habilidades que se espera que

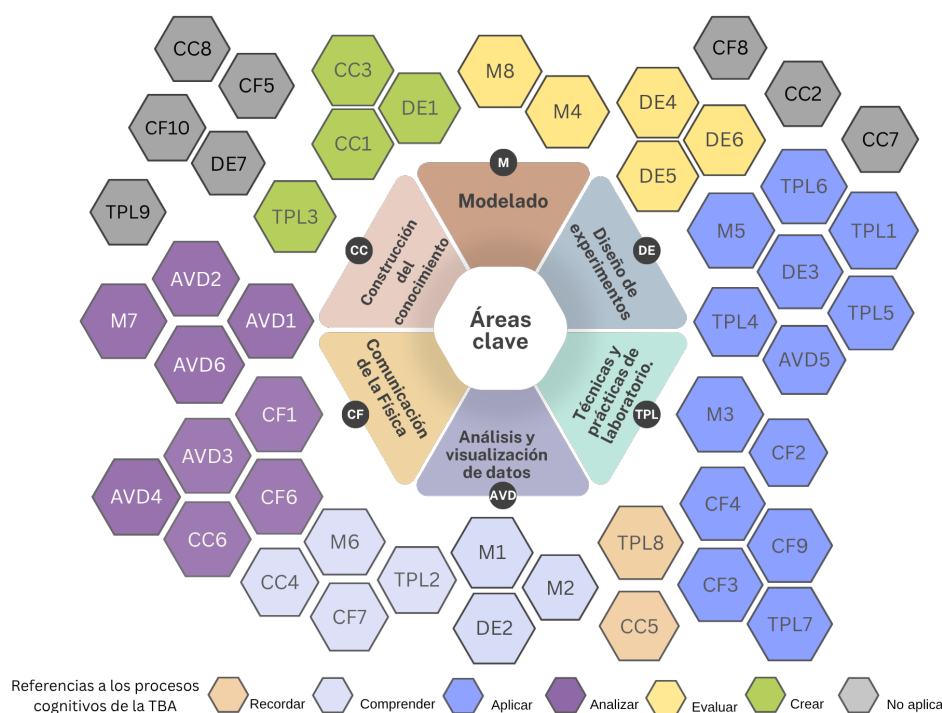


Figura 4.2: Diagrama con la agrupación de las habilidades esperadas (véase Apéndice D) pertenecientes a cada una de las áreas de enfoque para los cursos de laboratorio recomendadas por la Kozminski et al. (2014). Se agrupan de acuerdo con los procesos cognitivos de la TBA y áreas clave a la que pertenece.

el estudiante logre desarrollar de acuerdo con las áreas clave indicadas por la AAPT. Estas se han reagrupado considerando los seis niveles de procesos cognitivos de la TBA. Además, se agruparon aquellas habilidades que se consideró que no correspondían a este estudio, dadas sus características y el diseño del curso. Este nuevo ordenamiento permitió elaborar una nueva descripción de los niveles de la TBA, basada en las habilidades que se procura desarrollar en los cursos de física experimental. La [Tabla 4.2](#) muestra una versión preliminar de la guía de anotación suministrada a los anotadores para su evaluación. En dicha tabla se presenta la descripción de cada habilidad, junto con el código empleado para su identificación en la [Figura 4.2.](#)

Tabla 4.2: Guía de anotación que integra las habilidades científicas de la AAPT y las ordena según los niveles de los procesos cognitivos de la TBA.

Categoría	Código	Descripción
Recordar	CC5	Identifica los conceptos de Física más importantes de un experimento.
	TPL8	Realiza una descripción del aparato y procedimiento de medición.
Comprender	M1	Comprende cualitativa y cuantitativamente el sistema a modelar.
	M2	Selecciona un marco conceptual adecuado para la situación que se está modelando.
	DE2	Realiza una revisión bibliográfica del modelo científico para mejorar el diseño experimental.
	TPL2	Identifica las limitaciones e incertidumbres asociadas a los dispositivos de medición.
	M6	Identifica los errores sistemáticos o sesgos producidos por una investigación experimental.
	CC4	Describe las observaciones de manera clara y concisa.
	CF7	Interpreta componentes de comunicación propios de la Física como vocabulario, tablas, figuras, resultados numéricos y gráficos.

Continúa en la página siguiente

Tabla 4.2 – continuación

Categoría	Código	Descripción
Aplicar	M5	Calibra instrumentos para su uso.
	DE3	Construye y/o monta el dispositivo experimental.
	TPL1	Utiliza dispositivos de medición y realiza medidas coherentes.
	TPL5	Desarrolla habilidades prácticas manuales como nivelar, organizar, alinear, seleccionar pesos y tensiones adecuados, teniendo en cuenta el impacto en la recolección de datos.
	TPL6	Utiliza herramientas comunes como software de recolección de datos.
	M3	Utiliza múltiples representaciones de modelos: descripciones verbales, escritas, modelos matemáticos, computacionales, gráficos o diagramas.
	CF2	Utiliza varios elementos estándar de comunicación técnica como gráficos, tablas y resultados numéricos con incertidumbres y diagramas.
	CF3	Utiliza vocabulario técnico apropiado para el contenido, los materiales e instrumentos.
	CF4	Utiliza métodos de escritura y presentación científica.
	CF9	Usa su cuaderno de laboratorio como un registro para explicar detalles de su trabajo.
	TPL7	Registra observaciones llevando un cuaderno de laboratorio.
	TPL4	Cuantifica la incertidumbre de dispositivos analógicos o digitales.
	AVD5	Cuantifica las incertidumbres en los resultados y realiza la propagación de incertidumbres en cálculos.
Analizar	AVD3	Extrae información de los gráficos interpretando críticamente la validez y las limitaciones de los datos presentados.
	AVD4	Representa información de forma adecuada en un gráfico y extrae información realizando los ajustes correspondientes.
	CC6	Elabora argumentos e identifica tendencias basadas en datos experimentales.
	CF1	Presenta argumentos respaldados por evidencia experimental.
	CF6	Realiza conexiones lógicas entre contexto teórico, evidencia experimental y sus argumentos.
	AVD1	Utiliza métodos estadísticos para analizar datos.
	AVD2	Selecciona métodos de representación de datos de forma adecuada para extraer cantidades físicas de los parámetros de ajuste.
	AVD6	Compara sus resultados con modelos matemáticos, computaciones o simulaciones.
	M7	Utiliza modelos para predecir resultados.
Evaluar	DE4	Evalúa, realiza o sugiere mejoras para el dispositivo experimental.
	DE5	Reconoce las limitaciones del diseño como calidad, costo, tiempo, fuentes de incertidumbre y equipo disponible.
	DE6	Diagnostica y soluciona problemas básicos durante el montaje experimental.
	M4	Reconoce limitaciones y suposiciones del modelo, así como la incertidumbre debida a esto.
	M8	Evalúa si determinada cantidad física es razonable dado el sistema que se está modelando.

Continúa en la página siguiente

Tabla 4.2 – continuación

Categoría	Código	Descripción
Crear	TPL3	Elegir un dispositivo experimental apropiado para realizar la medición.
	CC1	Generar preguntas científicas y experimentos adecuados para responder a las mismas.
	CC3	Idear modelos o hipótesis falsables para explicar características observables en la naturaleza.
	DE1	Diseñar un procedimiento para probar un modelo o hipótesis.
Otros	CC2	Revisar preguntas que estén mal diseñadas o no sean comprobables.
	CF8	Criticar sus propias presentaciones, tanto por la calidad de los argumentos científicos como por el estilo.
	CF10	<i>Trabajo colaborativo:</i> Discutir las ideas en pequeños grupos como parte del proceso científico.
	DE7	Trabajar colaborativamente para diseñar y construir el experimento.
	CC8	<i>Ética y comunicación:</i> Divulgar los resultados de manera profesional y ética (evitar plagio).
	CF5	Comunicar resultados de forma ética y efectiva en forma oral o escrita.
	CC7	<i>Contextualizar y transferir conocimiento:</i> Transferir conocimiento entre diferentes contextos, aplicando los mismos a diferentes situaciones.
	TPL10	Aplicación de la física a sistemas familiares, dispositivos de consumo, dispositivos médicos y procesos industriales.

La búsqueda de acuerdos entre anotadores puso en evidencia que, aunque se trabaje solo con seis clases, resulta muy complejo considerar simultáneamente la diversidad de descripciones asociadas a cada una. Disminuir la cantidad de clases favorece una mejor discriminación al momento de determinar si un enunciado corresponde o no a la etiqueta de una clase. Incluso, resulta conveniente proporcionar ejemplos que orienten la inclusión o exclusión de un enunciado dentro de una clase, ya que, en algunos casos, la interpretación de la descripción de la clase podía dar lugar a ambigüedades. Este es el caso del código TPL8, cuya descripción corresponde a “Realiza una descripción del aparato y del procedimiento de medición” y que fue asociado a la categoría *Recordar*. La mención de los materiales utilizados implica recuperarlos de la memoria, listarlos o nombrarlos, acciones que se corresponden con este proceso cognitivo, ya que implican identificar y evocar información previamente conocida. Sin embargo, al explicar el procedimiento de medición, el estudiante no solo recuerda mientras escribe, sino que también describe un procedimiento aplicado. Esta descripción involucra el uso de diferentes instrumentos de medición y la realización de ajustes en el diseño experimental. Estas acciones están más asociadas con el desarrollo de habilidades prácticas de laboratorio, enmarcadas dentro de la clase *Aplicar* (ejecutar un procedimiento o implementar).

La falta de acuerdo o consenso inicial entre los anotadores, originada por las diversas interpretaciones de los textos y clases, impulsó la elaboración de una nueva guía de anotación más específica y más simple. Dicha guía se diseñó buscando disminuir ambigüedades, ofreciendo primero una definición de cada uno de los procesos cognitivos de la TBA para luego crear subcategorías que integraban los códigos de las habilidades propuestas por la AAPT. Se incorporó un nivel adicional de detalle, incluyendo comentarios descriptivos de cada habilidad, ejemplos de inclusión con su justificación, así como descripciones y ejemplos de exclusión para delimitar la correspondencia o no de una clase. Es pertinente señalar que no se hallaron ejemplos correspondientes a la categoría Crear en los informes consultados para la elaboración de la guía, dado que las prácticas del curso no contemplaban el grado de apertura requerido para este proceso cognitivo, de acuerdo con la descripción de habilidades. Tomando como referencia trabajos anteriores (Fussell et al., 2025), y con el fin de estandarizar el proceso, se acordó con todos los anotadores que la unidad de análisis sería el enunciado, y se excluyeron las descripciones de las figuras, los pies de página y los cálculos o desarrollos de ecuaciones. Si bien esta guía, con categorías y subcategorías cuyo detalle se expone en el [Apéndice E](#), resultó fundamental para alcanzar un mayor acuerdo entre anotadores, también planteó nuevas dificultades. Su extensión y la tendencia a anotar mediante subcódigos específicos, en lugar de recurrir a la categoría principal, hicieron que la tarea de anotación fuera altamente compleja, incrementando significativamente el tiempo y la atención requeridos durante el proceso.

Con el propósito de optimizar el esquema de clasificación y en respuesta a la retroalimentación proporcionada por los anotadores, se diseñó una nueva guía de anotación, que fue la versión definitiva empleada en este trabajo. Dicha guía quedó conformada por cinco clases, organizadas según los procesos cognitivos de la TBA. En el nivel *Aplicar*, se definió la clase APL, relacionada con el desarrollo de habilidades propias de la medición en laboratorio. Esta clase incluye el uso de instrumentos de medición, la realización de prácticas manuales durante el diseño experimental, la calibración de instrumentos y los ajustes necesarios para obtener medidas coherentes. En el nivel *Analizar*, se definieron dos clases: AND y ANG. La clase AND se vincula con el análisis de resultados experimentales mediante el uso de software de análisis de datos o de herramientas estadísticas. Incluye la comparación de resultados con modelos matemáticos o simulaciones, la predicción de resultados y el análisis

de las relaciones entre los datos experimentales. La clase ANG corresponde al análisis de gráficos de medidas experimentales, lo que implica extraer información de representaciones gráficas, analizar la validez y las limitaciones de los datos presentados, realizar ajustes, relacionar los parámetros de ajuste con cantidades físicas y comparar los resultados con modelos matemáticos. Finalmente, el nivel más alto utilizado fue *Evaluar*, en el que se definieron las clases EIM y EDE, asociadas, respectivamente, a la evaluación de los efectos de la incertidumbre en la medición y a la evaluación del diseño experimental. La clase EDE involucra la evaluación del diseño o del procedimiento experimental, mediante el reconocimiento de aciertos, ventajas y limitaciones, así como la formulación de mejoras para el dispositivo o el procedimiento. También incluye la identificación y resolución de problemas durante la ejecución del experimento. En cambio, la clase EIM refiere a la síntesis de información disponible considerando la incertidumbre, con el fin de extraer conclusiones, así como a la evaluación de los resultados mediante juicios que contemplen la incertidumbre y las fuentes de error.

En la [Tabla 4.3](#) se presenta una versión resumida de la guía de anotación final, donde se especifica la etiqueta correspondiente a cada clase, su descripción breve y los criterios de inclusión asociados (véase el instrumento completo en el [Apéndice F](#)). La guía de anotación completa incorpora un apartado de observaciones orientado a clarificar situaciones de ambigüedad, especificando por ejemplo clases en que el etiquetado de un enunciado es múltiple y presenta ejemplos de exclusión para facilitar la toma de decisiones. Cada uno de los criterios de inclusión definidos fue vinculado y alineado con las habilidades propuestas por la AAPT.

Luego de las distintas etapas de elaboración, revisión y ajuste de la guía de anotación, así como de la retroalimentación aportada por los anotadores durante su evaluación, se resolvió excluir los niveles extremos de la TBA. Se desestimó la categoría *Recordar*, por tratarse del proceso cognitivo más básico y por su escasa relevancia para los objetivos del curso y de esta investigación. Cabe agregar que, este nivel se asocia principalmente al *fundamento teórico*, sección que no fue considerada en el corpus, dado que el análisis se restringió a *metodología*, *análisis de resultados* y *conclusiones*. Del mismo modo, también se excluyó el nivel *Crear*, debido a la escasa evidencia de este proceso cognitivo en el análisis de la muestra de informes.

Tabla 4.3: Clases, códigos y criterios de inclusión definidos en la guía de anotación.

Etiqueta	Descripción breve	Criterios de inclusión
APLICAR: Realizar o usar un procedimiento en determinada situación. La categoría aplicar comienza con la ejecución estrecha y estructurada en la que el conocimiento procedimental se aplica casi de forma rutinaria. Luego continúa con una implementación cada vez menos estructurada, en la que debe seleccionarse el procedimiento para adaptarse a la situación.		
APL	Desarrolla habilidades propias de la medición en laboratorio. Nivel TBA: Aplicar.	Describe instrumentos de medición utilizados o el procedimiento necesario para realizar medidas coherentes. Se evidencia el desarrollo de algunas habilidades prácticas de laboratorio manuales. Utiliza herramientas comunes de recolección de datos, como por ejemplo videos, interfaz o software para extraer datos físicos. Calibra instrumentos para su uso.
ANALIZAR: Dividir el material en sus partes constituyentes y determinar relaciones entre las partes y/o con una estructura o propósito general.		
ANG	Analiza gráficos de medidas experimentales. Nivel TBA: Analizar.	Analiza gráficos extrayendo información de los mismos. Analiza críticamente la validez y limitaciones de los datos presentados en el gráfico. Analiza datos utilizando un software o un análisis estadístico sobre los datos representados en forma gráfica. (Por análisis estadístico básico se entiende aplicar mínimos cuadrados, resultados de la media y la desviación estándar.) Representa datos de forma adecuada para extraer cantidades físicas de los parámetros de ajuste. Realiza ajustes de curvas básicos y relaciona los parámetros de ajuste con cantidades físicas. Compara sus resultados con modelos matemáticos, computacionales o simulaciones.
AND	Analiza datos o resultados experimentales. Nivel TBA: Analizar.	Compara sus resultados con modelos matemáticos, computacionales o simulaciones. Analiza datos utilizando un software o análisis estadístico. Analiza para obtener por ejemplo la media, la desviación estándar. Realiza conexiones lógicas entre el contexto teórico, la evidencia experimental y sus explicaciones, fundamentos o razonamientos. Utiliza modelos para predecir resultados.
EVALUAR: Emitir juicios basados en criterios y estándares; detectar la efectividad de un procedimiento a medida que se implementa; detectar lo apropiado de un procedimiento para un problema dado.		
EDE	Evalúa el diseño (procedimiento o montaje) experimental. Nivel TBA: Evaluar.	Evalúa el diseño experimental, reconoce aciertos, ventajas y/o limitaciones en el diseño (procedimiento o montaje) experimental como por ejemplo calidad de datos, costo, tiempo, equipo disponible y condiciones de seguridad. Diagnostica, realiza o sugiere mejoras al dispositivo experimental y/o procedimiento. Sugiere mejoras a futuro, posteriormente a la realización del experimento. Identifica, propone y/o soluciona problemas durante el procedimiento experimental. Evalúa el dispositivo utilizado y propone un dispositivo apropiado para realizar la medición comprendiendo las limitaciones y/o incertidumbres asociadas a estos.
EIM	Evalúa el impacto de la incertidumbre en la medición. Nivel TBA: Evaluar.	Sintetiza la información disponible, basándose en los resultados e incertidumbres, para llegar a una conclusión coherente. Evalúa los resultados y emite un juicio teniendo presente la incertidumbre en la medición.

4.2. Construcción del corpus anotado.

Una vez consolidada la guía de anotación, se procedió a su aplicación para la construcción del corpus anotado. Si bien en el diseño y refinamiento del instrumento participaron tres evaluadores, la etapa final de anotación se desarrolló en conjunto con un único anotador externo. Luego de alcanzar un nivel de acuerdo intercodificador considerado adecuado para la tarea, se continuó con el proceso con el fin de ampliar la cantidad de datos anotados.

La anotación se realizó con el software de análisis cualitativo VERBI Software (2021), que permitió configurar el sistema de códigos, segmentar los informes y extraer los enunciados seleccionados para el análisis. A partir de este procedimiento, se elaboró un documento con los enunciados que debían ser etiquetados por el anotador externo. Junto con ello, se les proporcionó también el informe completo de origen como material de consulta, de modo que, en caso de ser necesario, pudieran recuperar el contexto original del fragmento.

En una primera instancia, se suministró al segundo anotador un conjunto de enunciados extraídos de cuatro informes, seleccionándose uno al azar de cada práctica de laboratorio del semestre. El proceso de anotación admitió el etiquetado múltiple, es decir, tras la lectura de cada enunciado, se asignaron tantas etiquetas como clases se identificaron en él. Por ejemplo, si un fragmento de texto evidenciaba el uso de elementos de medición y, simultáneamente, una evaluación sobre la pertinencia de dicho instrumento, el enunciado recibía ambos códigos: APL y EDE..

Una vez alcanzado el acuerdo en esta primera fase, se proporcionó un nuevo conjunto de enunciados extraídos de otros cuatro informes, manteniendo el mismo criterio de distribución, es decir, uno por cada práctica. De este modo, se trabajó con un total de ocho documentos de referencia, cuyos enunciados fueron utilizados para evaluar el grado de concordancia entre ambos anotadores mediante el coeficiente κ de Cohen.

Tabla 4.4: Nivel de acuerdo entre evaluadores humanos según el estadístico κ de Cohen para cada clase interpretado de acuerdo a Landis y Koch (1977)

Clase	κ de Cohen	Nivel de concordancia
APL	0.87	Casi perfecto
AND	0.62	Sustancial
ANG	0.70	Sustancial
EDE	0.89	Casi perfecto
EIM	0.86	Casi perfecto

La Tabla 4.4 presenta los valores del estadístico κ de Cohen obtenidos para cada clase, utilizado para evaluar la concordancia de la anotación entre evaluadores (Cohen, 1960). Como criterio de calidad se consideró aceptable, al menos, una concordancia sustancial. La interpretación de los resultados se realizó siguiendo los umbrales de Landis y Koch (1977) (véase Sección 3.2). Los resultados muestran una concordancia casi perfecta en la clases APL, EDE y EIM, y sustancial en ANG y AND.

Tabla 4.5: Distribución de enunciados anotados por clase antes y después de ampliar el corpus. El total de enunciados de las clases no coincide necesariamente con la suma porque un enunciado puede tener más de una etiqueta.

Clase	Cantidad de enunciados	
	Inicial	Final
APL	265	265
ANG	234	234
AND	369	369
EIM	128	197
EDE	140	245
Sin etiqueta	323	323
Total enunciados del corpus	1121	1594

La etapa inicial de anotación se desarrolló a partir de una distribución proporcional de informes por práctica. En una primera instancia, se anotaron ocho informes destinados a evaluar la concordancia entre ambos anotadores. Una vez alcanzado un nivel de acuerdo considerado adecuado para la tarea, el proceso continuó hasta completar un total de 24 informes. La cantidad de etiquetas obtenida en esta fase se presenta en la columna “Inicial” de la Tabla 4.5. Como puede observarse, algunos códigos presentaron una frecuencia de aparición menor, particularmente EIM (evaluación de la relevancia de la

incertidumbre en la medición) y EDE (evaluación del diseño experimental). Considerando que los enunciados correspondientes a estas categorías suelen concentrarse principalmente en el apartado de *Conclusiones* de los informes de laboratorio, se realizó una ampliación del corpus con el propósito de mejorar la representatividad de dichas clases. Para ello, se incorporó la anotación exclusiva de este apartado en 52 informes adicionales: 44 del curso 2023 y 8 del curso 2024.

Tabla 4.6: Distribución de la cantidad de enunciados pertenecientes a cada clase y por corpus.

Clase	Corpus		
	Entrenamiento	Desarrollo	Evaluación
APL	185	40	40
ANG	164	35	35
AND	258	55	56
EDE	171	37	37
EIM	138	29	30

Una vez consolidado el corpus anotado, se procedió a su partición en tres subconjuntos: entrenamiento, desarrollo y evaluación. El corpus de entrenamiento quedó conformado por el 70 % de los enunciados, mientras que los conjuntos de desarrollo y evaluación representaron, cada uno, el 15 % restante. En esta partición se procuró preservar la representatividad de cada clase en proporciones equivalentes entre los tres subconjuntos. La [Tabla 4.6](#) muestra la distribución de enunciados por clase en cada subconjunto, utilizada en las etapas posteriores de entrenamiento, ajuste y evaluación de los modelos.

Capítulo 5

Clasificación de enunciados: experimentos y resultados.

En este capítulo se describen las técnicas de PLN utilizadas en el desarrollo de este trabajo, junto con las pruebas realizadas y la configuración final de los modelos de AA empleados para la clasificación de texto. Además, se presentan los resultados obtenidos por los diferentes clasificadores en cada una de las clases analizadas y se examinan los errores de clasificación con el fin de comprender e identificar sus principales limitaciones.

5.1. Clasificación con técnicas de PLN.

Las primeras pruebas y aproximaciones a la clasificación automática se fundamentaron en la estructura de la TBA, utilizando los seis niveles de procesos cognitivos y empleando para ello algoritmos de aprendizaje automático clásicos, tales como Multinomial Naive Bayes (MNB), Multi-layer Perceptron (MLP) y Support Vector Machine (SVM) (véase [Sección 3.3](#)). En esta etapa, también se realizó la identificación de los k términos más relevantes (k-Best). Esta decisión metodológica se basó en trabajos previos en PLN en que se utiliza la TB o TBA y diversas técnicas de PLN para clasificar preguntas utilizadas en evaluaciones (Gani et al., [2023](#); Meher & Mall, [2023](#); Mohammed & Omar, [2020](#); Yahya et al., [2012](#)).

La búsqueda de mejoras en el modelado, la guía de anotación y el acuerdo entre anotadores condujo a una nueva modalidad de anotación basada en cinco clases descritas en la [Sección 4.2](#) (APL, AND, ANG, EDE y EIM). A partir de

esta redefinición, fue necesario reconfigurar también la etapa de clasificación automática.

En este punto es necesario precisar que, si bien el problema se define originalmente como de etiquetado múltiple, el abordaje automatizado se realizó mediante la descomposición de la tarea en problemas binarios independientes. Esto implicó transformar el corpus completo en una clasificación binaria para cada una de las cinco clases, donde el objetivo quedó restringido a determinar la presencia o ausencia de una etiqueta específica en cada enunciado, replicando este procedimiento de forma independiente para cada clase.

En una fase inicial, las pruebas se realizaron sobre el corpus de desarrollo utilizando los clasificadores clásicos MNB, MLP y SVM. A estos se añadió el modelo LR, dado que la revisión de la literatura en PER (Fussell et al., 2025) lo reporta como un algoritmo con buen desempeño en tareas de clasificación similares.

Con el objetivo de mejorar las métricas de evaluación de desempeño obtenidas con estos modelos, se ejecutaron pruebas experimentales aplicando técnicas de preprocesamiento. Una de ellas consistió en la eliminación de palabras vacías (*stop words*) con el objetivo de reducir la presencia de términos de alta frecuencia y bajo contenido semántico. Este tipo de palabras comprende, entre otras, artículos, pronombres y preposiciones, como “a”, “le”, “del” o “allá”. La lista empleada fue ajustada específicamente para este trabajo y puede consultarse en el [Apéndice G](#).

De forma complementaria, los modelos de SVM y LR, dadas sus características, permitieron evaluar el impacto de configurar la *ponderación de clases de forma balanceada* (Scikit-learn developers, s.f.) sugerida en el artículo Fussell et al. (2025). Esta decisión está fundada en el desbalance existente entre las clases del corpus, condición que induce a los modelos a sesgar sus predicciones hacia la categoría mayoritaria (la más frecuente) en detrimento de las demás. Mediante la incorporación de estos pesos se fuerza al algoritmo a ajustar su sensibilidad para identificar correctamente las etiquetas menos representadas.

Como parte de esta fase de preprocesamiento, también se realizaron pruebas en las que se incorporó la *lematización* de los enunciados antes de su clasificación. Este proceso consiste en reducir las palabras a su lema y tiene como objetivo disminuir la variación léxica del corpus. Para la implementación técnica de esta tarea se utilizó Stanza, una biblioteca de PLN en Python desarrollada por el grupo de investigación de Stanford NLP Group (2020). El

resultado de aplicar la lematización a un enunciado se muestra a continuación:

Enunciado original:

Estos datos eran recolectados por el cronómetro y por el fotosensor, el cual estaba conectado a la interfaz y eran traducidos en la PC mediante los programas mencionados.

Enunciado lematizado:

este dato ser recolectar por el cronómetro y por el fotosensor el cual estar conectado a la interfaz y ser traducido en la pc mediante los programa mencionado.

Figura 5.1: Ejemplo de un enunciado y su versión lematizada.

Las pruebas anteriores permitieron seleccionar a MLP y LR como los clasificadores clásicos con mejores métricas para las diferentes clases a clasificar en este trabajo.

De forma complementaria, se exploró un método de clasificación de texto que combina representaciones vectoriales y aprendizaje basado en instancias (Abdelmounaim & Madani, 2024; Li et al., 2018; Liapis et al., 2024). En este método, el conjunto de datos de entrada se transforma en *embeddings* que capturan sus características más relevantes. Luego, estos *embeddings* se utilizan como entrada de un clasificador k-NN, que asigna a cada instancia la etiqueta de clase correspondiente en función de las clases mayoritarias entre sus vecinos más cercanos en el espacio vectorial.

Para la construcción de las representaciones vectoriales de los enunciados (sentence embeddings) se utilizaron dos modelos: Multilingual-E5 (Wang et al., 2024) y EmbeddingGemma (300M) de Google (Google AI, 2025; Vera et al., 2025). Estas representaciones no solo se utilizaron como insumo para la clasificación, sino que también permitieron visualizar la distribución de los enunciados en el espacio vectorial y explorar su posible separación entre clases. En particular, el *Análisis de Componentes Principales* para la reducción de la dimensionalidad (PCA) permitió reducir la dimensionalidad de estas representaciones y obtener una visualización bidimensional de los enunciados, lo que facilitó observar si las clases tendían a formar agrupamientos distinguibles entre sí o, por el contrario, presentaban solapamientos. A modo de ejemplo, la Figura 5.2a muestra esta representación para la clase APL del corpus de entrenamiento. A partir de su observación, puede advertirse que, si bien los enunciados pertenecientes a esta clase tienden a agruparse en función de ciertas características, representados en color violeta, también se observa superposición con enunciados que no pertenecen a la clase, mostrados en color

verde. Esto sugiere que algunos enunciados no pertenecientes a APL comparten rasgos similares con aquellos que sí integran esta clase. Por su parte, la [Figura 5.2b](#) presenta la representación correspondiente a la clase EDE en el corpus de entrenamiento, donde se observa un mayor grado de solapamiento entre los enunciados que pertenecen y los que no pertenecen a esta clase. Esto indica que la discriminación de la clase EDE resulta más compleja que la de APL, debido a una menor separación entre ambos grupos en el espacio vectorial.

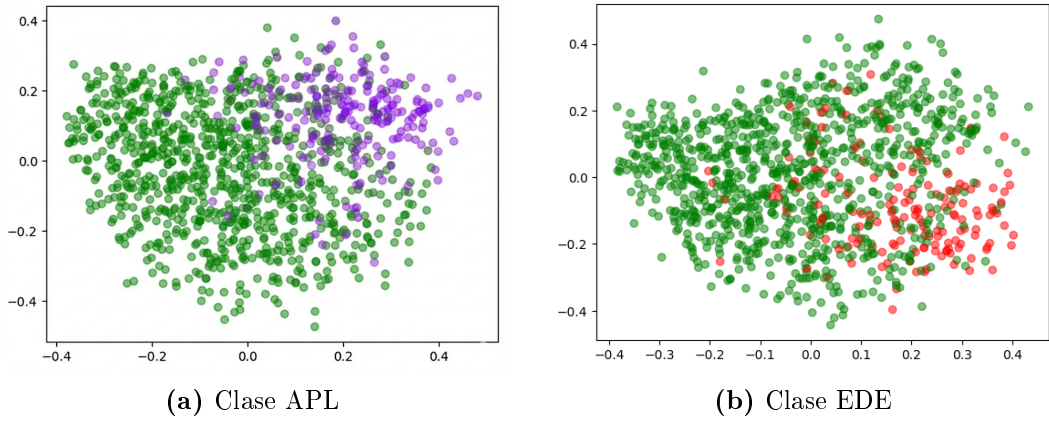


Figura 5.2: Representación en dos dimensiones de la distribución de los *sentence embeddings* en el corpus de entrenamiento utilizando el modelo Gemma. En ambas subfiguras, los puntos verdes representan enunciados que no pertenecen a la clase considerada y los puntos rojos representan enunciados pertenecientes a dicha clase: APL en la subfigura (a) y EDE en la subfigura (b).

Estas representaciones se utilizaron posteriormente como entrada para el algoritmo k -NN. Dado que el desempeño de este clasificador depende de la cantidad de vecinos considerados, fue necesario evaluar distintos valores de k , es decir, diferentes cantidades de vecinos más cercanos. Los resultados obtenidos indicaron que el mejor rendimiento se alcanzaba con $k = 5$ para *Embedding-Gemma* y para *Multilingual-E5*.

En cuanto al uso de métodos de clasificación basados en LLM, se inició utilizando el modelo *Llama 3 8B Instruct* (Meta, 2024). Su elección se fundamentó, en que se trata de un modelo de código abierto que, en comparación con otras alternativas, no requiere una infraestructura computacional de alta exigencia para su implementación. Para ello, se implementó una estrategia de diseño de instrucciones (*prompting*) a la que luego incorporó la técnica de aprendizaje con pocos ejemplos (*few-shot learning*). La selección de estos ejemplos se

realizó de dos formas, seleccionando manualmente los ejemplos y de manera dinámica, empleando el método de los k-vecinos más cercanos (k-NN) para identificar los casos más relevantes según el contexto.

A partir del análisis de los resultados preliminares obtenidos con el modelo *Llama 3 8B Instruct*, se advirtió que, en algunas categorías de clasificación, su rendimiento no resultaba superior al de los clasificadores clásicos. Por este motivo, se optó por ampliar el estudio incorporando modelos adicionales. En particular, se añadió el modelo *Gemini 2.5 Flash Lite* (Google DeepMind, 2025) mediante la API de Google, así como el modelo de código abierto *GPT:OSS 20B* (OpenAI, 2025b), publicado durante el transcurso de esta investigación. Al igual que *Llama 3 8B Instruct*, este último puede ejecutarse localmente, lo que favorece un mayor control tanto de los costos operativos como de la privacidad de los datos utilizados.

Con los LLM también se realizaron diversas pruebas, comenzando por el diseño del *prompt*, dado que pequeñas variaciones en su formulación producían resultados diferentes. La construcción de un *prompt* que permitiera alcanzar las mejores métricas posibles constituyó una tarea compleja, al igual que la selección manual de ejemplos relevantes. Para ello, se trabajó sobre los resultados obtenidos en el corpus de desarrollo y se examinaron especialmente los casos clasificados de manera incorrecta, en particular los FN y FP, con el fin de ajustar y formular reglas de clasificación que fueran, a la vez, suficientemente generales para aplicarse a distintos corpus y lo bastante específicas para la clase que se pretendía identificar.

Se incorporó, además, una técnica mixta en la que, junto con el *prompt*, se suministraban ejemplos seleccionados mediante *k*-NN a partir de las características obtenidas con los *sentence embeddings*. En la Figura 5.3 se muestran tres enunciados de cada clase seleccionados por este procedimiento. La lectura de estos ejemplos permite observar que todos remiten a un tema recurrente: el termómetro de alcohol, en relación con su incertidumbre, apreciación o precisión.

A modo de ejemplo, la Figura 5.4 ilustra cómo las variaciones introducidas en la construcción del *prompt* se reflejan en la métrica Macro F1-score, según la inclusión o no de ejemplos y el criterio adoptado para su selección. En el caso de Gemini, se presentan los resultados obtenidos mediante dos estrategias para proporcionar ejemplos al LLM en la tarea de clasificación. En la primera, se suministraron manualmente dos ejemplos pertenecientes a la clase APL y dos

Vecinos más cercanos de un enunciado encontrados con k-NN	
Enunciado a clasificar	
Etiqueta real: 0 <i>“Para las barras de errores en el eje de las ordenadas se usó la incertidumbre dada por el fabricante mientras que para el eje de las abscisas se usó la apreciación del termómetro de alcohol (1 kelvin)”.</i>	
Enunciados cercanos pertenecientes a la clase EDE (etiqueta : 1)	
1.1. <i>“La medición de la temperatura con un termómetro de alcohol nos brinda una incertidumbre la cual puede afectar de forma directa a los resultados obtenidos, en caso de haber realizado este experimento con un termómetro más preciso podríamos atenuar la afección de este error a los resultados obtenidos, los cuales podrían ser más cercanos a los resultados esperados teóricamente”.</i> 1.2. <i>“El termómetro de alcohol se utilizó como instrumento patrón para calibrar, se podría haber elegido un instrumento más preciso que disminuiría las incertidumbres de las medidas finales haciendo que los resultados fueran más certeros, pero esto no es necesario”.</i> 1.3. <i>“Debido al tipo de experimento y a la cantidad de medidas no fue necesario acudir a instrumentos más precisos que el termómetro a alcohol”.</i>	
Enunciados cercanos no pertenecientes a la clase EDE (etiqueta: 0)	
0.1. <i>“Para las barras de errores se usó en el eje de las ordenadas la incertidumbre dada por el fabricante y para el eje de las abscisas la apreciación del termómetro de alcohol (1 grado celsius)”.</i> 0.2. <i>“Las incertidumbres del termómetro de alcohol y del multímetro son de 1 °C y 30,7Ω respectivamente”.</i> 0.3. <i>“En consideración del análisis de datos para las próximas tres secciones, cabe mencionar que el termómetro de alcohol utilizado presenta una apreciación de 1 °C, por lo que se le atribuye una incertidumbre de tipo B igual a su estimación (0,5 °C)”.</i>	

Figura 5.3: Ejemplo de clasificación que muestra los tres vecinos semánticamente más cercanos, pertenecientes o no a la clase EDE, recuperados mediante el algoritmo k-NN a partir de representaciones obtenidas con *Sentence Embeddings Gemma*. La etiqueta 0 corresponde a enunciados que no pertenecen a la clase EDE y la etiqueta 1 a los que sí pertenecen.

no pertenecientes a ella, esta técnica corresponde al enfoque *Few-Shot* (FS). En la segunda, se combinó *sentence embeddings* con *k*-NN para seleccionar, en función de su similitud con el enunciado a clasificar, tres ejemplos de la clase APL y tres no pertenecientes a la clase. Como puede observarse en la [Figura 5.4](#), la métrica utilizada para evaluar el desempeño del clasificador basado en un LLM puede mejorar cuando se realizan ajustes adecuados.

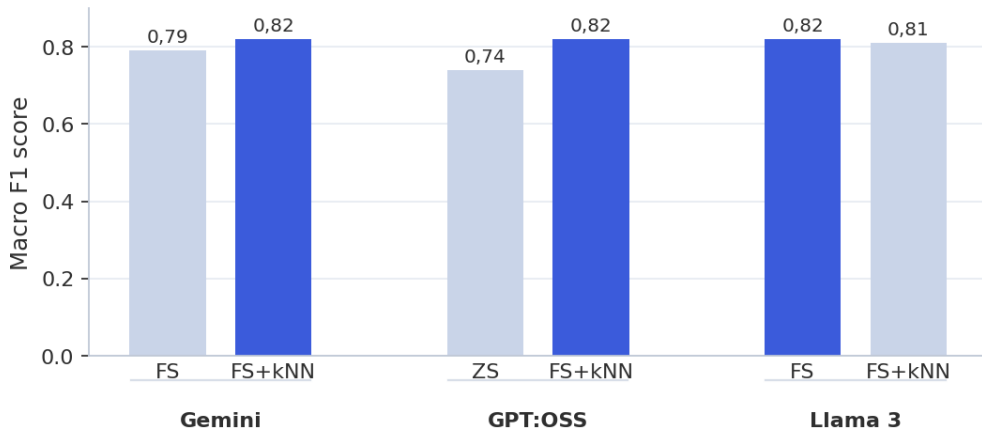


Figura 5.4: Comparación del *Macro F1-score* para la clase APL en los distintos LLM y estrategias evaluadas sobre el corpus de desarrollo. Se agrupan los resultados por familia de modelos: Gemini, GPT-OSS y Llama 3. Las siglas ZS y FS corresponden a las estrategias *zero-shot* y *few-shot*. Dentro de cada modelo, el tono más oscuro destaca el mejor desempeño alcanzado.

Para el modelo GPT se presenta la variación de la métrica Macro F1-score entre una configuración sin ejemplos, denominada *Zero-Shot* (ZS), y otra en la que se proporcionan tres ejemplos de cada clase seleccionados mediante *k*-NN.

En el caso de Llama 3, para la clase APL no se observaron variaciones significativas en el desempeño al incorporar ejemplos en el *prompt* con kNN.

El desarrollo de las pruebas experimentales y la gestión del código del proyecto se realizó en la plataforma *Google Colaboratory* (Google, 2017), entorno que facilitó la ejecución de los algoritmos de clasificación descritos anteriormente.

Una vez concluida esta etapa de exploración de distintos enfoques de clasificación, resumida en la [Tabla 5.1](#), se seleccionaron los modelos y configuraciones que alcanzaron el mejor desempeño en la clasificación de las cinco clases analizadas. Los resultados obtenidos permitieron advertir que ningún modelo presentaba un rendimiento superior de manera uniforme en todas las clases. Por el contrario, se observó una variabilidad en el desempeño, de modo que

ciertos algoritmos resultaban más adecuados para identificar algunas clases que otras.

Finalmente, considerando los resultados, se definió un conjunto final de clasificadores. Dicho conjunto quedó integrado por la Regresión Logística (LR), como representante de los métodos clásicos, por un clasificador que combina *Sentence Embeddings* con el algoritmo k -NN, y por los tres LLM mencionados anteriormente. Finalmente, las predicciones de este conjunto se integraron mediante una técnica de *voto mayoritario* (Kuncheva, 2004; Kyritsis et al., 2025), en la cual cada modelo asigna una clase a cada enunciado y la decisión final se determina a partir de la clase más votada entre las predicciones individuales. El uso de un número impar de clasificadores simplifica este procedimiento, ya que evita empates y permite establecer la clase final cuando al menos la mitad más uno de los modelos coincide en la misma clasificación.

Tabla 5.1: Pruebas con distintas técnicas de clasificación realizadas sobre el corpus de desarrollo.

Modelo	Pruebas realizadas	Configuración final
BoW (bolsa de palabras)	Frecuencia de verbos. Frecuencia de verbos y secuencias generales de palabras. Frecuencia de verbos y secuencias específicas de Física Experimental.	–
Clasificadores clásicos		
MNB	Por defecto y eliminación de <i>stop words</i> .	–
MLP	Por defecto, eliminación de <i>stop words</i> , ajuste de hiperparámetros y lematización.	<code>hidden_layer_sizes=(200, 100, 50), activation='logistic', max_iter=500; lematización con Stanza.</code>
LR	Por defecto, eliminación de <i>stop words</i> , ponderación de clases, ajuste de hiperparámetros y lematización.	<code>penalty='l2', class_weight='balanced', random_state=42, max_iter=10000; lematización con Stanza.</code>

Continúa en la página siguiente

Tabla 5.1: Pruebas con distintas técnicas de clasificación realizadas sobre el corpus de desarrollo.

Modelo	Pruebas realizadas	Configuración final
SVM	Por defecto, eliminación de <i>stop words</i> y – ponderación de clases.	
Clasificadores basados en instancias		
<i>Multilingual-e5-large</i> + k -NN	Clasificación a partir de <i>sentence embeddings</i> y búsqueda de vecinos más cercanos; prueba de distintos valores de k .	Mejor resultado con $k = 5$.
<i>EmbeddingGemma</i> (300M) + k -NN	Clasificación a partir de <i>sentence embeddings</i> ; prueba de distintos valores de k .	Mejor resultado con $k = 5$.
Grandes modelos de lenguaje (LLM)		
Llama 3 8B Instruct	Zero Shot, ajuste y comparación de variantes de <i>prompt</i> , Few Shot con ejemplos seleccionados manualmente y <i>prompt</i> con ejemplos dinámicos seleccionados mediante <i>sentence embeddings</i> y k -NN.	<i>Prompt</i> + Few Shot con ejemplos seleccionados manualmente para algunas clases, para otras sin ejemplos. Temperatura = 0
Gemini 2.5 Flash Lite	Zero Shot, ajuste y comparación de variantes de <i>prompt</i> , Few Shot con ejemplos seleccionados manualmente y <i>prompt</i> con ejemplos dinámicos seleccionados mediante <i>sentence embeddings</i> y k -NN.	<i>Prompt</i> + ejemplos dinámicos con <i>EmbeddingGemma</i> (300M) y $k = 3$. Temperatura = 0
GPT-OSS 20B	Zero Shot, ajuste y comparación de variantes de <i>prompt</i> , Few Shot con ejemplos seleccionados manualmente y <i>prompt</i> con ejemplos dinámicos seleccionados mediante <i>sentence embeddings</i> y k -NN.	<i>Prompt</i> + ejemplos dinámicos con <i>Multilingual-e5-large</i> y $k = 3$. Temperatura = 0

5.2. Configuración y evaluación de los algoritmos de AA utilizados.

En esta sección se describen las configuraciones de los algoritmos de AA que, luego de las pruebas efectuadas sobre el corpus de desarrollo, fueron finalmente seleccionados para esta investigación, junto con sus resultados en el corpus de evaluación. Los desempeños obtenidos en el corpus de desarrollo se

incorporan en esta sección como antecedentes, en la medida en que sirvieron de base para la selección final de los modelos considerados.

Los clasificadores clásicos que lograron mejores resultados sobre el corpus de desarrollo fueron MLP y RL. Al compararlos, como puede verse en la tabla [Tabla 5.2](#), el clasificador que logró un desempeño algo superior al mismo tiempo en las cinco clases fue LR, por lo que fue uno de los seleccionados para el análisis posterior. Para el modelo de LR, el mejor rendimiento se alcanzó adoptando los criterios de configuración sugeridos por Fussell et al. ([2025](#)).

Tabla 5.2: Comparativa de resultados de la métrica *Macro F1-score* entre MLP y LR sobre el corpus de desarrollo para las diferentes clases.

Clase	<i>Macro F1-score</i>	
	MLP	LR
APL	0,83	0,83
AND	0,75	0,75
ANG	0,80	0,79
EDE	0,76	0,78
EIM	0,73	0,74

El siguiente clasificador que se utilizó fue k-NN en su combinación con *Sentence Embeddings Gemma*. Este clasificador se configuró para utilizar los $k = 3$ ejemplos de entrenamiento más próximos al enunciado a clasificar, asignando la clase final por votación mayoritaria entre las clases de esos vecinos.

Los clasificadores restantes se basan en grandes modelos de lenguaje (LLM). En todos los casos se estandarizó la estructura de las instrucciones proporcionadas a los modelos con el objetivo de comparar su desempeño en condiciones equivalentes y evitar favorecer a alguno de ellos mediante variaciones en las indicaciones. Los *prompts* utilizados comparten una estructura común, detallada en la [Figura 5.5](#), que incluye cuatro componentes: asignación de rol, definición de la tarea, criterios de clasificación y restricción del formato de salida.

En primer lugar, se asignó al modelo el rol de clasificador experto en enunciados de Física Experimental. A continuación, se definió la tarea de clasificación binaria, seguida por la especificación de los criterios de inclusión y exclusión en la clase correspondiente. Finalmente, se estableció una restricción en el formato de salida que obliga al modelo a emitir una única respuesta, limitada a los valores ‘1’ o ‘0’, indicando respectivamente la pertenencia o no

del enunciado a la clase analizada.

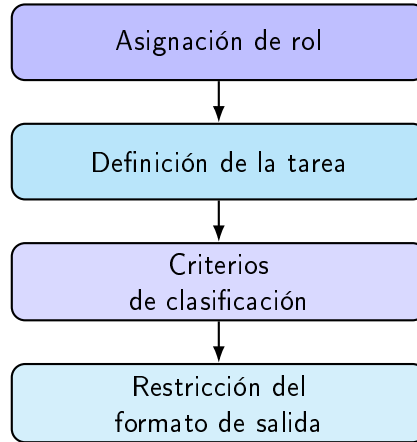


Figura 5.5: Estructura general del *prompt* utilizado para la tarea de clasificación.

Aunque la estructura general de los *prompts* se mantuvo constante, la estrategia de selección de ejemplos incorporados en ellos varió entre los modelos. Esta decisión se basó en los resultados obtenidos en el corpus de desarrollo, donde cada modelo mostró comportamientos diferentes. En el caso de *Llama 3 8B Instruct*, los ejemplos fueron seleccionados manualmente. Para *Gemini*, se utilizaron tres ejemplos de cada clase recuperados mediante un enfoque basado en k-NN y *sentence embeddings* generados con *EmbeddingGemma*, los cuales se presentaron de forma alternada. Por su parte, para *GPT:OSS* se seleccionaron tres ejemplos de cada clase utilizando k-NN combinado con *sentence embeddings* generados con *Multilingual-E5*.

La [Figura 5.6](#) presenta un esquema de la redacción de los *prompts* empleados. En todos los casos se mantienen los cuatro componentes mencionados. Además, para algunas clases fue necesario incorporar reglas adicionales de desempate con el fin de mejorar el desempeño de los modelos.

En cuanto al corpus empleado en este trabajo, quedó conformado por 1594 enunciados. A partir de este conjunto, se realizó una partición en tres subconjuntos de datos: 1115 enunciados para entrenamiento, 239 para desarrollo y 240 para evaluación. Los enunciados del conjunto de entrenamiento se utilizaron, además, como fuente de ejemplos en los *prompts* de los LLM. La distribución de enunciados pertenecientes a cada clase en cada subconjunto de datos se presentó en la [Tabla 4.6](#).

Tras las pruebas realizadas sobre el corpus de desarrollo y la selección de las mejores configuraciones para los distintos algoritmos de clasificación, se

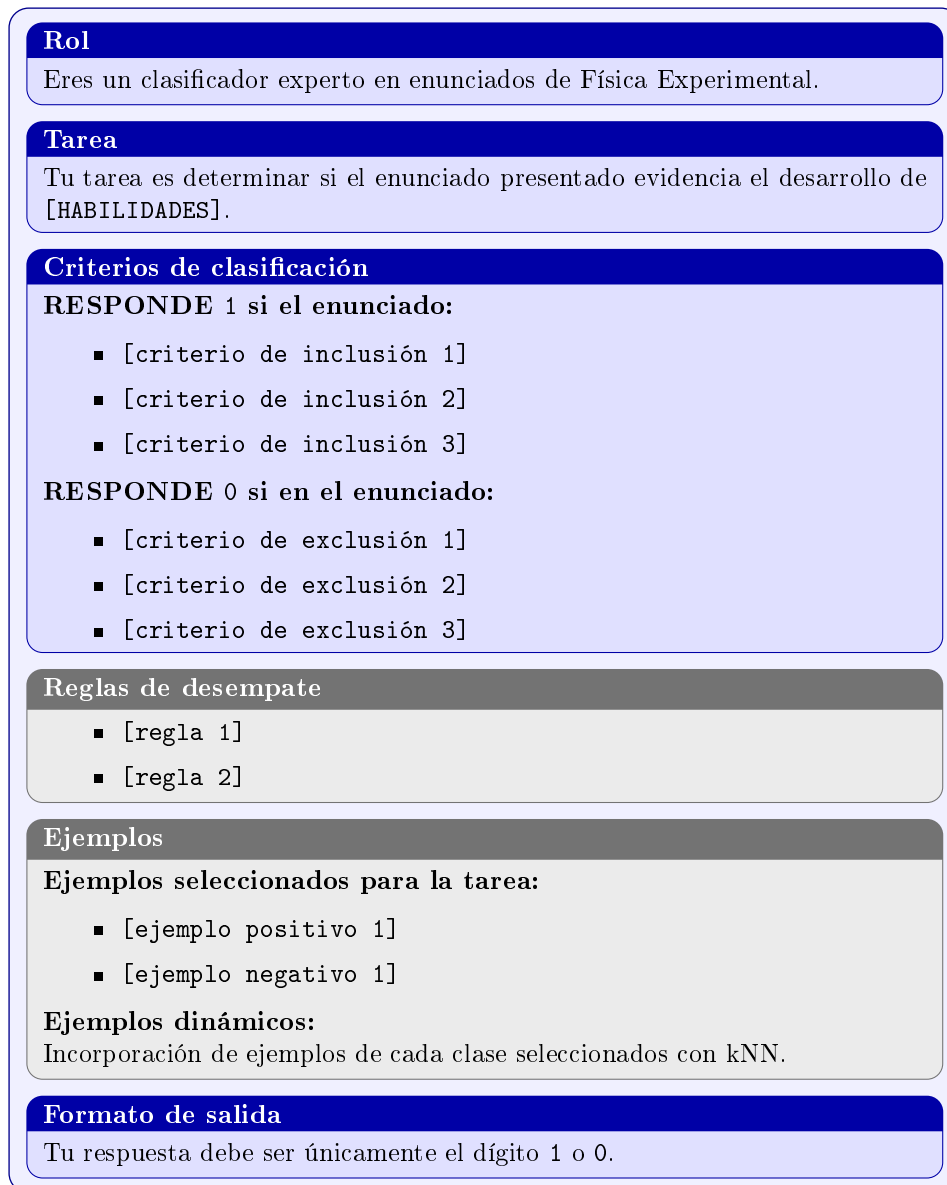


Figura 5.6: Estructura general del *prompt* utilizado para la clasificación binaria. Los bloques en color azul corresponden a componentes permanentes del *prompt*, mientras que los bloques en gris representan elementos opcionales, incorporados o excluidos según el modelo utilizado y su desempeño en las pruebas realizadas.

llevó a cabo la validación final sobre el corpus de evaluación. La [Figura 5.7](#) presenta el *Macro F1-score* obtenido por cada modelo en cada clase e incluye, como referencia, el valor de esta métrica calculado a partir de la comparación entre las clasificaciones realizadas por dos evaluadores humanos. Cabe precisar que el desempeño entre evaluadores humanos reportado es de referencia, dado que fue obtenido a partir de un subconjunto del corpus y no sobre la totalidad

de la muestra empleada para la evaluación de los modelos. Para la obtención de las métricas mencionadas se emplearon los enunciados clasificados por dos evaluadores humanos en el marco del cálculo del coeficiente κ de Cohen, que comprendió 315 enunciados correspondientes a ocho informes de laboratorio (dos por cada práctica).

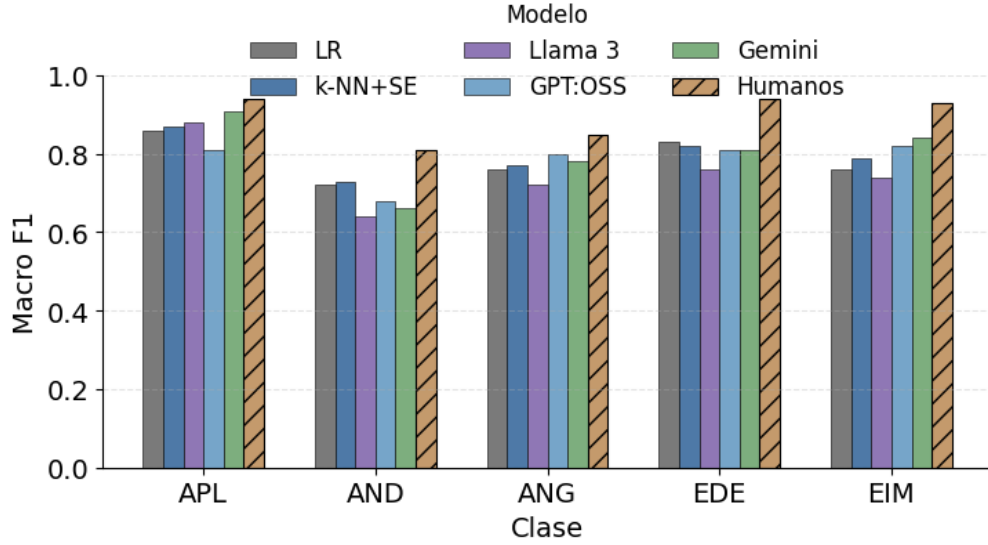


Figura 5.7: *Macro F1-score* para cada clase y modelo en el corpus de evaluación. Se incluye además, como referencia, el valor de esta métrica calculado a partir de la clasificación realizada por evaluadores humanos sobre una muestra de 315 enunciados.

En este trabajo se emplea la métrica *F1-score* para evaluar el desempeño de los modelos, ya que combina *precision* y *recall* y permite valorar el equilibrio entre ambas. Se optó por esta métrica en lugar de la *accuracy*, reportada en otros trabajos (Campbell et al., 2024; Fussell et al., 2025), debido a que el corpus presenta desbalance de clases. En efecto, para cada una de las cinco clases consideradas, la cantidad de enunciados que no pertenecen a la clase supera a la de aquellos que sí pertenecen. La métrica reportada en la ?? corresponde al *Macro F1-score*. Esta métrica calcula el *F1-score* de cada clase por separado y luego promedia esos resultados. En este trabajo, dicho cálculo se realiza para la clasificación binaria de los enunciados, considerando si pertenecen o no a cada una de las cinco clases analizadas. El detalle de las métricas obtenidas puede consultarse en el [repositorio en línea](https://github.com/lspm1102/Metodologia-automatizada-analisis-informes-laboratorio.git)¹.

Una primera lectura de los resultados de la Figura 5.7 permite advertir que

¹<https://github.com/lspm1102/Metodologia-automatizada-analisis-informes-laboratorio.git>

ningún modelo presenta un desempeño superior al de los demás en la mayoría de las cinco clases analizadas. Además, puede observarse que los clasificadores clásicos, como LR y k -NN, alcanzan un desempeño similar al de los LLM, a pesar de que estos últimos requieren mayores recursos computacionales para su ejecución. En contraste, los LLM presentan la ventaja de no demandar una alta cantidad de datos anotados para su utilización.

En la [Figura 5.7](#) también se observa que la clase APL, vinculada al desarrollo de habilidades prácticas propias de la medición en laboratorio, es en general, aquella en la que los algoritmos de clasificación alcanzan el mejor desempeño, con valores prácticamente comparables a los obtenidos por evaluadores humanos. Además, esta clase fue una de las que registró los niveles más altos de concordancia entre los evaluadores humanos (véase la [Tabla 4.4](#)), lo que sugiere que resulta más fácil de discriminar que otras. Esto podría vincularse con características propias de la clase APL, en la que predominan acciones asociadas al uso de instrumentos, la toma de medidas y la configuración del montaje experimental. A su vez, su ubicación predominante en el nivel del proceso cognitivo *Aplicar*, según la TBA, también podría favorecer su identificación.

Los resultados muestran que la clase AND, vinculada al análisis de datos, es, en general, la que presenta el menor rendimiento entre los clasificadores basados en AA. En este caso, también se observa que los LLM alcanzan un desempeño inferior al de modelos como LR o k -NN. Además, AND fue una de las categorías que registró menores niveles de concordancia entre evaluadores humanos. Su definición comprende enunciados relacionados con la comparación de resultados con modelos teóricos, el ajuste de parámetros, el uso de herramientas estadísticas y la formulación de conclusiones basadas en datos. La dificultad observada sugiere que la identificación de estos enunciados no depende solo de reconocer verbos asociados al proceso cognitivo *Analizar* en la TBA, sino también de interpretar el sentido del enunciado en su contexto y establecer si efectivamente se ha realizado un análisis de los datos, y no una mera descripción.

A modo de ejemplo se cita el siguiente enunciado:

“Para mejorar la precisión del período elaboramos un análisis estadístico, realizando 50 mediciones en idénticas condiciones de forma de ampliar considerablemente la muestra.”

Este enunciado pone de manifiesto que la sola mención de la realización

de un “análisis” no implica, por sí misma, que se esté evidenciando explícitamente un proceso de análisis de datos. Este aspecto fue considerado tanto en el diseño de la guía de anotación como en la discusión orientada a alcanzar el acuerdo entre anotadores. En consecuencia, para que un enunciado fuera etiquetado como perteneciente a una determinada clase, debía aportar evidencia de que la acción correspondiente había sido efectivamente realizada. Por esta razón, se estableció que no serían clasificados dentro de esta categoría aquellos enunciados que únicamente expresaran una intención o una acción proyectada a futuro.

Las clases EDE y EIM se enmarcan dentro del nivel del proceso cognitivo *Evaluar* de la TBA. La clase EDE busca identificar enunciados que evidencien una evaluación del diseño o del procedimiento experimental, en los que los estudiantes valoren el método empleado, juzguen posibles modificaciones o sugieran mejoras. Por su parte, la clase EIM se asocia con la evaluación de los efectos de la incertidumbre sobre las medidas realizadas. En ambos casos, el rendimiento de los clasificadores resulta algo inferior en comparación con el acuerdo observado entre clasificadores humanos. Al analizar las matrices de confusión obtenidas para estas clases e interpretadas de acuerdo con la [Figura 3.3](#), se observó que los LLM tienden a cometer más errores de tipo falso positivo (FP)¹. Este resultado sugiere una tendencia de los modelos a clasificar como pertenecientes a estas clases enunciados que, según la anotación experta, no deberían incluirse en ellas.

Dado que ningún clasificador mostró un desempeño claramente superior en la clasificación de las diferentes clases, los resultados se integraron mediante un proceso de votación por mayoría simple. Según este criterio, un enunciado fue asignado a una clase cuando al menos tres clasificadores coincidieron en su pertenencia a ella.

La [Figura 5.8](#) permite analizar, para cada clase, el desempeño del esquema de votación por mayoría simple en relación con el mejor modelo individual y con el valor de referencia de la métrica *Macro F1-score*, determinado a partir de la comparación entre dos evaluadores humanos sobre una muestra del corpus. En ella se observa que, para las clases AND y EDE, el esquema de votación supera el desempeño del mejor modelo individual, mientras que para la clase ANG alcanza un desempeño igual al de dicho modelo. En cuanto a las clases APL y

¹Los resultados completos de las matrices de confusión se encuentran disponibles en el [repositorio](#) asociado a esta tesis.

EIM, el resultado de la votación es algo inferior al de Gemini, que es el modelo de mejor desempeño en ambas clases, sin embargo, sigue siendo superior al del resto de los modelos para esas mismas clases (véase la [Figura 5.7](#)). En términos generales, este esquema permite unificar y compensar las diferencias asociadas a las características de cada modelo particular, además de utilizar un mismo procedimiento para todas las clases, sin necesidad de seleccionar en cada caso el modelo de mejor desempeño según la clase que se desea analizar.

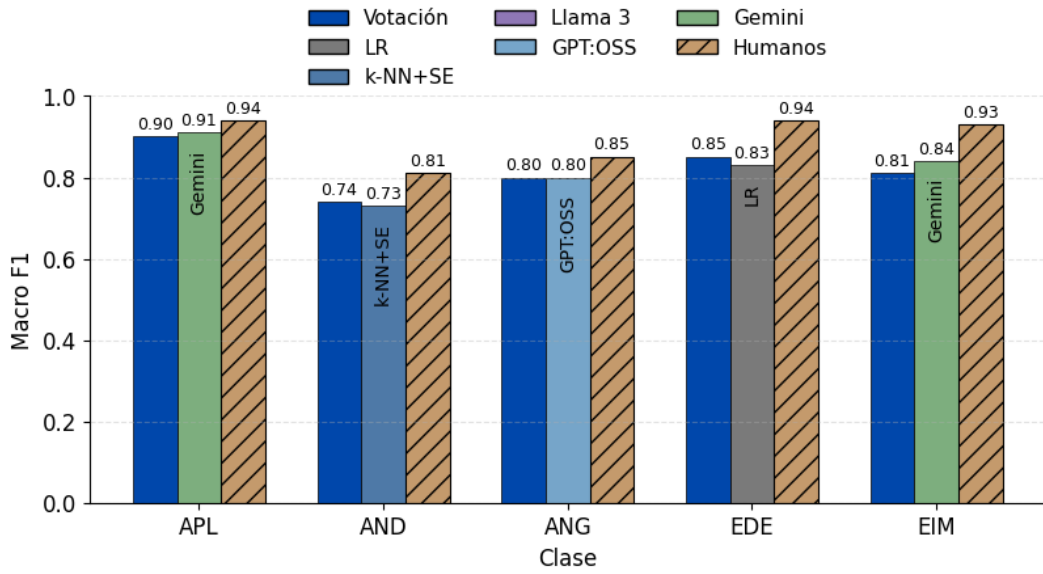


Figura 5.8: Comparación del *Macro F1-score* por clase en el corpus de evaluación entre la votación por mayoría simple, el modelo con mejor desempeño en cada clase y el acuerdo entre dos anotadores humanos, calculado sobre una muestra de 315 enunciados clasificados independientemente.

La [Figura 5.8](#) también muestra que las clases APL y EDE se destacan dentro del esquema de votación, aunque por razones diferentes. En APL, la clasificación automática alcanza el valor más próximo a la métrica de referencia estimada a partir del acuerdo entre evaluadores humanos, mientras que en EDE la votación por mayoría simple muestra un comportamiento más robusto al superar el desempeño del mejor modelo individual. Considerando estos resultados, en las secciones siguientes se analiza el tipo de errores cometidos por los clasificadores en el corpus de evaluación para ambas clases, distinguiendo entre falsos positivos y falsos negativos. La selección de APL y EDE responde además, a que remiten a dimensiones estrechamente vinculadas del trabajo experimental: por un lado, APL reúne enunciados asociados al desarrollo de habilidades prácticas de laboratorio, tales como el montaje experimental, la

utilización de instrumentos y su calibración, y por otra parte, EDE corresponde a un proceso cognitivo de mayor complejidad, que involucra la evaluación del procedimiento experimental, de los instrumentos empleados o de las técnicas de medición utilizadas.

5.3. Análisis de errores en la clasificación automática.

El análisis de las métricas permitió identificar aquellas clases para las que los clasificadores presentan un rendimiento más alto o más bajo, pero estas métricas no explican las razones por las cuales ciertos enunciados son clasificados por los diferentes algoritmos de manera distinta a como lo haría un experto. Para comprender estas discrepancias fue necesario revisar los enunciados a los que se les asignó la clase cuando no correspondía, denominados *falsos positivos* (FP), así como aquellos que fueron excluidos de la clase a pesar de corresponder a ella, es decir, los *falsos negativos* (FN).

A continuación se llevará a cabo el análisis de errores para las clases APL y EDE, asociadas a los procesos cognitivos *Aplicar* y *Evaluar* de la TBA. La selección de estas clases responde, por un lado, a que se encuentran entre aquellas en las que, en términos generales, los clasificadores de AA alcanzan mejores desempeños y, por otro, a que representan el proceso cognitivo de menor y de mayor nivel jerárquico dentro de los tres procesos cognitivos considerados en este trabajo. Además, ambas remiten a dimensiones estrechamente vinculadas del trabajo experimental. Con este fin, se extrajeron todos los enunciados identificados como FN y FP en cada clase. Durante el análisis de FP y FN, los enunciados fueron leídos y agrupados manualmente en áreas temáticas. La identificación de estas áreas se realizó con asistencia de GPT 5.2 Thinking (OpenAI, 2025a), Gemini 3 en su versión de razonamiento (Google, 2026a) y NotebookLM (Google, 2026b).

5.3.1. Análisis de errores sobre el corpus de evaluación para la clase APL.

La clasificación automática aplicada sobre el corpus de evaluación permitió identificar, para cada modelo, la cantidad de enunciados clasificados correcta

o incorrectamente. Para la clase APL, asociada al proceso cognitivo *Aplicar* y al desarrollo de habilidades de medición en el laboratorio (véase [Tabla 4.3](#)), la [Tabla 5.3](#) presenta la cantidad de enunciados clasificados incorrectamente por cada modelo. A partir de estos resultados, se observa que no existen grandes diferencias entre los clasificadores respecto de la cantidad total de errores. Dentro de esta comparación, GPT presenta el mayor número de errores, mientras que *Sentence Embeddings* combinado con k -NN registra el menor.

Tabla 5.3: Total de enunciados falsos positivos y negativos en la clasificación del corpus de evaluación, por modelo, para la clase APL. La suma de los falsos negativos y positivos no coincide con la cantidad de enunciados clasificados por cada modelo dado que la clasificación se realiza de forma binaria e independiente.

Modelo	Cantidad
Gemini 2.5 Flash Lite	18
GPT:OSS 20B	25
Llama 3 8B Instruct	17
Regresión logística	19
Sentence Embeddings & k -NN	16
Falso negativo	21
Falso positivo	20

El análisis de los enunciados correspondientes a la clase APL permitió identificar 21 FN y 20 FP. En ambos grupos, 7 enunciados presentaron coincidencia entre tres o más clasificadores automáticos. Por lo tanto, de acuerdo con el sistema de votación por mayoría simple, esos enunciados serían efectivamente considerados errores de clasificación.

El análisis de los **enunciados FN** permitió distinguir tres situaciones recurrentes en las que los clasificadores tendieron a excluir enunciados que, de acuerdo con la clasificación humana, pertenecían a la clase. En mayor proporción, se identificaron enunciados que *justificaban la toma de medidas* a partir de dos tipos de fundamentos: las características de los instrumentos o materiales que integraban el dispositivo experimental, y decisiones metodológicas sobre el procedimiento sustentadas en consideraciones teóricas. A modo de ejemplo puede leerse el siguiente enunciado:

“La cantidad de medidas tomadas es la mínima que permite obtener una calibración relativamente buena, y los valores de temperatura en los que se midió la resistencia dentro del intervalo, se determi-

naron teniendo en cuenta el comportamiento de cada dispositivo a calibrar para lograr mediciones significativas.”

En el enunciado previo se observa que los estudiantes desarrollan habilidades vinculadas con la medición, en tanto describen las condiciones necesarias para obtener mediciones coherentes. Si bien no se mencionan instrumentos específicos, el enunciado refiere explícitamente a las magnitudes medidas, lo que evidencia un criterio de control de las condiciones de medición.

Entre los enunciados FN, con una frecuencia similar al caso anterior, se identificaron enunciados que describen la *configuración del dispositivo experimental*, en algunos casos remitiendo explícitamente a una figura o descripción del montaje, como ejemplo puede verse el siguiente enunciado:

“Para esta práctica se diseña un dispositivo como se muestra en la Figura 2, el cual permite estudiar el movimiento de un objeto cuando se desplaza por un plano inclinado.”

En este enunciado no se detallan los instrumentos, montaje o procedimiento, sin embargo, la referencia a una figura que ilustra la configuración permite asumir que los estudiantes realizaron el armado correspondiente. Por esta razón, se entiende que el enunciado sugiere de manera indirecta el desarrollo de habilidades prácticas y manuales asociadas a la preparación del dispositivo experimental.

La tercera situación en la que los enunciados son excluidos de la clase APL por los clasificadores automáticos se observa cuando estos refieren al *registro o la toma de medidas*, que implica el uso de instrumentos de medición aunque no se haya indicado explícitamente en el enunciado el modo en que dichas acciones se llevaron a cabo. A modo de ejemplo:

“Para esto se realizó la práctica de registrar el periodo de oscilación de varios ángulos.”

Por otra parte, el análisis de los **enunciados FP** revela que, la mayor proporción de estos enunciados refieren a *mejoras a futuro o limitaciones en el dispositivo experimental o procedimiento*, en algún caso con mención de los instrumentos de medida utilizado, como por ejemplo:

“Para mejorar la medición se concluye que el uso de un termómetro digital sería más preciso.”

En segundo lugar, se encontraron enunciados FP que hacen *énfasis en el registro de medidas con mención a algún instrumento o programa informático*, la sola mención de éstos no asegura el desarrollo de habilidades practicas propias del trabajo en laboratorio como puede observarse en el siguiente ejemplo:

“Siendo R_0 el valor de la resistencia eléctrica del termistor medido a la temperatura inicial del agua T_0 .”

Este análisis también permitió identificar cinco enunciados que los clasificadores automáticos asignaron a la clase APL, pero que no fueron incluidos en dicha clase en la codificación humana. La relectura de estos casos sugiere la existencia de errores de clasificación también en el proceso manual, probablemente asociados al volumen de enunciados que deben ser leídos y clasificados. En particular, se observó que estos enunciados debieron haber sido incorporados originalmente a la clase APL, por lo que su clasificación automática es la esperada.

El resto de los enunciados identificados como falsos positivos (FP) corresponde a situaciones menos frecuentes y de carácter aislado, por lo que no configura un patrón reconocible.

5.3.2. Análisis de errores sobre el corpus de evaluación para la clase EDE.

El análisis de los falsos negativos (FN) y falsos positivos (FP) de la clase EDE permitió identificar y agrupar los casos en que la clasificación de los modelos de AA no coinciden con la clasificación humana. Del total de los enunciados identificados, 24 son FN y 55 FP.

El estudio de los **enunciados FN** revela que los errores de clasificación se concentran en formulaciones que describen *limitaciones del dispositivo experimental* o que explicitan cómo el *diseño, el control y/o la ejecución del experimento inciden en los resultados*.

En el primer caso, se observan enunciados que reconocen restricciones asociadas al montaje, los materiales o los instrumentos, por ejemplo:

“[...] se ven afectados por el rozamiento del aire, imperfecciones en la tanza que generan fricción y el hecho de que la masa no sea puntual.”

Tabla 5.4: Total de enunciados del corpus de evaluación clasificados incorrectamente por cada modelo para la clase EDE. El total de errores corresponde a la cantidad de enunciados identificados como falsos positivos o falsos negativos. Este valor no coincide con la cantidad total de enunciados evaluados, ya que cada modelo clasifica todos los enunciados de forma binaria e independiente, según pertenezcan o no a la clase EDE.

Modelo	Cantidad
Gemini 2.5 Flash Lite	24
GPT OSS 20B	28
Llama 3 8B Instruct	46
Regresión logística	20
Sentence Embeddings & k-NN	21
Falso negativo	24
Falso positivo	55

“[...] no dejar al metal el tiempo necesario en el baño térmico para que alcance su temperatura, o quizás el tiempo en contacto con el agua del calorímetro no haya sido el suficiente para alcanzar la temperatura de equilibrio.”

Por otra parte, también se identifican enunciados que evalúan el dispositivo y/o el diseño experimental con foco en la *ejecución del procedimiento*, como el siguiente:

“Al realizar la práctica se colocó el péndulo levemente torcido respecto al eje del fotosensor y eso llevó a la obtención de datos erróneos al final de los 30 minutos”

En menor proporción aparecen evaluaciones generales en las que se *reconoce la presencia de errores, pero sin identificar su causa* o el aspecto específico del diseño o del procedimiento involucrado, por ejemplo:

“Esto debe ser a causa de algún mínimo error en la práctica, ya que en el gráfico se presentaba un poco de batido.”

Se encontró que, de los 24 enunciados identificados como FN, en 10 de ellos existió coincidencia entre tres o más modelos, esto implica que, tras aplicar el proceso de votación, esos casos continuarían siendo clasificados como FN.

Por otra parte, se identificaron 55 enunciados FP, para los cuales predominan motivos distintos a los observados en los FN. El análisis de contenido

permitió reconocer, en orden de frecuencia, los siguientes conjuntos: (i) enunciados que *evalúan o emiten conclusiones acerca de los resultados* más que el dispositivo o el procedimiento experimental; (ii) enunciados que *describen el montaje y/o el procedimiento* sin realizar una evaluación de instrumentos, limitaciones, diseño o ejecución de la práctica; y (iii) enunciados centrados en el *análisis de datos y/o gráficos*.

A continuación se presentan enunciados ilustrativos. En cada caso se indica el conjunto al que pertenece el enunciado de la clase FP:

Evaluación de resultados – “*De esta manera se logró calcular ambos coeficientes de amortiguamiento viscoso con una buena precisión en función de los datos obtenidos*”

Descripción del montaje/procedimiento – “*La configuración experimental fue lograda, como muestra la figura 1, mediante un baño térmico y dos vasos térmicos, estos últimos actúan como calorímetros.*”

Análisis de datos – “*Para concluir, a pesar de que el caso del cronómetro fue el único que tuvo concordancia en el valor de la aceleración gravitatoria con el valor de referencia, notamos que los valores de los otros dos casos no estaban muy alejados del mismo, y sus porcentajes de error no fueron altos, por lo que, los consideramos como valores aceptables.*”

Es importante señalar que, de los 55 enunciados clasificados como FP, sólo 8 presentan coincidencia entre tres o más modelos de AA. Esto sugiere que, al aplicar un esquema de votación por mayoría simple, únicamente esos 8 enunciados serían finalmente clasificados como pertenecientes a la clase EDE, aun cuando constituyen FP. Este resultado contribuye a explicar el aumento en el valor de la métrica *Macro F1-score* para esta clase tras el proceso de votación.

Además, se puede observar que la proporción de FP es aproximadamente el doble que la de FN. Tal como se observa en la [Tabla 5.4](#), el modelo Llama3 identifica más del doble de enunciados como pertenecientes a la clase EDE en comparación con el resto. En consecuencia, se generan 23 FP que no coinciden con la clasificación de ningún otro modelo. En su mayoría, estos enunciados *evalúan o informan resultados* en lugar de evaluar el dispositivo o el procedimiento experimental, o bien se centran en el *análisis de datos y/o gráficos*.

5.4. Síntesis del capítulo.

Los resultados presentados muestran que es posible abordar la identificación automática de habilidades científicas en informes de laboratorio mediante la combinación de distintas técnicas de PLN. Los desempeños alcanzados por los diferentes clasificadores varían según la clase que se pretende identificar, por lo que no es posible establecer una única elección del modelo más adecuado para todas las clases. Estas diferencias se abordaron mediante la incorporación de un esquema de votación por mayoría simple, orientado a contemplar las particularidades para la clasificación propias de los distintos modelos.

El análisis de los resultados de la clasificación automática permitió ir más allá de la comparación entre métricas. La revisión de los errores puso en evidencia que las dificultades no dependen únicamente del algoritmo empleado, sino también de la forma en que determinadas habilidades quedan expresadas en la escritura estudiantil. En particular, cuando los enunciados describen acciones de manera indirecta, proyectan mejoras a futuro o presentan formulaciones ambiguas entre descripción, análisis y evaluación, aumentan las posibilidades de desacuerdo entre clasificadores automáticos e incluso respecto de la clasificación humana. De esta forma, los hallazgos obtenidos permiten valorar tanto el alcance de los modelos implementados como los posibles errores de clasificación al abordar, en los próximos capítulos, el análisis de los informes del semestre.

Capítulo 6

Objetivos y habilidades: entre el diseño y la percepción.

El curso Física Experimental 1, foco de este trabajo, se desarrolla en el laboratorio con grupos de entre 12 y 15 estudiantes, quienes asisten quincenalmente a una instancia presencial de tres horas. Entre una clase y la siguiente, la plataforma educativa les permite el acceso a materiales de consulta en diversos formatos para preparar el trabajo experimental, así como con un espacio de consultas a docentes e intercambio entre estudiantes. Durante las clases de laboratorio trabajan preferentemente en parejas o tríos para desarrollar la actividad experimental con la orientación del docente. Existen cuatro temas en los que se hace énfasis durante el curso: mediciones e incertidumbre, modelado de problemas físicos, tratamiento estadístico de datos y método de mínimos cuadrados. Las prácticas realizadas también abordan los contenidos de los programas de las unidades curriculares Física 1 y 2.

Tabla 6.1: Objetivos del curso Física Experimental 1 extraídos del Programa del Curso (véase Anexo 4).

Objetivo	Descripción
1	Reconocer las magnitudes físicas a medir y seleccionar los parámetros asociados al sistema físico de interés
2	Implementar el montaje experimental siguiendo los lineamientos propuestos.
3	Aplicar una metodología para medir las magnitudes físicas y parámetros de la experiencia.
4	Analizar los puntos débiles en el montaje experimental y/o el procedimiento y definir mejoras para los mismos.
5	Identificar y analizar las diversas fuentes de incertidumbre en la realización de medidas de magnitudes físicas, y evaluar su impacto en el resultado.
6	Utilizar herramientas de análisis de datos experimentales, en particular las estadísticas y de ajustes por mínimos cuadrados e interpretar los resultados obtenidos al aplicarlas.
7	Reconocer los distintos instrumentos de medición, recordar sus características de funcionamiento y utilizarlos en la medida de cantidades físicas en el procedimiento experimental.
8	Realizar y analizar gráficos de medidas experimentales y de modelos teóricos.
9	Analizar los resultados experimentales e interpretarlos en comparación con el modelo físico propuesto, y obtener conclusiones sobre el diseño de la experiencia y sus resultados.
10	Presentar resultados de forma escrita en un formato preestablecido, con varias instancias a lo largo del curso para mejorar y reforzar los aprendizajes.
11	Elaborar una presentación y exponer de forma oral los resultados obtenidos en una de las experiencias.

En la [Tabla 6.1](#) se han extraído los objetivos del curso que se encuentran en el programa (ver anexo 4). Los mismos se han desglosado para su clasificación según los procesos cognitivos y dimensiones del conocimiento de la TBA. La [Figura 6.1](#) muestra los resultados de la clasificación de estos objetivos del curso de acuerdo con la TBA y en ella se evidencia que la mayor parte de los objetivos del curso se ubica los procesos cognitivos de las categorías *Aplicar*, *Analizar* y *Evaluar*, distribuidos mayormente entre el *conocimiento conceptual* y el *conocimiento procedimental*. Esta distribución resulta coherente con las clases (APL, ANG, AND, EIM y EDE) seleccionadas para el análisis de los informes de laboratorio, las cuales se asocian a los mismos procesos cognitivos.

Respecto de esta clasificación, es pertinente introducir algunas precisiones. El objetivo 1, “*seleccionar parámetros asociados al sistema físico de interés*”, podría considerarse inicialmente dentro de la categoría *Aplicar*, sin embargo, dado que en este contexto implica distinguir lo relevante de lo irrelevante, se

lo ubicó en la categoría *Analizar*.

DIMENSIÓN DEL CONOCIMIENTO	DIMENSIÓN DEL PROCESO COGNITIVO					
	1. RECORDAR	2. COMPRENDER	3. APLICAR	4. ANALIZAR	5. EVALUAR	6. CREAR
A. CONOCIMIENTO FÁCTICO	Reconocer las magnitudes físicas (1). Reconocer instrumentos de medición y recordar las características de su funcionamiento (7).					
	Identificar diversas fuentes de incertidumbre (5).	Interpretar los resultados obtenidos a partir del uso de herramientas estadísticas (6).	Utilizar herramientas de análisis de datos experimentales (6).	Seleccionar parámetros asociados al sistema físico de interés (1). Analizar las fuentes de incertidumbre en la realización de medidas de magnitudes físicas (5). Analizar gráficos de medidas experimentales y modelos teóricos (8). Analizar los resultados experimentales e interpretarlos en comparación con el modelo físico propuesto (9).	Evaluar el impacto de la incertidumbre en la medición (5). Obtener conclusiones sobre sus resultados (9).	
B. CONOCIMIENTO CONCEPTUAL						
C. CONOCIMIENTO PROCEDIMENTAL			Implementar el montaje experimental siguiendo los lineamientos propuestos (2). Aplicar una metodología para medir las magnitudes Físicas y parámetros de la experiencia (3). Utilizar diferentes instrumentos de medición (7). Realizar gráficos (8). Presentar resultados de forma escrita en un formato preestablecido , con varias instancias a lo largo del curso para mejorar y reforzar los aprendizajes (10).	Analizar los puntos débiles del montaje experimental y/o el procedimiento (4).	Definir mejoras para el montaje experimental y/o el procedimiento (4). Obtener conclusiones sobre el diseño de la experiencia (9).	Elaborar una presentación y exponer de forma oral los resultados obtenidos en una de las experiencias (11).
D. CONOCIMIENTO METACOGNITIVO						

Figura 6.1: Clasificación de los objetivos del curso Física Experimental 1 en la tabla de la Taxonomía de Bloom revisada por Anderson y Krathwohl (2001) (TBA). Entre paréntesis se ha hecho referencia al objetivo del curso indicado en la [Tabla 6.1](#). Fuente: Elaboración propia.

Por su parte, el objetivo 10, “*presentar resultados de forma escrita en un formato preestablecido, con varias instancias a lo largo del curso para mejorar y reforzar los aprendizajes*”, puede resultar ambiguo al momento de clasificarlo. La referencia a instancias de mejora podría sugerir que se espera que los estudiantes sean más conscientes y responsables de su propio aprendizaje. Sin embargo, la mejora derivada únicamente de la retroalimentación no garantiza por sí misma, un componente metacognitivo. Para ubicar este objetivo dentro del conocimiento metacognitivo sería necesario explicitar acciones como justificar cambios realizados, identificar debilidades propias, reflexionar sobre el proceso experimental y/o escritura, o explicar cómo la retroalimentación le permitió mejorar su trabajo.

Por otra parte, el segmento del objetivo 4 que refiere a “*Definir mejoras para el montaje experimental y/o el procedimiento*” se ha clasificado a nivel del proceso cognitivo *Evaluar*, puesto que implica juzgar con cierto criterio previamente a proponer alternativas o mejoras. Esto se diferencia de “*Crear*” que implicaría el rediseño del experimento.

Del análisis de los objetivos del curso también se desprende que la *calibración de instrumentos* no se explicita como objetivo. Esto es relevante al considerar que dicha calibración forma parte de los objetivos de la Práctica 2 (véase Anexo 3). Sin embargo, podría interpretarse como implícita en objetivo 7, que refiere a lograr “*utilizar diferentes instrumentos de medición*”.

Tabla 6.2: Objetivos del curso agrupados por proceso cognitivo (TBA) y clase.

Proceso cognitivo	Clase	Objetivos (números de la Tabla 6.1)
Aplicar	APL	2, 3, 7.
Analizar	ANG	8, 9.
	AND	1, 9.
Evaluar	EDE	4, 9.
	EIM	5.

Las clases establecidas en la guía de anotación para el análisis de texto mediante aprendizaje automático abarcan un subconjunto de los objetivos del curso. Tal como se observa en la [Tabla 6.2](#), ocho de los once objetivos quedaron contemplados en las clases definidas.

Por otra parte, con el objetivo de relevar la percepción de los docentes

del curso acerca de la importancia atribuida a sus objetivos, se aplicó una encuesta en formato electrónico (véase [Apéndice H](#)) dirigida a docentes del Instituto de Física de la Facultad de Ingeniería que, en alguna oportunidad, han dictado el curso Física Experimental 1. La encuesta fue respondida por 10 docentes. De ellos, el 70 % manifestó haber dictado el curso en más de cuatro oportunidades, mientras que el 20 % indicó haberlo dictado entre dos y tres veces. Un docente, por su parte, reportó haberlo dictado en una única ocasión. Además, solo cuatro de los participantes integraron el equipo docente del semestre 2024, correspondiente al período del cual se extrajeron los informes analizados. Los restantes docentes lo dictaron en otros semestres.

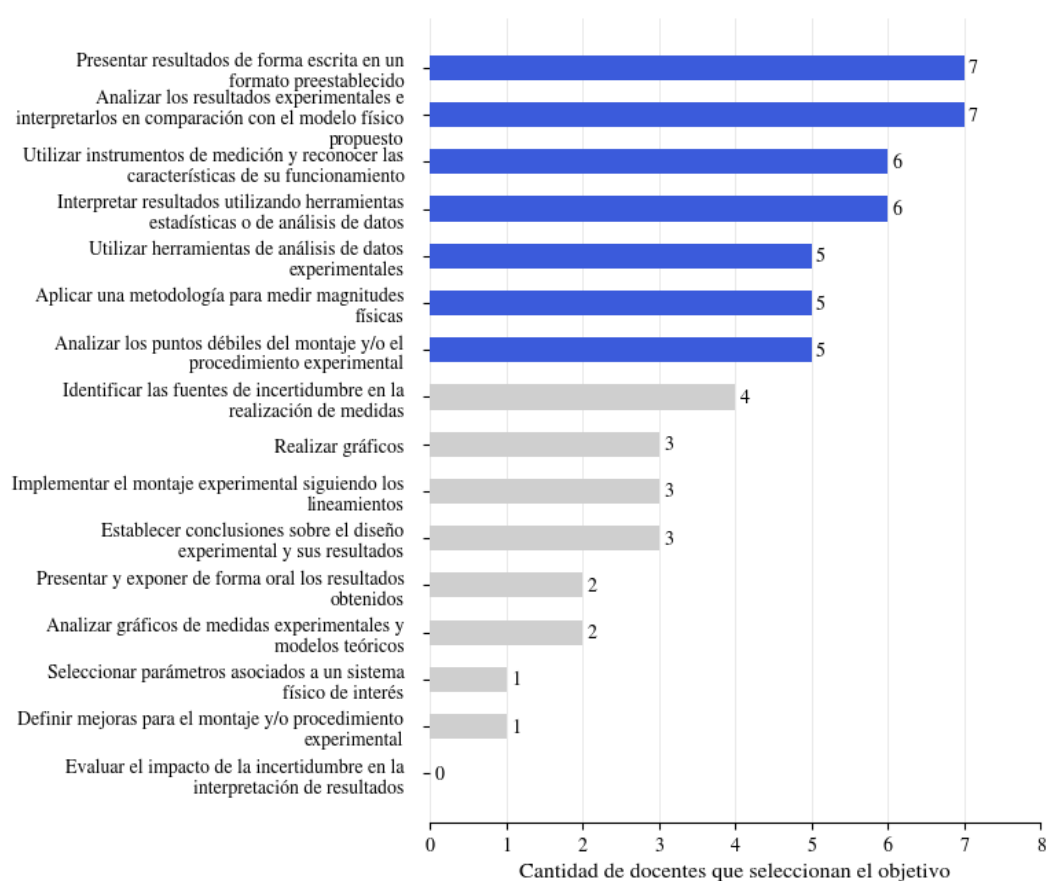


Figura 6.2: Frecuencia de selección de los objetivos identificados por los docentes como los más importantes para un curso de Física Experimental 1. Los objetivos se presentan ordenados según su frecuencia de selección. El eje horizontal indica la cantidad de docentes que seleccionó cada objetivo y el eje vertical presenta los objetivos.

La percepción docente sobre los objetivos principales del curso muestra que existe un conjunto de estos que son considerados prioritarios por al menos el

50 % de los encuestados. Este grupo reúne objetivos vinculados con el desarrollo de habilidades prácticas propias de la medición en laboratorio, el análisis de datos experimentales y el análisis del montaje y/o del procedimiento experimental. De acuerdo con la clasificación presentada en la [Figura 6.1](#), estos objetivos se ubican en los procesos cognitivos *Comprender*, *Aplicar* y *Analizar*. Además de estos objetivos, se destaca la importancia asignada a que los estudiantes elaboren un informe de laboratorio siguiendo un formato preestablecido.

Resulta destacada la escasa selección docente, de objetivos asociados al nivel *Evaluar*, como *definir mejoras para el montaje y/o procedimiento experimental* y *evaluar el impacto de la incertidumbre en la interpretación de resultados*.

La consulta a docentes acerca de la necesidad de incorporar algún otro objetivo de aprendizaje reveló aportes diferentes. Hay quienes señalaron el solapamiento entre objetivos del programa. Esto puede explicarse, por ejemplo, al leer los objetivos *Analizar gráficos de medidas experimentales y modelos teóricos* y *Analizar los resultados experimentales e interpretarlos en comparación con el modelo físico propuesto*. Si bien el segundo tiene un alcance más amplio, dado que refiere al análisis de resultados experimentales en general, presentados en formatos como tablas, gráficos o resultados numéricos, el primero se acota al tratamiento de resultados representados gráficamente. Esto puede llevar a que ambos objetivos puedan percibirse como parcialmente superpuestos, pero cabe señalar que esta segmentación de objetivos proviene de la formulación de los mismos en el programa del curso. También es posible que esta percepción de superposición se relacione con la falta de distinción entre distintos niveles de complejidad en los procesos cognitivos involucrados. Por ejemplo, *realizar gráficos* supone un proceso menos complejo que *analizar gráficos*. Si esta diferencia no se explicita o no se considera al momento de responder, ambos objetivos pueden llegar a interpretarse como similares.

Otro aporte sugirió incluir entre los objetivos el “*aprender a trabajar en equipo*”. El laboratorio es un espacio importante para fomentar o reforzar las habilidades de trabajo en equipo y colaboración, y forma parte de las recomendaciones de la AAPT (Kozminski et al., [2014](#)). Si bien este objetivo no figura explícitamente en el programa del curso, la metodología de enseñanza favorece el trabajo en equipo tanto en las instancias de laboratorio, desarrolladas en parejas o tríos, como durante la elaboración de los informes. Esto puede con-

tribuir al desarrollo de habilidades vinculadas al trabajo en equipo, aunque no garantiza por sí solo que dichas habilidades se desarrollen efectivamente.

Con el propósito de identificar las habilidades que se busca desarrollar en el curso de Física Experimental 1, se elaboró una lista basada en las recomendaciones de la AAPT (Kozminski et al., 2014), organizada según los procesos cognitivos de la TBA. En la encuesta electrónica dirigida a docentes, ya descrita anteriormente, y en otra respondida por veintiocho estudiantes (véase [Apéndice I](#)), se solicitó a los participantes que seleccionaran las cinco habilidades que, a su juicio, se priorizan en el curso. Para adecuarlas al formato del instrumento, las habilidades se presentaron en una versión abreviada. Para la aplicación de la encuesta a estudiantes, se contó con la colaboración de los docentes del curso y, adicionalmente, esta se publicó en EVA (Entorno Virtual de Aprendizaje), la plataforma virtual de apoyo al curso.

Los resultados de la [Figura 6.3](#) evidencian que, si bien estudiantes y docentes coinciden en la importancia atribuida a algunas de las habilidades, también se advierten algunas diferencias significativas. La mayor diferencia se presenta en la habilidad asociada a *evaluar cómo la incertidumbre influye en la validez de los resultados*, la cual es seleccionada con mayor frecuencia por los docentes que por los estudiantes, lo que sugiere una distinta percepción respecto del rol de la incertidumbre en la validez y calidad de la medición. Esta diferencia también resulta llamativa porque dicha habilidad se encuentra estrechamente vinculada con el objetivo 5 de la [Tabla 6.1](#). Sin embargo, al consultar a los docentes sobre la priorización de este objetivo, como se muestra en la [Figura 6.2](#), se observa que ninguno lo seleccionó. Esta discrepancia constituye un aspecto relevante para futuras investigaciones, posiblemente mediante entrevistas a docentes.

Los resultados de esta encuesta también permiten observar que los docentes asignan mayor relevancia a habilidades asociadas con la instrumentación y el análisis del trabajo experimental, mientras que los estudiantes priorizan aspectos vinculados con el marco conceptual y el tratamiento cuantitativo de la incertidumbre. Estas diferencias pueden reflejarse en la forma en que los estudiantes elaboran sus informes de laboratorio, condicionando qué aspectos priorizan en la redacción, qué consideran evidencia relevante, qué justifican y qué reportan con mayor detalle. También cabe preguntarse si esta percepción del estudiantado está mediada por el tipo de retroalimentación recibida durante la corrección de los informes. Es posible que las habilidades sobre las

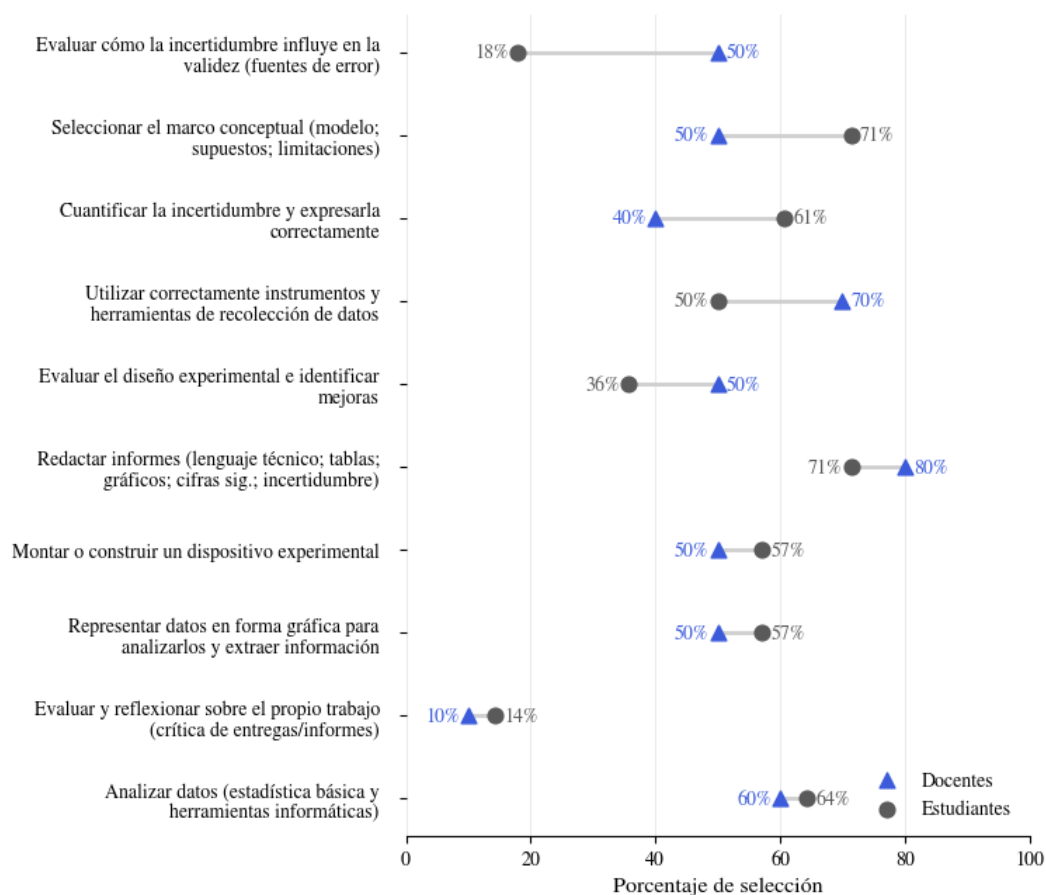


Figura 6.3: Comparación entre docentes y estudiantes acerca de la prioridad en las habilidades a desarrollar en el curso Física Experimental I. El gráfico compara el porcentaje de selección de cada habilidad para docentes (azul, $n = 10$) y estudiantes (gris, $n = 28$).

que se formulan más observaciones o correcciones tiendan a ser consideradas prioritarias, mientras que aquellas que reciben menos comentarios se perciban como secundarias.

En esta línea, resulta importante mencionar que la habilidad de *evaluar y reflexionar críticamente sobre el propio trabajo* ha sido considerada de baja prioridad por estudiantes y docentes, lo cual es coherente con la clasificación de objetivos del curso realizada en la [Figura 6.1](#) y que sugiere un área de mejora.

En la encuesta dirigida a docentes también se indagó si, a su juicio, las herramientas de evaluación implementadas en el curso se encuentran alineadas con los objetivos de aprendizaje propuestos. Dado el carácter abierto de la pregunta, las respuestas se analizaron mediante la identificación de tendencias generales. A partir de este análisis, se observó que poco más de la mitad de

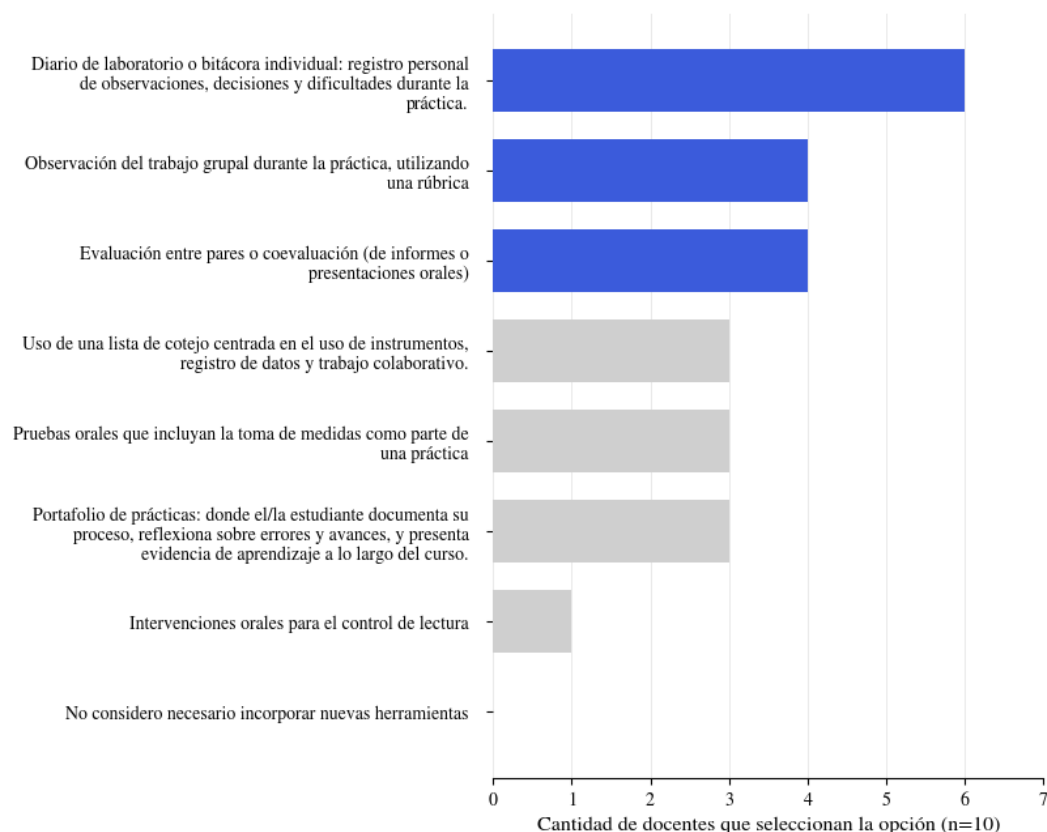


Figura 6.4: Distribución de respuestas de docentes a la pregunta sobre la incorporación de herramientas adicionales de evaluación en el curso de Física Experimental 1. La pregunta fue de *selección múltiple* sin límite en la cantidad de opciones que podían marcarse, por lo que se muestran las frecuencias que refieren al número de docentes que seleccionó cada alternativa.

los participantes considera que los informes de laboratorio, actualmente utilizados como herramienta de evaluación, constituyen la opción más viable en función de la modalidad del curso. Sin embargo, también plantea la necesidad de realizar ajustes y formula algunas sugerencias. Algunos docentes manifiestan la necesidad de homogeneizar criterios de evaluación, así como incorporar estrategias que permitan apreciar el proceso de trabajo colaborativo durante la elaboración de los informes, de modo de evitar la sobrecarga de tareas de alguno de los integrantes de los equipos de trabajo. Otros docentes señalan que el informe escrito tiende a privilegiar dimensiones asociadas a la redacción y al análisis de resultados, pero presenta limitaciones para evaluar el diseño experimental. Por otra parte, en menor proporción, un conjunto de respuestas considera que las herramientas de evaluación no están alineadas con los objetivos, ya sea por considerar que se sobrevalora el informe o porque considera que

las herramientas de evaluación disponibles requieren mejoras para que logren evidenciar los aprendizajes esperados.

En relación con la pregunta previa, se consultó a los docentes si consideraban necesario incorporar otras herramientas de evaluación brindándose las opciones que se muestran en la [Figura 6.4](#) y permitiendo agregar otras opciones. El resultado muestra la incorporación de un diario de laboratorio al curso podría ser de ayuda para la evaluación, y en menor proporción, han seleccionado que sería oportuno incorporar rúbricas para la evaluación del trabajo grupal durante la práctica o actividades que impliquen la evaluación entre pares.

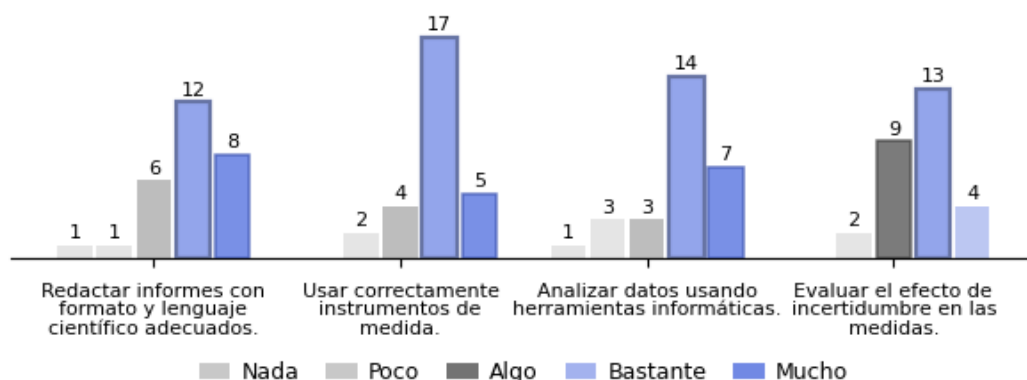


Figura 6.5: Distribución de respuestas sobre en qué medida se consideran distintas habilidades en la evaluación del curso. Pregunta de selección múltiple.

Otra de las preguntas dirigidas a los estudiantes indagó en qué medida perciben que las habilidades presentadas en la [Figura 6.5](#) son consideradas en la evaluación del curso. En tres de las habilidades consultadas, la mayoría de las respuestas, se concentran en las opciones *bastante* o *mucho*, lo que resulta alineado con la prioridad asignada a las habilidades a desarrollar. Las respuestas también sugieren que, para los estudiantes, el equipo docente valora en menor medida la *evaluación del efecto de la incertidumbre en la medición*. Esto se corresponde con lo observado en la [Figura 6.2](#) y en la [Figura 6.3](#), donde dicha habilidad no aparece como prioritaria ni para docentes ni para estudiantes, aunque entre los docentes fue seleccionada en mayor medida que el objetivo al que se encuentra vinculada.

La consulta sobre posibles cambios en la evaluación permitió recoger diversos aportes. Entre ellos, se destacó la incorporación de otros programas para

el tratamiento de datos, como por ejemplo Python, como una herramienta a considerar a futuro. También se sugirió que las instancias de evaluación incluyan componentes que permitan valorar el desempeño individual, dado que el trabajo colaborativo en la elaboración de informes no siempre se distribuye de manera equitativa entre los integrantes del grupo.

En conjunto, este capítulo permite advertir que, aunque los objetivos del curso, las habilidades y las formas de evaluación responden a una misma propuesta de enseñanza, su importancia no es interpretada del mismo modo por docentes y estudiantes. El análisis mostró diferencias entre las habilidades que los docentes consideran centrales y aquellas que el estudiantado identifica como más relevantes. Estos resultados dan lugar al capítulo siguiente, en el que se utiliza el aprendizaje automático para analizar los informes elaborados por los estudiantes a lo largo de todo el semestre e identificar en ellos los enunciados correspondientes a las clases definidas en la guía de anotación. Este abordaje permitirá examinar en qué medida las habilidades asociadas a las clases definidas se hacen presentes en la escritura estudiantil. En particular, para el análisis del capítulo siguiente se seleccionaron las clases APL y EDE.

Capítulo 7

Habilidades identificadas en el curso de física experimental.

En este capítulo se presenta la conformación del corpus con los enunciados extraídos del conjunto de informes del semestre, en particular de las secciones *Metodología*, *Análisis de datos* y *Conclusiones*. Sobre este corpus se aplicaron los cinco modelos de clasificación seleccionados para este trabajo, con el objetivo de identificar cuáles de esos enunciados podían ser asignados o no, a alguna de las cinco clases definidas en la guía de anotación. Posteriormente, las clasificaciones producidas por los modelos se integraron al esquema de votación por mayoría simple para determinar, en cada caso, la pertenencia o no de los enunciados a cada clase. Los resultados obtenidos mediante este procedimiento constituyen la base del análisis presentado en este capítulo. Luego, se profundiza el análisis de los resultados de clasificación para las clases APL y EDE mediante un análisis cualitativo de las habilidades que se evidencian en los informes de laboratorio. Este análisis busca relaciones entre las producciones escritas del estudiantado con las guías de trabajo de las prácticas de laboratorio, las orientaciones para la elaboración y evaluación de los informes y los objetivos del curso.

7.1. Los informes del curso: descripción y clasificación automática de enunciados.

Se recabaron 489 informes del curso Física Experimental 1 correspondientes al primer semestre de 2024. La distribución de informes por práctica se

presenta en la [Tabla 7.1](#). En ella se observa que la cantidad de informes no es la misma en las distintas prácticas, dado que la entrega se realiza en grupos. Además, pueden presentarse situaciones que, por diversas razones, impliquen la redistribución de integrantes entre equipos o la falta de entrega de algún informe, a pesar de que su presentación es obligatoria para la aprobación del curso.

Tabla 7.1: Cantidad de enunciados extraídos para el análisis de las secciones *Metodología*, *Análisis de datos* y *Conclusiones* correspondientes a informes del primer semestre de 2024.

N° Práctica	Informes	Enunciados
1	126	5095
2	124	5060
3	121	6022
4	118	4783

Para someter los enunciados a la clasificación mediante técnicas de PLN, los informes atravesaron una etapa de preprocesamiento. En primer lugar, los archivos PDF fueron convertidos a formato de texto plano (`.txt`) y anonimizados. Luego, se conservaron únicamente las secciones definidas para el análisis en este trabajo y se eliminaron los textos correspondientes a los pies de figuras y tablas, así como los desarrollos de cálculos. La extracción de enunciados por informe y sección se realizó de forma automática mediante un código desarrollado en Python. En esta etapa también se excluyeron los enunciados de menos de cinco palabras, ya que, por su contenido, difícilmente aportaran información relevante para el análisis, entre ellos se encontraban expresiones del tipo “Despejando de la ecuación XX:”. No obstante, estos enunciados se almacenaron en un archivo aparte y fueron revisados manualmente para determinar si correspondía su incorporación al corpus.

La [Tabla 7.2](#) presenta la cantidad de enunciados clasificados automáticamente según clase y práctica¹, mientras que la [Figura 7.1](#) muestra el porcentaje que estos representan sobre el total de enunciados de cada práctica. A partir de esta figura, se observa que la clase AND concentra la mayor proporción de enunciados en las cuatro prácticas. Además, no se identifica una variación

¹El desglose de los resultados de la votación puede consultarse en el repositorio <https://github.com/lspm1102/Metodologia-automatizada-analisis-informes-laboratorio.git>

Tabla 7.2: Cantidad de enunciados clasificados con AA por práctica y por clase.

Práctica	APL	ANG	AND	EIM	EDE
P1	1184	648	1450	778	605
P2	930	786	1859	499	455
P3	1104	937	1811	375	347
P4	821	370	1328	497	531
Total	4039	2741	6448	2149	1938

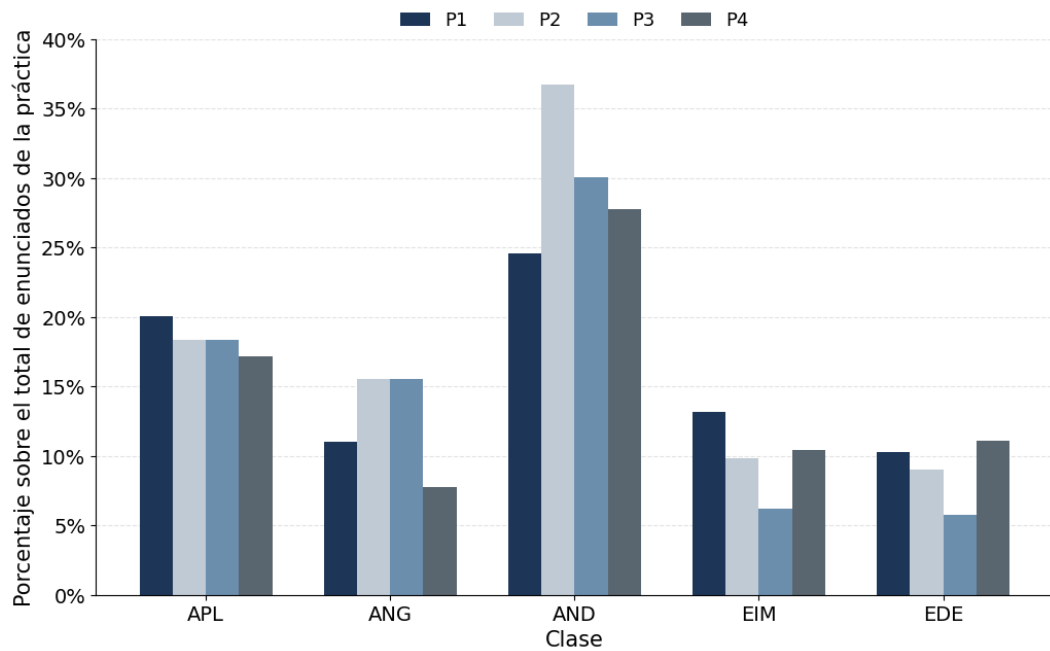


Figura 7.1: Comparación de la distribución porcentual de enunciados por clase y por práctica (indicadas con P1, P2, P3 y P4).

uniforme de los porcentajes entre prácticas para el conjunto de las clases, con excepción de APL, que presenta una leve tendencia descendente. Por otra parte, las clases EIM y EDE, vinculadas al proceso cognitivo *Evaluar* de la TBA, son las que registran las menores cantidades de enunciados en todas las prácticas. Estos resultados guardan coherencia con los observados en la [Tabla 4.5](#), correspondiente a la distribución de enunciados anotados por evaluadores humanos, donde también se advierte un predominio de la clase AND y una menor presencia de las clases EIM y EDE. La menor cantidad de enunciados correspondientes a estas clases, asociadas con procesos cognitivos de orden superior, podría vincularse con el hecho de que este tipo de enunciados se identifica

principalmente en la sección *Conclusiones* de los informes de laboratorio. Dado que esta sección suele tener una menor extensión que *Metodología y Análisis de datos*, es esperable que concentre una cantidad más reducida de enunciados de este tipo.

En suma, la distribución de enunciados por clase según la práctica considerada permite advertir diferencias, pero no una evolución progresiva claramente definida a lo largo del semestre. Este resultado sugiere que la presencia relativa de las distintas clases estaría más asociada a las particularidades de cada práctica. A partir de ello, resulta pertinente profundizar en el contenido de los enunciados que resultaron de esta clasificación.

7.2. Habilidades prácticas propias de la experimentación y medición en laboratorio (clase APL).

En esta sección se analizan los enunciados clasificados en la clase APL, vinculada al desarrollo de habilidades prácticas propias de la experimentación y la medición en laboratorio. Para ello, se considera el conjunto de enunciados de los informes correspondientes a las distintas prácticas del semestre que fueron clasificados de forma automática como pertenecientes a esta clase. A partir de estos enunciados, se busca reconocer las características de las habilidades se evidencian en los informes de los estudiantes. Con este fin, se examinan tanto su distribución como las temáticas que emergen en cada práctica, y se vinculan los resultados con las orientaciones de las guías de trabajo, los objetivos del curso y las recomendaciones de la AAPT.

Para la **Práctica N°1**, titulada *Determinación de la aceleración de la gravedad local* (véase Anexo 3) la clasificación automática permitió identificar enunciados pertenecientes a la clase APL y vinculados con la categoría *Aplicar* de la TBA. Esta clase está ligada al *desarrollo de habilidades propias de la experimentación y medición en laboratorio*. Se logró clasificar automáticamente 1184 enunciados pertenecientes a 126 informes de laboratorio. La mayoría de ellos (1054) pertenecía a la sección *Metodología*, mientras que 119 se localizaron en la sección *Análisis de datos* y 11 en las *Conclusiones*. A su vez, se observaron informes con un mínimo de 3 y un máximo de 19 enunciados de esta clase, obteniendo una mediana y un promedio de 9 enunciados por informe.

Un conteo de la frecuencias de palabras de los enunciados seleccionados mediante AA para la clase APL permitió llegar a la nube de palabras de la [Figura 7.2a](#), que muestra las 40 palabras más frecuentes. El mayor tamaño de las palabras representa la frecuencia con la que aparecen en los enunciados pertenecientes a esta clase. Se puede identificar que los enunciados tienen un fuerte vínculo con la medición, lo cual era esperado de acuerdo con la definición de la clase. De esta forma se identificaron 913 hallazgos de los lemas *medir* o *medición* y las siguientes palabras más frecuentes se corresponden con magnitudes que se midieron durante la práctica como *período* o *periodo* (545), *ángulo* (434), y *masa* (317). En tercer lugar en orden de frecuencia, se encuentran los instrumentos utilizados durante la práctica como el *cronómetro* (295) y el *fotosensor* (285).

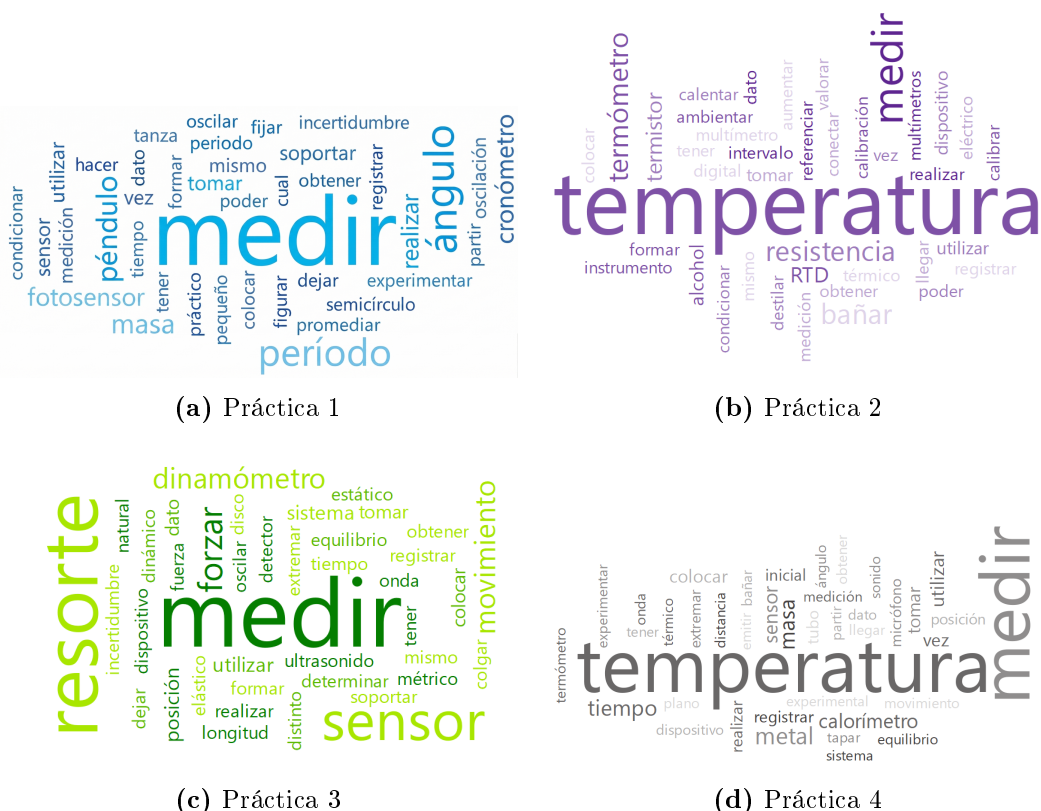


Figura 7.2: Nubes de palabras de las cuatro prácticas analizadas, con las 40 palabras más frecuentes según su lema.

La lectura de los enunciados permitió identificar que los mismos refieren a diferentes fases de la Práctica N°1 en la que se utiliza un péndulo simple. Las temáticas que pueden identificarse en los enunciados clasificados son las

siguientes:

- **Montaje y modelado experimental**, descripción del dispositivo y condiciones del modelo experimental. En esta temática se encuentran enunciados que refieren a la *importancia de buscar replicar las condiciones del modelo teórico* y se evidencia a través de expresiones como “tanza inextensible y sin masa”, “hilo de masa despreciable”, “seleccionamos la masa esférica para aproximarla a una masa puntual”, “se midió el largo del péndulo considerando el largo de la tanza y la medida del radio de la esfera (ya que el centro de masa de la esfera se encuentra en su centro geométrico)”. Otras expresiones como “se colocó el sensor en un mástil con forma de L”, “se sostuvo la masa con una superficie plana rígida intentando hacer el menor contacto posible con la ésta, hasta llevarla al ángulo de amplitud requerido”, muestran que los estudiantes son capaces de construir un dispositivo, nivelar, alinear, organizar, utilizar materiales adecuados incluso teniendo en cuenta el impacto de estas elecciones en la recolección de datos.
- **Descripción de los instrumentos utilizados y control de variables**. De acuerdo con la clase APL, se espera que los estudiantes sean capaces, como mínimo, de *utilizar instrumentos de medida simples*, como una balanza, un semicírculo o un cronómetro. Además, deberían ser capaces de *emplear herramientas de recolección de datos*, tales como videos, interfaces u otros programas informáticos, con el fin de extraer datos físicos. En esta práctica, estas habilidades se manifiestan en el uso del fotosensor y de una interfaz, en la correcta colocación del sensor y del cronómetro, en la medición de ángulos con el semicírculo y en el control de ciertas condiciones experimentales. Entre estas últimas, se destacan la necesidad de que la oscilación se mantenga en un único plano, que la masa sea soltada sin ejercer fuerza externa y que la misma persona registre tanto el ángulo como el tiempo con el cronómetro. Este tipo de acciones puede identificarse en expresiones como: “necesidad de que la oscilación se mantenga en un único plano”, “soltar la masa sin ejercer ningún tipo de fuerza externa” y “la persona que registró el ángulo y el tiempo con el cronómetro fue siempre la misma”.
- **Comparación de métodos e instrumentos de medición**. A esta temática pertenecen enunciados que establecen una *diferenciación en la*

toma de datos con diferentes instrumentos de medida, y que, dadas sus características, también refieren a la realización de mediciones simultáneas, sus ventajas, sus limitaciones y la incidencia del factor humano en la medición. Estas referencias pueden identificarse, por ejemplo, en segmentos de texto como: “se utilizan en simultáneo, dos elementos distintos para la medición: fotosensor y cronómetro” o “las medidas en cuestión fueron tomadas con el fotosensor y el cronómetro, ya que con ambas tenemos resultados ligeramente distintos, debido a la capacidad de reacción del humano”. De esta forma, en el texto no solo se comparan instrumentos y métodos, sino que también se integra la *reflexión sobre el factor humano en la medición* y se busca asegurar que las mediciones se realicen en iguales condiciones, lo que evidencia una reflexión sobre la *incidencia de los instrumentos y del operador en los resultados obtenidos*.

- **Aplicación de estrategias adecuadas para la toma de medidas.** Estos enunciados describen las estrategias para adquirir los datos como el barrido de ángulos, promediado de oscilaciones, cantidad de repeticiones, medición de larga duración, criterios para mejorar la medición. Por ejemplo, se encuentran segmentos como: “dejar que oscile el péndulo con un ángulo inicial de 10 ± 1 hasta obtener un total de 50 periodos medidos con fotosensor”, “se dejó caer el péndulo con distintos valores de θ inicial y se midió el tiempo que tardaba en completar 5 oscilaciones, así, el fotosensor promedió el período”, “se realizaron 30 repeticiones para luego hacer un análisis estadístico”, “se vuelve a medir el período de oscilación del sistema, pero dejándolo oscilar durante varios minutos, de manera de observar cómo va variando el período tras un largo lapso”. Estos enunciados muestran que existen toma de decisiones a la hora de experimentar, tanto para encontrar el ángulo óptimo como para buscar reducir el error experimental.
- **Justificación procedimental.** Incluye enunciados con justificación de las decisiones metodológicas con base en estadística y reducción de la incertidumbre. Se encuentran enunciados que contienen afirmaciones del tipo: “Con el fin de realizar un análisis estadístico que nos de información relevante, se decidió tomar un número grande de medidas del periodo” o “A la hora de la práctica, procuramos mantener las mismas condiciones tanto en la ejecución como en la medición, de esta manera, minimiza-

mos los errores humanos” o “Se mantendrá este montaje durante toda la práctica para asegurar la replicabilidad y la consistencia de los resultados entre sí”.

Para el desarrollo de esta primera práctica, los estudiantes disponen de una guía de trabajo experimental y, para la elaboración del informe, de una segunda guía (ambas pueden consultarse en el Anexo 3). En esta última se solicita que, en la sección *Metodología*, se incluyan los aspectos del procedimiento que se consideren más relevantes y se analicen las condiciones necesarias para satisfacer las hipótesis planteadas. Con el fin de orientar esta reflexión, se proponen, a modo de ejemplo, las siguientes preguntas:

1. ¿Por qué se deciden registrar un número grande de períodos en lugar de medir sólo un período?
2. ¿Por qué deben realizar el experimentos en iguales condiciones?
3. ¿Por qué deben medir tantas veces lo mismo?
4. Establezcan al menos una diferencia entre medir el tiempo con cronómetro y con sensor.

Estas indicaciones para la confección del informe están estrechamente ligadas a las áreas temáticas de los enunciados clasificados en forma automática. Los ítems 1, 2 y 3 con la temática *Diseño de la toma de medidas* y con la *Justificación procedimental*. El ítem 4 con la *Comparación de métodos e instrumentos de medición*.

Se seleccionó el ítem 1 y se buscó identificar los enunciados de la clase APL que hacen referencia a un número grande de mediciones o a la repetición de medidas, pero que además incluyen una justificación, explicación o motivo por el cual se llevan a cabo. Se realizó el análisis de los enunciados en forma asistida con los modelos online GPT 5.2 *Thinking* (OpenAI, 2025a), a Gemini en su modelo *Razonamiento 3 PRO* (Google, 2026a) y a Notebook LM (Google, 2026b) una herramienta de IA de Google que permite realizar el análisis de datos basado únicamente en la información que se le proporciona y en modelos de Gemini. A pesar de que los tres modelos recibieron la misma indicación, no identificaron la misma cantidad de enunciados. Notebook LM seleccionó 51 enunciados correspondientes a 47 informes, Gemini identificó 66 enunciados

correspondientes a 56 informes y GPT, fue más estricto en sus interpretación de la justificación y encontró 31 enunciados correspondientes a 30 informes. De esta forma se procedió a validar manualmente las selecciones realizadas por los modelos de IA. Se identificaron coincidencias entre los modelos y también enunciados que habían sido identificados únicamente por uno de los modelos. Además se eliminaron aquellos enunciados que se consideró que no debían incluirse, dado que no justificaban el motivo por el cual realizar un número grande de mediciones, en su lugar, incluían el por qué realizar las mediciones en iguales condiciones experimentales. En total, se identificaron 70 enunciados correspondientes a 62 informes (el 49 % de los informes analizados), todos pertenecientes a la sección *Metodología* de los informes con excepción de 4 enunciados.

En relación con las justificaciones ofrecidas por los estudiantes acerca de la necesidad de realizar múltiples mediciones, la [Tabla 7.3](#) muestra que la justificación más frecuente corresponde a aquellos enunciados que vinculan este procedimiento con la realización de un estudio estadístico de los datos, haciendo en ocasiones referencia a una distribución gaussiana. El estudio estadístico señalado por los estudiantes es, de hecho, uno de los objetivos mencionados en la guía de trabajo experimental: “*familiarizarnos con los procedimientos estadísticos para el tratamiento de datos*”.

Tabla 7.3: Justificaciones identificadas en los enunciados sobre la necesidad de realizar múltiples mediciones.

Justificación	Frecuencia
Realizar un estudio estadístico de los datos, en algunos casos con referencia a una distribución gaussiana.	28
Acercarse al valor real o al valor esperado.	12
Disminuir el error o la incertidumbre.	11
Mejorar la precisión o la exactitud.	10
Justificaciones varias que no pueden agruparse en una misma categoría.	9
Total	70

De los 70 enunciados identificados, 26 fueron detectados por uno de los tres modelos de IA usados, 29 por dos modelos y 15 por los tres modelos. Esto muestra que existe variabilidad en la interpretación de los textos por los diferentes modelos de IA.

La **Práctica N°2** del semestre, denominada *Caracterización de dispositi-*

vos de medición de temperatura (véase Anexo 3) tiene por objetivo comprender las características del funcionamiento de los dispositivos de medición de temperatura incluyendo la calibración del mismo. El análisis automático identificó 930 enunciados pertenecientes a la clase APL correspondientes a 124 informes. El promedio es de 7,5 enunciados por informe y la mediana es de 7, con un mínimo de 3 y un máximo de 16 enunciados por informe. Los enunciados pertenecientes a esta clase, al igual que los de la primera práctica, se concentran en la sección *Metodología* en un 91 %, un 7 % en la sección *Análisis de datos* y un 2 % en la *Conclusiones*.

Por otra parte, al realizar la nube de palabras de la Figura 7.2b se encontró dentro de las palabras más frecuentes aquellas que refieren a magnitudes físicas e instrumentos asociadas al proceso de medición en esta práctica como: temperatura (780 hallazgos), medir (501), resistencia (360), termómetro (358). También se identifican verbos como: *tomar*, *calentar*, *conectar*, *registrar*, *realizar* y *calibrar* que se vinculan con la categoría *Aplicar* de la TBA, y por lo tanto están alineados con la clase APL.

Al igual que para la Práctica N°1 se encontraron enunciados que refieren al montaje y modelado experimental, la descripción de instrumentos de medida y control de variables, diseño de la toma de medidas, justificación procedimental y la comparación de métodos y/o instrumentos de medición. A estas temáticas es necesario agregar la *calibración de instrumentos de medida*.

Se utilizaron los mismos tres modelos de IA empleados anteriormente (*GPT*, *Gemini* y *NotebookLM*) para identificar los enunciados que abordan la calibración de instrumentos para su uso, una de las habilidades a desarrollar de acuerdo con las recomendaciones de la AAPT (Kozminski et al., 2014). Se solicitó a los modelos que identificaran todos los enunciados que refirieran a calibrar, en cualquiera de sus conjugaciones, o a calibración, así como aquellos que aludieran a un instrumento patrón o de referencia, ya que una búsqueda previa había evidenciado el uso de expresiones como “instrumento patrón” o “temperatura de referencia” durante el proceso de calibración. *Gemini* identificó 183 enunciados correspondientes a 94 informes diferentes, *GPT* identificó 178 enunciados correspondientes a 92 informes diferentes y *NotebookLM* identificó 181 enunciados correspondientes a 91 informes diferentes. Tras unificar manualmente las respuestas de los tres modelos, se identificó un total de 177 enunciados correspondientes a 93 informes. En la Práctica N°2 se examinó una muestra de 124 informes, por lo que se obtuvo que el 79 % de ellos hace referen-

cia a la calibración de instrumentos. Si bien este porcentaje es elevado, cabría esperar uno mayor, dado que el objetivo central de la práctica es la calibración de dispositivos de medición de temperatura. Por otro lado, estos enunciados provienen de la clase APL, que fue la que registró las mejores métricas en cuanto a clasificación con AA. No obstante, es fundamental advertir que los resultados podrían estar condicionados por la metodología de clasificación empleada, ya que, aunque esta clase presenta las mejores métricas de desempeño, los clasificadores utilizados no logran recuperar la totalidad de los enunciados correspondientes a ella.

En relación con la calibración de instrumentos, los enunciados de la clase APL pueden agruparse según la forma en que hacen referencia a la calibración, de acuerdo con los siguiente enfoques:

- **Metodológico y procedimental:** se centran en las condiciones necesarias para que la calibración sea considerada *adecuada* o *válida*, incluyendo criterios como el número de mediciones y el rango de temperaturas.
- **Estándar de referencia:** aluden a la presencia de un instrumento patrón o de referencia indispensable para la calibración.
- **Configuración experimental:** priorizan la descripción del montaje y la disposición física de los componentes involucrados en la calibración (sensores, baño térmico, conexiones, etc.).
- **Fundamentación teórica:** (con poca frecuencia) refieren a la calibración como la determinación de la relación funcional entre la resistencia eléctrica y la temperatura.
- **Menciones generales:** incluyen referencias genéricas a *calibración* o *calibrar* sin aportar detalles, generalmente haciendo referencia a los objetivos de la práctica.

La **Práctica N°3** denominada *Oscilaciones amortiguadas*, tiene como objetivos diseñar una experiencia que permita determinar la constante elástica de un resorte utilizado en un sistema masa–resorte con amortiguamiento viscoso (véase Anexo 3), evaluar el coeficiente de amortiguamiento en diferentes medios y condiciones y, por último, aplicar conceptos y herramientas de análisis de datos. En lo que respecta al montaje experimental, la propuesta de

trabajo no requiere que los estudiantes diseñen la experiencia en su conjunto, como podría interpretarse a partir de uno de los objetivos de la práctica. En este caso, el diseño se refiere más específicamente a la elaboración del montaje experimental y a la definición de un procedimiento de trabajo a partir de los materiales proporcionados. En particular, en esta práctica, se proporcionan menos orientaciones que en las prácticas anteriores para guiar el desarrollo de la actividad, aunque sí se indican algunos de los instrumentos a utilizar, como el sensor de fuerza y el sensor de movimiento. En cuanto a la guía para la elaboración del informe de esta práctica, no se incluyen preguntas específicas sobre la metodología, como sí ocurre en las dos primeras prácticas. En cambio, se solicita describir el montaje y el procedimiento, así como explicar las consideraciones que deben tenerse presentes al momento de realizar la práctica.

La clasificación automática para la clase APL en la práctica N° 3 identificó 1104 enunciados en 121 informes (media: 9 enunciados por informe; mediana: 9). El 89 % se ubica en *Metodología*, el 11 % en *Análisis de datos* y menos del 1 % en *Conclusiones*. La búsqueda de palabras frecuentes en este caso también permite identificar aquellas asociadas al proceso de medición, entre ellas se destacan *medir* (615) *resorte* (555), *sensor* (454) (véase [Figura 7.2c](#)).

En esta tercera práctica se identifican nuevamente las mismas áreas temáticas que en las dos prácticas anteriores. Se reconocen enunciados que describen la configuración física del sistema masa-resorte y las condiciones ideales del modelo experimental, a través de expresiones como por ejemplo “colgando un resorte de masa despreciable”, las cuales se corresponden con la temática *Montaje y modelado experimental*.

En mayor proporción, se identifican enunciados agrupados dentro de la temática *Descripción de instrumentos y control de variables*, referidos a las características técnicas de los sensores, sus incertidumbres y las precauciones adoptadas para mantener constantes determinadas variables.

La temática *Comparación de métodos e instrumentos de medición* abarca mayormente enunciados que diferencian entre la toma de medidas de forma estática y dinámica, es decir, aquellos que contrastan métodos de medición en condiciones de equilibrio, frente a métodos que implican el uso de sensores durante la oscilación del sistema. Esta diferenciación responde a una solicitud explícita en la guía de trabajo, en la que se solicita a los estudiantes considerar ambos casos.

Finalmente, se reconocen enunciados vinculados con el *Diseño de la toma*

de medidas y la estrategia de adquisición, en los que se mencionan aspectos como las frecuencias de muestreo, el software de recolección de datos utilizado y/o la simultaneidad de las mediciones. Además, se identifican enunciados correspondientes a la temática *Justificación procedimental y decisiones metodológicas*. En estos últimos, los estudiantes ofrecen razones metodológicas que fundamentan las decisiones adoptadas durante el experimento, explicando el “por qué” o el “para qué” de determinadas acciones (evitar errores, asegurar condiciones controladas o mejorar la calidad de la medición), como, por ejemplo: “*El plato (...) fue elegido porque, al ocupar más área, facilitaba el correcto funcionamiento del sensor*” o “*se tuvo cuidado de no generar un “batido” no deseado*”.

Los enunciados correspondientes a la última práctica del semestre, la **Práctica N°4** pueden pertenecer a tres actividades experimentales diferentes (véase Anexo 3). La clasificación automática permitió identificar 821 enunciados de 118 informes, con un promedio de 7 enunciados por informe y una mediana de 6. De estos, 225 corresponden a 34 informes de la práctica *Determinación de la velocidad del sonido*, 301 enunciados correspondientes a 37 informes de la práctica *Determinación del calor específico de un metal* y 295 a 47 informes de la *Determinación del coeficiente de fricción cinética* para el movimiento de un objeto sobre un plano inclinado.

El análisis asistido por IA de los enunciados correspondientes a esta práctica se realizó a partir de la consulta de la temática predominante en cada una de las prácticas. Se encontró que, si bien los distintos modelos empleados no coinciden en el orden exacto de predominancia, sí identifican de manera consistente las temáticas más relevantes en la práctica denominada ***Determinación del coeficiente de rozamiento cinético***. Entre ellas se destacan el *diseño de la toma de medidas y la estrategia de adquisición*, el *montaje y modelado experimental* y la *justificación procedimental y decisiones metodológicas*.

En esta cuarta práctica experimental de laboratorio se observa una mayor presencia de enunciados asociados a las temáticas *montaje y modelado experimental* y *diseño de la toma de medidas y la estrategia de adquisición*. Entre ellos, se encuentran aquellos que describen detalladamente la configuración geométrica del sistema, como el ángulo de la rampa y la posición y alineación del sensor, así como otros que refieren a la estrategia de adquisición de datos, explicitando la realización de múltiples lanzamientos o “tiradas”, con series de 20, 25, 30, 41 o incluso hasta 50 repeticiones bajo condiciones idén-

ticas. Además, estos últimos suelen incluir especificaciones sobre la frecuencia de muestreo del sensor y la preparación del programa de adquisición de datos.

En estrecha relación con las temáticas anteriores, se identifica la *justificación procedimental y decisiones metodológicas* como otra de las temáticas relevantes, dado que se encuentran enunciados que explicitan las razones de ciertas precauciones adoptadas, tales como soltar el objeto desde el reposo para evitar velocidades iniciales no deseadas o la elección de distancias mínimas con el fin de evitar la zona ciega del sensor.

La presencia de estas temáticas guarda relación con el diseño didáctico que el curso propone para esta práctica de laboratorio. En la guía de trabajo puede observarse que, si bien el objetivo principal es determinar el coeficiente de fricción mediante la aplicación de distintas herramientas de análisis estudiadas en el curso, en la sección *procedimiento experimental* de la guía, se solicita a los estudiantes que *diseñen el dispositivo* que les permita estudiar el movimiento de un cuerpo sobre un plano inclinado. Si bien se les proporciona un ejemplo ilustrativo, no se les brinda una configuración experimental completa, lo que requiere que los estudiantes tomen decisiones relativas al montaje, la estrategia de medición y el control de variables.

En otra de las actividades experimentales posibles de la **Práctica N°4** del semestre, orientada a la *Determinación de la velocidad del sonido*, el análisis de los enunciados asistido por los distintos modelos de IA permitió identificar la presencia predominante de tres temáticas: el *montaje y modelado experimental*, la *descripción de instrumentos y control de variables* y el *diseño de la toma de medidas y la estrategia de adquisición*.

Si bien en la guía de trabajo proporcionada a los estudiantes la descripción del procedimiento experimental es más detallada que en la práctica destinada a la determinación del coeficiente de rozamiento (aunque no se incluye una figura de la configuración experimental), el *Montaje y modelado experimental* constituye una temática central, dado que el experimento depende de la modificación física de un tubo para que ocurran distintos fenómenos físicos. En los informes, los estudiantes detallan aspectos como las dimensiones del tubo (longitud L), la posición del micrófono (centrado en uno de los extremos) y la configuración de los extremos (abierto o cerrado), indicando además el material utilizado para cerrar el tubo cuando corresponde.

En relación con la temática *Descripción de instrumentos y control de variables*, los enunciados destacan tanto la configuración de la tasa de muestreo

en el software utilizado, expresada en muestras por segundo, como el control de la temperatura ambiente, dado que esta variable incide en la variación de la velocidad del sonido.

Por otra parte, el *Diseño de la toma de medidas* también presenta una alta ocurrencia, ya que la obtención de los datos requiere una estrategia de medición que involucra la generación de señales acústicas. Por este motivo, se encuentran enunciados en los que se describen parámetros específicos, tales como los rangos de frecuencia empleados, el tipo de señal utilizada, la duración de los registros de audio y la cantidad de repeticiones realizadas.

La última actividad experimental posible correspondiente a la Práctica N°4 es la ***Determinación del calor específico de un metal***. El análisis de los enunciados asistido por IA identifica las mismas temáticas generales previamente consideradas para la clase APL, sin embargo, no existe una coincidencia entre los modelos respecto de cuáles son las tres más relevantes. A pesar de ello, los tres modelos empleados coinciden en señalar que una de las temáticas predominantes en los enunciados de la clase APL para esta actividad experimental es la *Descripción de instrumentos y control de variables*.

Este resultado es esperable, dado que en la guía de trabajo se solicita explícitamente que cada grupo decida cómo llevar a cabo la toma de medidas, cuáles magnitudes serán consideradas variables y cuáles se mantendrán constantes, así como que analice las limitaciones experimentales, los factores controlables y no controlables y su posible incidencia en los resultados (véase Anexo 3, Sección 3.5). En correspondencia con esto, los enunciados asociados a esta temática hacen referencia explícita a los instrumentos de medición utilizados, a sus especificaciones técnicas, así como a consideraciones relativas a la precisión y la incertidumbre. Además, se identifican enunciados centrados en el control de las condiciones iniciales necesarias para el correcto funcionamiento del método, tales como la espera de estados de estabilidad, el mantenimiento de magnitudes constantes y la adopción de precauciones para evitar la introducción de sesgos.

Por otra parte, el análisis comparativo entre los distintos modelos de IA evidencia divergencias respecto de cuál constituye la siguiente temática más relevante en la práctica *Determinación del calor específico*. Sin embargo, pueden reconocerse ciertas coincidencias parciales. En particular, ChatGPT y NotebookLM destacan el *Diseño de la toma de medidas y la estrategia de adquisición*, mientras que Gemini y ChatGPT coinciden en identificar el *Montaje y*

modelado experimental.

Si bien ambas temáticas están estrechamente vinculadas, los enunciados que refieren al *montaje y modelado experimental* se centran en describir detalladamente la construcción física del sistema a estudiar. En particular, responden a qué elementos lo constituyen, cómo está configurado y cuáles son los supuestos del modelo, tales como considerar al sistema como aislado o cerrado, asumir el equilibrio térmico o justificar el tratamiento del vaso térmico como calorímetro. En cambio, los enunciados asociados al *diseño de la toma de medidas* se vinculan con aquello que se registra, el orden en que se realiza dicho registro y los criterios bajo los cuales una medición es considerada válida. También incluyen referencias a la repetición de las mediciones. Esto puede observarse en expresiones como: “*Se repite el experimento para otras temperaturas de agua hasta tomar seis medidas con seis diferentes temperaturas de agua*” o “*Se aguardó hasta que el termómetro alcanzara la temperatura de equilibrio, determinada por su estabilidad temporal*”

Cabe señalar que Gemini y NotebookLM también identifican a la *Justificación procedimental* como una temática destacada en esta actividad experimental, dado que los enunciados ponen de manifiesto la preocupación de los estudiantes por asegurar las condiciones necesarias para suponer que el calorímetro no interviene de manera significativa en el proceso termodinámico. Esto se evidencia a través de enunciados del tipo: “*Para que el sistema actúe como un calorímetro se tapó con su tapa hermética para evitar pérdidas de calor*”, así como en aquellos en los que se argumenta por qué el metal no debe entrar en contacto con las paredes del vaso térmico ni con el termómetro, con el fin de garantizar que la transferencia de calor se produzca exclusivamente con el agua. Se identificaron también, enunciados que justifican el uso de agitadores para alcanzar una temperatura uniforme antes de realizar la lectura final, como, por ejemplo: “*El baño térmico fue llenado con agua destilada, se colocaron los metales dentro y se calentó el agua hasta llegar a la temperatura deseada, agitando el agua continuamente para homogeneizar la temperatura del sistema*”

La observación de las nubes de palabras frecuentes obtenidas para la clase APL en las cuatro prácticas realizadas en el curso, presentadas en la [Figura 7.2](#), permite identificar que en todas ellas se destaca el verbo *medir*. Este resultado indica que el foco de los enunciados de esta clase se encuentra alineado con la categoría *Aplicar* de la TBA y con las *habilidades prácticas propias de la experimentación y medición en el laboratorio* que se busca identificar en los

informes escritos por los estudiantes. En las nubes de palabras también se identifica un predominio de magnitudes físicas específicas de cada práctica: *período* y *ángulo*, en la práctica en que se utiliza un péndulo simple para determinar la aceleración gravitatoria; *temperatura* y *resistencia*, en la práctica de caracterización de dispositivos de medición de temperatura; *fuerza*, *posición* y *longitud*, en la práctica de oscilaciones amortiguadas; y *temperatura*, *tiempo*, *masa*, *ángulo*, *posición* y *distancia*, en las tres posibles actividades experimentales de la cuarta práctica. En todas ellas se evidencia una fuerte presencia de instrumentos de medición (*fotosensor*, *termómetro*, *cronómetro*, *dinamómetro*, *sensor*, *RTD*, *termistor*), lo que sugiere la importancia que los estudiantes otorgan a describir qué se mide y con qué instrumento, lo que se asocia con la temática *Descripción de instrumentos y control de variables*. Palabras como *registrar*, *tomar*, *promediar*, *intervalo*, *vez*, *repetir* y *obtener* aparecen reiteradamente en las cuatro nubes, lo que evidencia la necesidad de explicitar la estrategia de adquisición de datos, en correspondencia con la temática *Diseño de la toma de medidas y estrategia de adquisición*. Con menor frecuencia aparecen términos asociados al control de condiciones y a la justificación de decisiones (*fixar*, *ambientar*, *forzar*, *extremar*, *equilibrio*, *estático*), lo que sugiere que, si bien no es la temática dominante, la *Justificación procedimental* atraviesa las prácticas y se vincula con la validez del modelo. Se puede afirmar, entonces, que los enunciados de la clase APL se organizan principalmente en torno a la medición, el uso de instrumentos, el diseño de la toma de medidas y el control de las condiciones experimentales.

La [Figura 7.3](#) presenta una síntesis de la verificación cruzada realizada para respaldar la identificación de las «habilidades propias de la experimentación y medición en laboratorio». En el triángulo superior se sintetizan los núcleos temáticos efectivamente evidenciados en los informes de los estudiantes a partir del análisis con AA, tales como: el montaje del dispositivo conforme al modelo teórico, el uso de instrumentos de medición y el control de variables, la comparación de métodos e instrumentos de medida, el diseño de la estrategia de toma de datos, la justificación de decisiones metodológicas y la calibración de instrumentos para su uso. En el triángulo inferior izquierdo se incluyen componentes relacionadas a estas habilidades provenientes de las recomendaciones de la AAPT para la enseñanza de la física en laboratorios introductorios, mientras que en el triángulo inferior derecho se ubican los objetivos del curso vinculados a estas habilidades. A partir de la figura, se observa una alineación

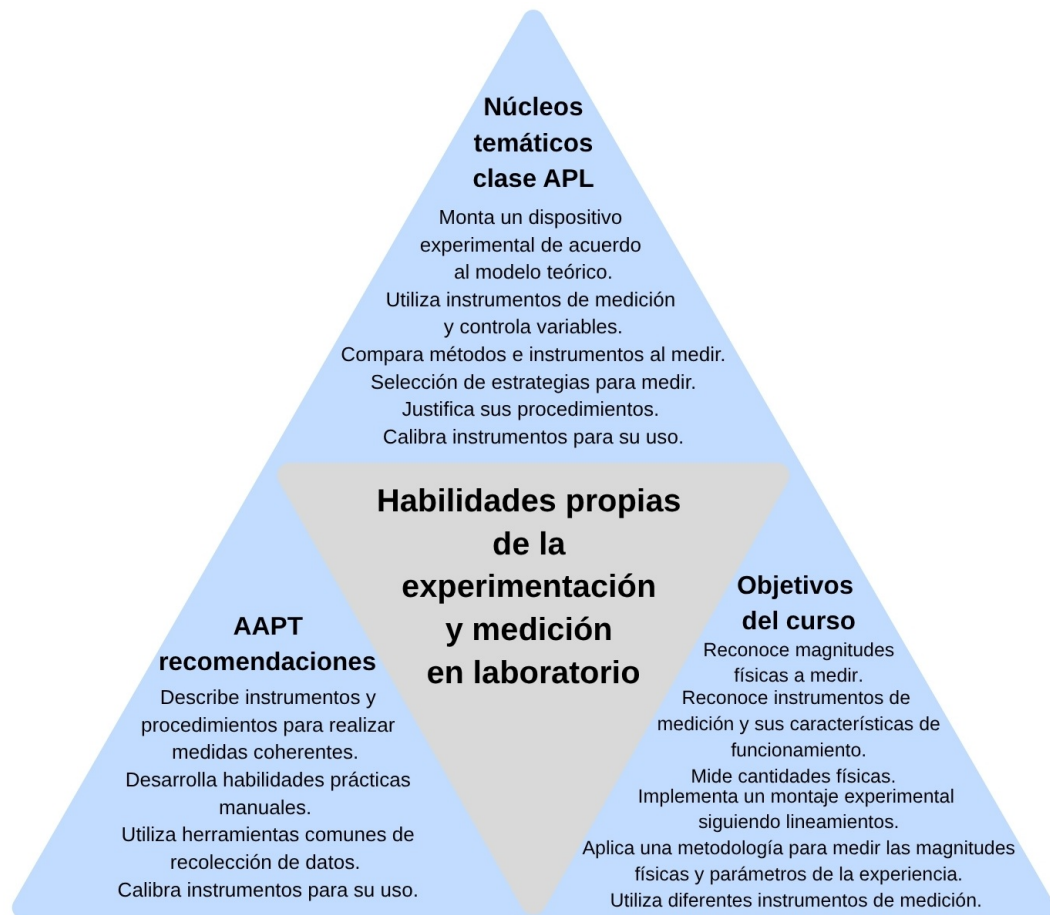


Figura 7.3: Esquema de verificación cruzada entre (i) habilidades evidenciadas en los informes, (ii) recomendaciones de la AAPT para los cursos de laboratorio introductorios y (iii) objetivos del curso de Física Experimental 1 de FING, en torno a las habilidades de experimentación y medición en laboratorio.

entre la evidencia encontrada en los informes, los lineamientos para el diseño de cursos introductorios de física experimental (AAPT) y la planificación curricular (objetivos del curso). En general, los núcleos temáticos relacionados con la habilidad de experimentación y medición en laboratorio relevados de los informes son consistentes con las recomendaciones y objetivos del curso. La calibración de instrumentos es la única habilidad que no se explicita en los objetivos del curso ni se identifica en el análisis de enunciados de la Práctica 1. Sin embargo, esta habilidad logra evidenciarse en las prácticas siguientes, por lo que el conjunto de recomendaciones y objetivos vinculados a esta dimensión se alcanza a lo largo del desarrollo del curso.

7.3. Habilidad para evaluar el diseño y procedimiento experimental (clase EDE).

En esta sección se analizan los enunciados clasificados en la clase EDE, vinculada a la habilidad para evaluar el diseño y el procedimiento experimental. Para ello, se consideró el conjunto de enunciados de los informes del semestre que fueron clasificados de forma automática dentro de esta clase. Si bien la presencia de esta clase se reconoce en las distintas prácticas del curso, el análisis en profundidad se centra en la Práctica N.º 1 y en la Práctica N.º 4, con el propósito de comparar cómo se manifiesta esta habilidad al inicio y al final del semestre. A partir de ello, se busca identificar qué aspectos de esta habilidad se evidencian en los informes de los estudiantes.

Los clasificadores automáticos identificaron enunciados correspondientes a la clase EDE en todas las prácticas del curso, aunque con variaciones en su proporción. Estas diferencias se observan en la [Figura 7.4](#), especialmente en la **Práctica N°1**, donde los modelos presentan discrepancias en la cantidad de enunciados asignados a esta clase. Esta variabilidad sugiere que algunos modelos son más restrictivos que otros al incluir enunciados dentro de la clase EDE, lo que podría vincularse con una mayor tasa de falsos positivos (FP), ya observada en los resultados del corpus de evaluación (véase la [Tabla 5.4](#)).

Luego de someter los resultados de los modelos al esquema de votación por mayoría simple, se identificaron 605 enunciados correspondientes a la clase EDE en 124 informes de la **Práctica N°1**, titulada “*Determinación de la aceleración de la gravedad local*”. El corpus utilizado para la clasificación automática se construyó a partir de 126 informes, por lo tanto, este resultado indica que solo en dos informes no se encontraron enunciados asignados a esta clase. Entre los informes en los que se identificó la clase EDE, se registró un promedio de 5 enunciados por informe, con un mínimo de 1 y un máximo de 15. Esto sugiere que, aunque existe variabilidad en la cantidad de texto vinculado con esta habilidad, su desarrollo está presente en prácticamente la totalidad de los informes.

La lectura de los enunciados clasificados dentro de la clase EDE en la Práctica N°1 permitió advertir que los clasificadores automáticos seguían incorporando, en mayor proporción, incluso luego de la etapa de votación, enunciados que no correspondían con el nivel del proceso cognitivo *Evaluar*, sino más bien con *Aplicar* o *Analizar*, es decir, falsos positivos (FP). En particular, se observó

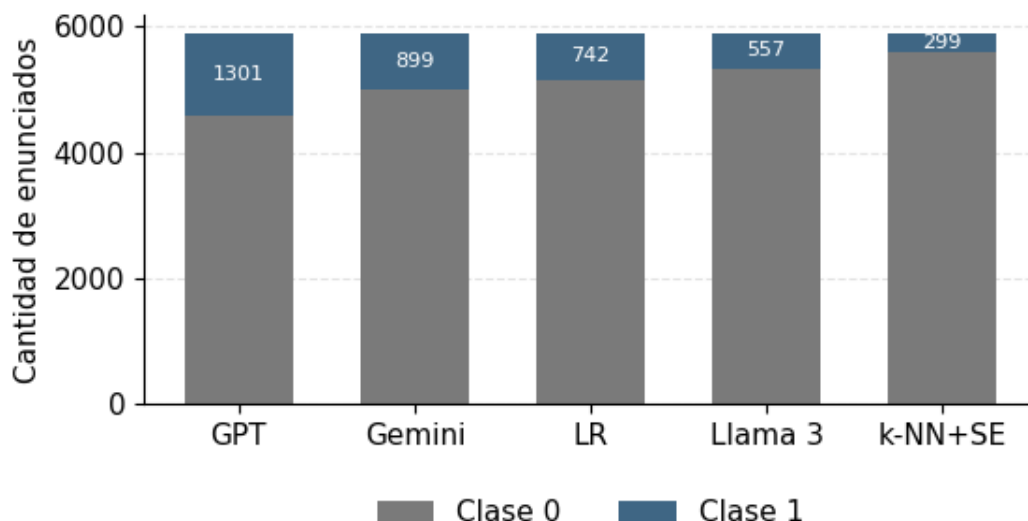


Figura 7.4: Distribución de la cantidad de enunciados clasificados como pertenecientes (clase 1) o no pertenecientes (clase 0) a la clase EDE, según el clasificador, en la Práctica N°1.

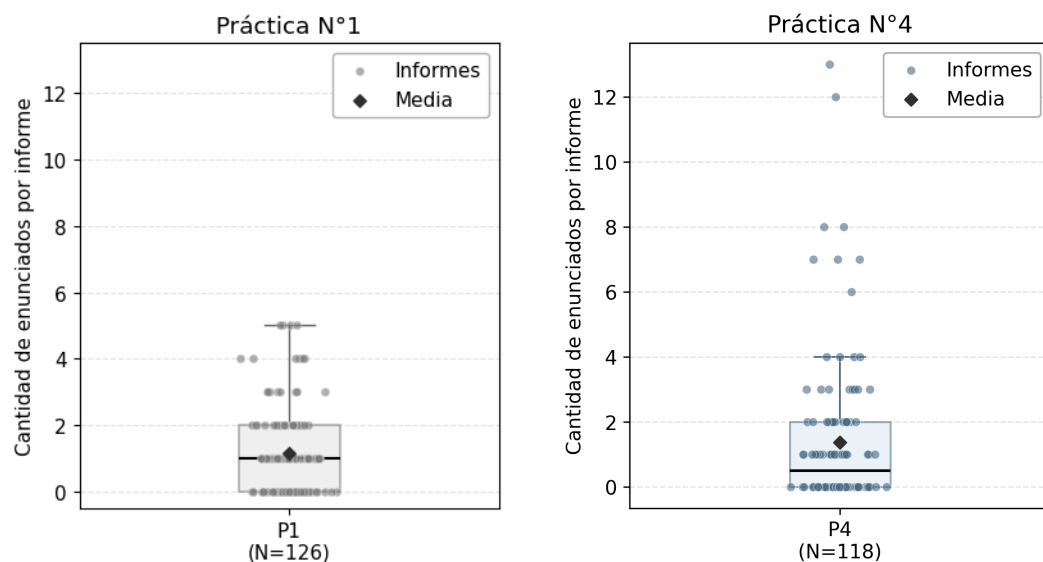
la inclusión en mayor proporción de enunciados que describen el procedimiento realizado y que no evidencian una evaluación del diseño experimental, de la metodología empleada ni de la resolución de imprevistos. También se identificó la presencia recurrente de enunciados centrados en el *Análisis* de datos, en la justificación de mediciones y procedimientos a partir de criterios estadísticos y el reporte de resultados de la medición haciendo referencia a la incertidumbre, concordancia o exactitud de las mediciones. Estos enunciados FP se centran en la evaluación de las mediciones, más que a una evaluación de los instrumentos, del dispositivo experimental o del procedimiento implementado.

Debido a la identificación reiterada de este tipo de enunciados, se realizó la lectura y clasificación de los 605 enunciados de la Práctica 1. Se encontró que, de ellos, 145 enunciados, pertenecientes a 75 informes diferentes, podrían clasificarse efectivamente dentro de la clase EDE. El resto corresponde a enunciados en los que aparecen descripciones que no pertenecen a la clase EDE, vinculada al proceso cognitivo *Evaluar*. De estos 145 enunciados, 56 se encuentran en la sección *metodología* de los informes, 31 en la sección de *análisis* de resultados y 58 en las *conclusiones*.

La [Figura 7.5a](#) muestra la distribución de la cantidad de enunciados por informe luego de la revisión realizada sobre los resultados. Se encontró que para esta clase, en la Práctica 1 del curso los estudiantes redactan en general al

menos una afirmación referida a la evaluación de la ejecución del experimento. Sobre la [Figura 7.5a](#) cada punto gris representa cada uno de los informes con su correspondiente cantidad de enunciados de esta clase. En ella se observa una alta concentración de informes con baja o nula cantidad de enunciados de la clase EDE, lo que evidencia que una baja producción relacionada al proceso cognitivo superior *Evaluar*, al que se vincula esta clase.

La baja presencia de enunciados de la clase EDE puede interpretarse en relación con el diseño de la guía de laboratorio de la Práctica N°1 (véase Anexo 3), ya que esta no plantea actividades que requieran a los estudiantes poner en juego el proceso cognitivo *Evaluar* en relación a las condiciones experimentales en que se desarrolló la experiencia. Pero, por otra parte, si se consulta la *Guía general para la elaboración de informes* (véase Anexo 2), en ella se solicita que en las conclusiones finales incluyan, “*por ejemplo: (...) ¿Qué evaluación hace de la experiencia realizada? ¿Se podría mejorar su experimento de alguna manera?*”. Además, para la elaboración del informe los estudiantes también cuentan con un tercer documento de consulta, la *Guía y pautas de corrección del Informe – Péndulo simple y Estadística* (véase 3) pero en ella no se evidencia la solicitud de evaluación de la actividad experimental realizada.



(a) Distribución de enunciados por informe en la Práctica N°1 (P1)

(b) Distribución de enunciados por informe en la Práctica N°4 (P4)

Figura 7.5: Distribución de la cantidad de enunciados de la clase EDE por informe en la primera y última práctica del semestre. Se asignó valor cero a los informes en los que no se identificaron enunciados.

Por otra parte, los enunciados pertenecientes a la clase EDE identificados en la Práctica N°1 pueden agruparse de acuerdo a tres áreas temáticas predominantes en función del tipo de evaluación que los estudiantes realizan sobre el dispositivo o procedimiento experimental. Estas áreas son:

- **Validez del diseño y ejecución** Esta clase agrupa enunciados orientados a evaluar la pertinencia del diseño experimental en relación con el objetivo de la experiencia, así como la adecuación del procedimiento seguido. También incluye valoraciones sobre si el montaje o la estrategia experimental permiten efectivamente medir la magnitud de interés o juicios sobre la correcta ejecución del procedimiento.
- **Limitaciones.** El estudiante reconoce limitaciones del dispositivo experimental o del procedimiento y las vincula con su impacto en los resultados, trascendiendo la mera descripción de errores o incertidumbres. Se incluyen enunciados que evalúan restricciones vinculadas a aspectos instrumentales como a condiciones experimentales o supuestos del modelo teórico utilizado. En ellos se identifican factores que pueden afectar la validez o precisión de los resultados, atribuidos explícitamente al diseño o al montaje.
- **Propuestas de mejora, alternativas, control de variables.** Comprende enunciados en los que se proponen modificaciones al diseño experimental o al procedimiento con el fin de mejorar la calidad de las mediciones o la validez de los resultados. Incluye sugerencias de mejora, alternativas metodológicas y estrategias de control de variables.

El análisis de los enunciados permite identificar distintos niveles de calidad en las evaluaciones realizadas por los estudiantes. A modo ilustrativo, se presenta la siguiente formulación:

“Para mejorar la precisión de futuros experimentos, sugerimos no solo tomar mediciones adicionales con mayor cuidado, sino que en la medida de lo posible, reducir las fuentes de error.”

En este caso, aunque se plantea una mejora, no se explicita qué aspecto del dispositivo o del procedimiento es evaluado como insuficiente. Una evaluación de mayor calidad no debería limitarse a enunciar la necesidad de “reducir las fuentes de error”, sino que debería especificar cuáles son dichas fuentes.

Con el fin de comparar la evolución de la calidad y cantidad de evaluación realizada por los estudiantes al final del semestre se decidió analizar la **Práctica N°4**, última actividad experimental del semestre. En este caso es importante señalar que existían tres posibles actividades experimentales: “*Determinación del calor específico de un metal*”, “*Movimiento en un plano inclinado: determinación de la fricción cinética*” o “*Determinación de la velocidad del sonido*” (véase Anexo 3).

Las tres guías analizadas presentan una estructura general similar. En todos los casos se identifican, al inicio del documento, las secciones “Objetivos”, “Fundamento teórico” y “Procedimiento experimental”. En la guía de “Velocidad del sonido” y “Calor específico” también se incluye bibliografía y apéndice, aunque en la práctica “Coeficiente de fricción” están ausentes. La principal diferencia en la estructura de las guías radica en las orientaciones dadas para el análisis de datos. Mientras que las guías de “Calor específico” y “Coeficiente de fricción” incluyen un apartado específico y numerado de “Tratamiento de datos”, la guía de “Velocidad del sonido” integra las consignas de análisis y trabajo con gráficos dentro del procedimiento experimental. Se analizaron las guías de estas prácticas en búsqueda de orientaciones hacia procesos cognitivos de orden superior asociados a *Evaluar* el diseño y/o procedimiento experimental, considerando las tres áreas temáticas identificadas en la Práctica N°1. Se encontró que, con respecto a la *validez del diseño y ejecución* las guías de “Calor específico” y “Coeficiente de fricción” estas solicitan una alta participación del estudiantado al solicitar que diseñar, discutir y elegir el diseño más adecuado para llevar a cabo la experiencia. En cambio, la guía de “Velocidad del sonido” no solicita evaluar el diseño, simplemente indica los pasos a seguir para realizar el procedimiento. En relación con las *limitaciones*, la guía de la práctica Calor específico” solicita explícitamente analizar las limitaciones experimentales y discutir las fuentes de incertidumbre. De forma similar, la guía Coeficiente de fricción” propone discutir cuáles serían las principales fuentes de incertidumbre asociadas a la práctica. Además, en ambas guías se solicita evaluar la relación entre la teoría y la realidad física haciendo exactamente la misma pregunta: “¿*El sistema cumple el modelo físico?*”, en cambio, la guía de “Velocidad del sonido” se limita a solicitar si los resultados obtenidos son concordantes con el valor de la referencia sin pedir una evaluación sobre si el dispositivo cumple verdaderamente el modelo. En cuanto a la tercera área temática, denominada *propuestas de mejora o alternativas*, tanto la guía de

“Calor específico” como la de “Coeficiente de fricción” solicitan explícitamente que los estudiantes analicen cómo se podría mejorar el montaje experimental”. Por su parte, la guía de “Velocidad del sonido” no solicita en ningún momento proponer mejoras o alternativas a la configuración del experimento. En consecuencia, aunque las guías comparten una estructura general semejante, se observan diferencias en el modo en que orientan la actividad del estudiantado. En dos de las tres guías analizadas, las consignas promueven una participación más reflexiva y activa, favoreciendo el desarrollo de procesos cognitivos de mayor complejidad vinculados con la evaluación del dispositivo experimental, la valoración de su adecuación al modelo teórico, la identificación de posibles fuentes de error y la formulación de propuestas de mejora del montaje. Esta mayor especificidad en las consignas incide en cómo los estudiantes escriben el informe tratando de responder a las mismas.

Para la **Práctica N°4**, el total de enunciados extraídos de los informes del semestre fue de 4783, distribuidos en 118 informes. Tras someter estos enunciados al proceso de clasificación automática mediante los cinco clasificadores y, posteriormente, al proceso de votación, se obtuvieron 531 enunciados clasificados como pertenecientes a la clase EDE, correspondientes a 107 informes.

Dado que previamente se había observado, que estos clasificadores tendían a incluir falsos positivos dentro de esta clase, se procedió a la lectura de la totalidad de los enunciados clasificados, con el fin de validar su pertenencia a la clase EDE. Esta revisión manual permitió identificar 162 enunciados que efectivamente pertenecían a dicha clase, distribuidos en 59 informes. Esto representa una disminución considerable respecto de la clasificación inicial e indica que, en aproximadamente el 50 % de los informes de la práctica, se evidencian en el texto escrito enunciados vinculados con el desarrollo de esta habilidad. En cuanto a los 162 enunciados, 76 de ellos se localizaron en la sección *Conclusiones*, 74 en la sección *Análisis de datos* y 12 en la sección *Metodología*.

La [Figura 7.5b](#) muestra la distribución de la cantidad de enunciados de la clase EDE por informe en la práctica N°4, considerando el total de informes (N=118), incluyendo aquellos en los que no se identificaron enunciados, donde cada punto representa un informe individual.

Se observa una alta concentración de informes que no contienen enunciados de esta clase o que presentan únicamente uno. La comparación entre la distribución de informes de la [Figura 7.5b](#) y la de la [Figura 7.5a](#) sugiere que, hacia la última práctica del semestre, se incrementa la presencia de enuncia-

dos vinculados con esta clase en los informes elaborados por el estudiantado, al tiempo que la distribución en P4 muestra una mayor variabilidad, con un número mayor de informes que presentan más de dos enunciados. Estos resultados muestran que, si bien las habilidades asociadas a la clase EDE se manifiesta en algunos informes de la Práctica N°4 con mayor desarrollo, su presencia no es generalizada en el conjunto de producciones analizadas.

Al identificar a qué práctica experimental corresponden los informes de los enunciados clasificados en la clase EDE, se encontró que 27 informes corresponden a la práctica “Calor específico”, 20 a “Coeficiente de fricción” y 12 a “Velocidad del sonido”. En cuanto a los enunciados que fueron contabilizados para conocer sobre qué práctica se estaba realizando la evaluación, 110 corresponden a “Calor específico”, 36 a “Coeficiente de fricción” y 16 a “Velocidad del sonido”.

Esta distribución guarda relación con el diseño didáctico de las guías de laboratorio. En particular, la guía correspondiente a la práctica “Calor específico” solicita de forma explícita, y en más de una oportunidad, que los estudiantes emitan un juicio acerca de la validez del diseño experimental, la ejecución del experimento, sus limitaciones y las fuentes de incertidumbre. La lectura de las guías permitió observar que, en esta práctica, este tipo de solicitudes aparece en siete oportunidades, algunas ubicadas en la sección referida al procedimiento experimental y otras en la sección de tratamiento de datos. En cambio, la guía de “Coeficiente de fricción” formula solicitudes de este tipo en dos oportunidades, ambas dentro de la sección de tratamiento de datos, mientras que la guía de “Velocidad del sonido” no incluye menciones explícitas al respecto. Estos resultados sugieren que la presencia de enunciados clasificados como EDE guardan relación con las oportunidades que ofrece cada propuesta experimental para explicitar la realización de evaluaciones sobre el diseño, la ejecución y las limitaciones del experimento.

Durante la lectura de los enunciados se encontraron las mismas áreas temáticas identificadas en la Práctica N°1. En cuanto a la *validez del diseño y ejecución*, se encuentran tres tipos de enunciados. Algunos refieren a si las hipótesis del modelo experimental se cumplen, incluso se identifican enunciados que responden a la pregunta “¿El sistema cumple el modelo físico?” que aparece en una de las guías de laboratorio. Otros refieren a la evaluación del diseño experimental y los instrumentos utilizados, y otros a una valoración del procedimiento empleado.

Con respecto a las *limitaciones*, se encuentran enunciados que refieren a condiciones experimentales asumidas que no se cumplieron, como por ejemplo que se trabajaba con un sistema aislado, entre otras. También se encontraron enunciados referidos a restricciones instrumentales como la resolución del instrumento, y otros que refieren a fallas técnicas, de montaje o condiciones experimentales que no podían controlarse, como por ejemplo el sonido ambiente.



Figura 7.6: Esquema de verificación cruzada entre (i) habilidades evidenciadas en los informes, (ii) recomendaciones de la AAPT para los cursos de laboratorio introductorios y (iii) objetivos del curso de Física Experimental 1 de FING, en relación con las habilidades vinculadas al diseño experimental.

Dentro de los enunciados clasificados también se encuentran los que incluyen *propuestas de mejora*. Entre ellos están los que refieren a alternativas en la metodología, como por ejemplo “*habría que realizar el experimento en una zona ideal*” o “*trabajar en un ambiente en el cual se pudiera alcanzar un si-*

lencio absoluto para evitar la contaminación sonora". También se identifican enunciados referidos a la mejora del montaje o los instrumentos, evidenciada en segmentos como *"la implementación de termómetros digitales puede mejorar significativamente la precisión de las mediciones de temperatura"* o *"contar con agitador magnético para garantizar temperatura uniforme"*. Otros enunciados refieren al control de variables, como *"se podría tener un control más estricto de la inclinación de la tabla"*

Además, se pudo identificar una temática adicional referida a la *resolución de problemas en tiempo real*, que incluye enunciados sobre acciones de control implementadas por los estudiantes durante el laboratorio para resolver un problema, identificándose segmentos como *"tal problema se pudo resolver colocando cinta alrededor de las caras huecas"*

La triangulación entre las recomendaciones de la AAPT para los cursos de laboratorio, los objetivos del curso y los núcleos temáticos identificados en los enunciados correspondientes a la clase EDE permite observar que, si bien los objetivos del curso se encuentran alineados con las habilidades asociadas al diseño experimental, presentan un menor grado de especificidad que los núcleos temáticos identificados en la clase EDE y que las recomendaciones de la AAPT. Esta diferencia se evidencia en una menor explicitación, dentro de los objetivos del curso, de procesos vinculados con la evaluación del diseño experimental, el reconocimiento de limitaciones y la búsqueda de soluciones a problemas emergentes durante el desarrollo de la práctica, como se muestra en la [Figura 7.6](#).

Si bien los objetivos del curso presentan un menor grado de especificidad, los núcleos temáticos identificados en la lectura de los informes evidencian una alineación con las recomendaciones de la AAPT para los cursos de laboratorio. Esto sugiere que dichas habilidades se promueven y desarrollan a lo largo del curso, aunque la presencia de enunciados de esta clase se identificó aproximadamente en la mitad de los informes de cada práctica y parece estar condicionada por el diseño de las guías de trabajo de laboratorio.

Capítulo 8

Conclusiones y consideraciones finales.

En este trabajo se desarrolló una metodología automatizada, basada en técnicas de PLN, para identificar y caracterizar las habilidades científicas presentes en los informes de laboratorio elaborados por estudiantes de un curso introductorio de Física Experimental. A partir de este análisis, se examinó en qué medida dichas habilidades quedan evidenciadas en los textos producidos por el estudiantado, así como su relación con los objetivos del curso y con el diseño didáctico de las guías de trabajo utilizadas. Al mismo tiempo, se incorporaron las perspectivas de docentes y estudiantes con el propósito de contrastar las habilidades identificadas en los informes con aquellas que los distintos actores consideran más relevantes y efectivamente priorizadas en el curso.

El desarrollo de esta metodología respondió al desafío que supone analizar un gran volumen de producciones escritas durante un semestre en cursos universitarios, cuyo abordaje exclusivamente cualitativo demanda mucho tiempo y dificulta la construcción de un procedimiento que pueda sostenerse y reproducirse a lo largo del tiempo. Es así que se recurrió al uso de herramientas de Inteligencia Artificial (IA) y, en particular, a técnicas de Procesamiento del Lenguaje Natural (PLN) para la clasificación de textos, lo que permitió construir una propuesta metodológica para la investigación en enseñanza de la física.

A continuación, se presentan los principales hallazgos, las reflexiones finales y las consideraciones a futuro, organizados a partir de dos dimensiones de

este trabajo. La primera refiere a las habilidades científicas identificadas en el curso de física experimental y la segunda se vincula con el diseño metodológico desarrollado para analizar dichas habilidades.

8.1. Las habilidades científicas en física experimental.

Uno de los aportes centrales de esta investigación fue la construcción de un marco teórico orientado a delimitar y organizar las habilidades científicas que se espera que el estudiantado desarrolle en un curso introductorio de física experimental. Siguiendo las recomendaciones de la AAPT (Kozminski et al., 2014), y de acuerdo con la taxonomía de Bloom revisada por Anderson y Krathwohl (2001), el marco teórico elaborado permitió agrupar dichas habilidades en función de los procesos cognitivos que implican. A partir de esta agrupación, se delimitaron cinco clases para la clasificación y se elaboró una guía de anotación destinada a identificar y clasificar las habilidades científicas presentes en los informes de laboratorio.

Este abordaje permitió analizar las producciones escritas tanto por práctica como en el conjunto del semestre, desplazando la mirada desde una lectura fragmentada de informes individuales hacia una comprensión más amplia del curso como espacio de formación en física experimental.

El análisis de los objetivos del curso junto con lo resultados de las encuestas a docentes y estudiantes permitió reconocer coincidencias y diferencias en torno a las habilidades priorizadas y a la forma en que estas son percibidas por ambos grupos. En líneas generales, los resultados mostraron diferencias entre aquellas habilidades que el equipo docente considera centrales y aquellas que el estudiantado reconoce como prioritarias. Entre los docentes, se destacó la valoración de las habilidades vinculadas con la medición en el laboratorio, asociadas con la clase APL. En cambio, las habilidades ligadas a procesos cognitivos de orden superior, como la evaluación del diseño experimental, del procedimiento o de la relevancia de la incertidumbre en la medición, fueron seleccionadas con menor frecuencia o no fueron seleccionadas como prioritarias.

También se identificaron algunas coincidencias. Tanto docentes como estudiantes señalaron la importancia de que el estudiantado aprenda a elaborar informes de laboratorio de acuerdo con criterios previamente establecidos.

Por otra parte, el análisis de los informes no solo permitió reconocer la presencia de determinadas habilidades, sino también advertir la influencia del diseño del curso en su desarrollo y en el modo en que se expresan en la escritura estudiantil.

El análisis de los enunciados del semestre, y en particular de la clase APL, mostró que estos pueden agruparse en cinco temáticas asociadas con habilidades prácticas propias del trabajo de laboratorio y vinculadas con la medición. Entre ellas se encuentran el uso de instrumentos y el control de variables, el diseño de la toma de medidas, la justificación del procedimiento para realizar mediciones coherentes, la comparación de métodos e instrumentos de medición, y el uso de instrumentos y herramientas comunes para la medición. A su vez, la clasificación de enunciados permitió profundizar el análisis e indagar aspectos específicos presentes en la escritura estudiantil, por ejemplo, hizo posible buscar evidencias vinculadas con la calibración de instrumentos. En conjunto, se observó una alineación entre la evidencia hallada en los informes, las recomendaciones de la AAPT para los cursos introductorios de Física Experimental y los objetivos del curso. Al mismo tiempo, los resultados de las encuestas mostraron que los docentes atribuyen una prioridad mayor al desarrollo de estas habilidades, lo que también es reconocido por los estudiantes. El análisis también permitió advertir que algunos de estos aspectos pueden no aparecer en forma explícita en los objetivos del curso, aunque sí se hacen presentes en ciertas prácticas experimentales y en los informes elaborados por los estudiantes, como ocurrió con la calibración de instrumentos de medida.

En cuanto a la clase EDE, el análisis del contenido de los enunciados clasificados en forma automática mostró que estos pueden agruparse en temáticas asociadas con la evaluación del diseño y del procedimiento experimental. Entre ellas se encuentran la valoración de la validez del diseño y de la ejecución, el reconocimiento de limitaciones del dispositivo experimental o del procedimiento, la formulación de propuestas de mejora, alternativas metodológicas y control de variables, y la resolución de problemas emergentes durante el desarrollo de la práctica. Si bien fue necesario una revisión de los enunciados clasificados en forma automática, encontrándose que existía una amplia proporción de enunciados falsos positivos, el resultado obtenido luego de este proceso permitió profundizar en cómo estas habilidades aparecen en la escritura estudiantil, encontrándose que la presencia de esta clase en los informes no resultó generalizada y se identificó aproximadamente en la mitad de los informes de las

prácticas analizadas. De este modo, se observó una alineación entre las evidencias halladas en los informes, las recomendaciones de la AAPT para los cursos introductorios de laboratorio y los objetivos del curso, aunque estos últimos presentan un menor grado de especificidad respecto de los procesos de evaluación del diseño experimental, el reconocimiento de limitaciones y la búsqueda de soluciones durante la práctica. El análisis también permitió advertir que la presencia de enunciados de esta clase parece estar condicionada por el diseño didáctico de las guías de laboratorio, ya que las prácticas que solicitan explícitamente evaluar el montaje, discutir limitaciones o proponer mejoras son aquellas en las que se identificó una mayor cantidad de evidencias vinculadas con la clase EDE. En cuanto a la percepción de docentes y estudiantes, se encontró que la importancia atribuida a las habilidades científicas asociadas con procesos cognitivos de orden superior, como *Evaluar*, varía entre los docentes, sin que se observe una tendencia predominante. También se observó que *proponer mejoras en el montaje o en el procedimiento experimental* no fue reconocido por docentes ni estudiantes como una habilidad principal dentro del curso. Sin embargo, se encontraron evidencias vinculadas con este objetivo en los informes y orientaciones dadas a los estudiantes en las guías de laboratorio.

Finalmente, sería conveniente que futuras investigaciones aborden dimensiones como el trabajo colaborativo y la evaluación en los cursos de Física Experimental, incorporando otras fuentes de información, como observaciones de clase, registros de trabajo grupal o entrevistas a estudiantes y docentes, con el fin de analizar con mayor profundidad cómo se desarrollan en el contexto del laboratorio.

8.2. El desarrollo de una metodología automatizada.

El desarrollo de este trabajo permitió advertir que la implementación de metodologías automatizadas de clasificación de texto requiere una tarea previa de construcción y validación de un conjunto de datos anotado. Este proceso implicó definir categorías, elaborar una guía de anotación y realizar distintas rondas de revisión, lo que demandó tiempo y recursos humanos. Además, la búsqueda de un buen desempeño de los clasificadores requirió múltiples ajus-

tes y pruebas. Sin embargo, una vez consolidada la metodología, la automatización permitió reducir de forma importante los tiempos de clasificación en comparación con la lectura manual y la codificación cualitativa realizada por evaluadores humanos.

En este trabajo no se encontró un método de clasificación que presentara un desempeño superior al resto en todas las clases analizadas. Ante esta situación, la estrategia de votación permitió construir una clasificación más robusta, que mejoró el desempeño de los clasificadores para algunas clases. También se observó que aquellas clases en las que los evaluadores humanos pudieron delimitar con mayor claridad si un enunciado pertenecía o no a la clase fueron, en general, las que alcanzaron mejores métricas en la clasificación automática.

En el análisis de errores de las clases APL y EDE se identificaron patrones recurrentes que aportan insumos para revisar la guía de anotación y, en el caso de los grandes modelos de lenguaje, los *prompts*.

Por otra parte, la elección entre una metodología basada en clasificadores más clásicos y modelos basados en grandes modelos de lenguaje no debería resolverse únicamente en función de su desempeño, sino también atendiendo a los tiempos de procesamiento, a la disponibilidad de recursos informáticos y al tipo de trabajo previo que exigen. Los clasificadores clásicos requieren una etapa inicial extensa de construcción del conjunto de entrenamiento y de validación de la guía de anotación, dado que dependen en mayor medida de contar con una cantidad suficiente de ejemplos previamente anotados. Por otra parte, los métodos de clasificación basados en grandes modelos de lenguaje (LLM) disminuyen la necesidad de reunir grandes cantidades de ejemplos anotados, pero plantean otras exigencias vinculadas con los recursos computacionales disponibles.

Otro aspecto relevante, que también incide en los tiempos de trabajo, es la preparación de los datos. La extracción de enunciados a partir de informes entregados en formato PDF y la limpieza posterior de esos archivos demandaron una dedicación considerable. Por ello, si se pretende desarrollar a futuro estudios de este tipo, resulta conveniente solicitar la entrega de los informes en un formato estandarizado.

En relación con el proceso de anotación, en este estudio los enunciados fueron clasificados por una única persona, lo que introduce sesgos vinculados con la interpretación individual de las categorías. Es necesario reconocer, entonces, que toda tarea de clasificación de textos depende, en alguna medida,

de quien realiza la clasificación. A esta limitación se suman, en el caso de los grandes modelos de lenguaje, los posibles sesgos propios introducidos durante su entrenamiento y desarrollo, los cuales pueden incidir en la forma en que estos modelos interpretan, priorizan o clasifican determinados enunciados. En futuros trabajos, sería deseable ampliar la cantidad de anotadores involucrados en la construcción del corpus anotado, ya que la participación de más personas permitiría reducir la incidencia de sesgos personales involuntarios y fortalecer la validez del proceso de clasificación.

Una línea de investigación futura acerca del uso de técnicas de PLN en investigaciones sobre enseñanza de la física consiste en desarrollar procedimientos que permitan estimar la incertidumbre asociada a los métodos de clasificación automática. Aunque en esta investigación el desempeño de los clasificadores se evaluó mediante métricas de clasificación y comparación con evaluadores humanos, una aproximación más cercana a las formas de validación propias de la física experimental requeriría cuantificar la incertidumbre de las métricas obtenidas y analizar su incidencia en la interpretación de los resultados.

8.3. Consideraciones finales.

Los resultados de esta investigación permiten situar los aportes del presente trabajo en una doble dimensión. Por un lado, el análisis realizado contribuyó a caracterizar las habilidades científicas que se promueven en un curso introductorio de Física Experimental y a reconocer cómo estas se expresan en los informes de laboratorio elaborados por el estudiantado. Por otro lado, el desarrollo de una metodología automatizada basada en técnicas de PLN permitió avanzar en la construcción de un procedimiento sistemático para el análisis de producciones escritas en el campo de la enseñanza de la física.

Desde esta perspectiva, la investigación mostró que los informes de laboratorio constituyen una fuente valiosa para estudiar ciertas habilidades científicas, aunque no permiten abordar todas las involucradas en el curso. Al mismo tiempo, los resultados obtenidos evidencian que el diseño didáctico de las guías de trabajo incide en las oportunidades que tienen los estudiantes para poner en juego determinadas habilidades y en el modo en que estas quedan registradas en la escritura, aportando elementos para revisar tanto las propuestas de enseñanza como las formas y herramientas de evaluación del curso.

Finalmente, esta tesis deja abierta una línea de trabajo orientada a pro-

fundizar en el uso de la Inteligencia Artificial (IA) y, en particular, del Procesamiento del Lenguaje Natural (PLN) para la investigación en enseñanza de la física, al tiempo que plantea nuevas preguntas sobre el uso actual de estas herramientas en los cursos de laboratorio. Estas preguntas resultan relevantes en relación con la elaboración de informes, el análisis de datos y la posible incidencia de la IA en las formas en que los estudiantes interpretan, escriben y comunican los resultados experimentales. En particular, se abre el desafío de analizar con qué criterios los estudiantes utilizan estas herramientas y en qué medida las producciones escritas permiten reconocer habilidades efectivamente desarrolladas por ellos, considerando que parte de esas producciones podrían estar asistidas por IA.

Referencias

- AAPT. (1998). Goals of the Introductory Physics Laboratory. *American Journal of Physics*, 66(6), 483-485. <https://doi.org/10.1119/1.19042>
- Abdelmounaim, K., & Madani, M. A. (2024). A Hybrid Ensemble Approach Integrating Machine Learning and Deep Learning with Sentence Embeddings for Webpage Content Classification. En A. Idrissi (Ed.), *Modern Artificial Intelligence and Data Science 2024: Tools, Techniques and Systems* (pp. 85-97). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-65038-3_7
- Abreu, M., Suárez, Á., Stari, C., & Marti, A. C. (2026). Exploring the potential of ChatGPT for feedback and evaluation in experimental physics. <https://arxiv.org/abs/2603.20412>
- Alaniz Ferreira, N., Díaz Rodríguez, F., & Martínez Acuña, A. (2025). *Reconocimiento y extracción de entidades a partir de archivos del pasado reciente* [Tesis de grado]. Facultad de Ingeniería, Universidad de la República. Uruguay.
- American Association of Physics Teachers. (1998). Goals of the Introductory Physics Laboratory. *American Journal of Physics*, 66(6), 483-485. <https://doi.org/10.1119/1.19042>
- American Association of Physics Teachers. (s. f.). PSSC: 50 Years Later [COMPADRE. Recuperado el 15 de abril de 2026].
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. Addison Wesley Longman, Inc.
- Antúnez, G. C., Pérez, S. M., & Petrucci, D. (2008). Concepciones de los docentes universitarios sobre los trabajos prácticos de laboratorio. *Revista Brasileira de Pesquisa em Educação em Ciências*. <https://www.redalyc.org/articulo.oa?id=571666045005>

- APFU. (2002). *El rol del laboratorio en la enseñanza de la Física* (Documento de trabajo N.º 2) (Presentado en el XII Encuentro Nacional de Profesores de Física, Carmelo, 20 de setiembre de 2002). Asociación de Profesores de Física del Uruguay (APFU). <https://apfu.uy/wp-content/uploads/2024/03/Documento-de-Trabajo-2-Carmelo-2002.pdf>
- Artstein, R., & Poesio, M. (2008). Inter-Coder Agreement for Computational Linguistics. *Computational Linguistics*, 34(4), 555-596. <https://doi.org/10.1162/coli.07-034-R2>
- Baeza, J. A., Contreras, G. A., & Fernandez, N. (2023). A Taxonomy for Physics Teachers' Practice. *The Physics Teacher*, 61(6), 496-499. <https://doi.org/10.1119/5.0069341>
- Ballesta, M., & Sguilla, S. (2025). La irrupción de la inteligencia artificial en el campo de la enseñanza y el aprendizaje de la Física en Uruguay. Perspectivas del estudiantado de bachillerato. En L. Bengochea Martínez, D. Meziat Luna, M. García Astete & A. Zúñiga Meléndez (Eds.), *Educación científica y ciencia ciudadana en la era digital* (pp. 499-503).
- Bandyopadhyay, A., Kim, H., & Charoenphol, P. (2023). Facilitate Improved Student Learning through Bloom's Taxonomy-Based Assignments in an Undergraduate Fluid Mechanics Course. *Proceedings of the 2023 ASEE Annual Conference & Exposition*. <https://doi.org/10.18260/1-2--43672>
- Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). What Is AI? En C. Bartneck, C. Lütge, A. Wagner & S. Welsh (Eds.), *An Introduction to Ethics in Robotics and AI* (pp. 5-16). Springer. https://doi.org/10.1007/978-3-030-51110-4_2
- Bates, S. P., Galloway, R. K., Riise, J., & Homer, D. (2014). Assessing the quality of a student-generated question repository. *Physical Review Special Topics - Physics Education Research*, 10(2). <https://doi.org/10.1103/physrevstper.10.020105>
- Becerra Labra, C., & Silva Arias, L. A. (2024). Trabajo de laboratorio en física y química a través de un enfoque problematizado. *Revista de Enseñanza de la Física*, 36(1), 53-66. <https://doi.org/10.55767/2451.6007.v36.n1.45312>
- Bhakti, Y., Arthur, R., & Supriyati, Y. (2025). Development of Physics HOTS Assessment Instruments in High School: A Comprehensive Literature Review. *Jurnal Ilmu Pendidikan Fisika*.
- Bishop, C. M. (2009). *Pattern Recognition and Machine Learning*. Springer.

- Bloom, B. S. (1990). *Taxonomía de los objetivos de la educación. La clasificación de las metas educacionales. Manuales I y II*. (10.^a ed.). El Ateneo.
- Borges, A. T. (2002). Novos rumos para o laboratório escolar de ciências. *Caderno Brasileiro de Ensino de Física*, 19(3), 291-313.
- Bralin, A., Morpew, J., Rebello, C., & Rebello, N. S. (2023-20 de julio). Analysis of Student Essays in an Introductory Physics Course Using Natural Language Processing. *Physics Education Research Conference (PERC) 2023*, 58-63.
- Bralin, A., Sirnoorkar, A., Zhang, Y., & Rebello, N. S. (2024). Mapping the literature landscape of artificial intelligence and machine learning in physics education research. *Physics Education Research Conference 2024*, 52-59.
- Buick, J. (2011). Physics Assessment and the Development of a Taxonomy. *European Journal of Physics Education*, 2(1), 12-27.
- Calderón, S., Núñez, P., Di Laccio, J., & Iannelli, L. M. (2015). Aulas Virtuales de bajo costo, usando TIC. *Eureka sobre Enseñanza y Divulgación de las Ciencias*, 12(1), 212-226.
- Campbell, J., Ansell, K., & Stelzer, T. (2024). Evaluating IBM's Watson natural language processing artificial intelligence as a short-answer categorization tool for physics education research. *Phys. Rev. Phys. Educ. Res.*, 20, 010116. <https://doi.org/10.1103/PhysRevPhysEducRes.20.010116>
- Carrascosa, J., Gil Pérez, D., Vilches, A., & Valdés, P. (2006). Papel de la actividad experimental en la educación científica. *Caderno Brasileiro de Ensino de Física*, 23(2), 157-181. <https://periodicos.ufsc.br/index.php/fisica/article/view/6274/12764>
- Castiblanco, O., & Vizcaíno, D. (2008). La experiencia del laboratorio en la enseñanza de la física. *Revista Digital Educación en Ingeniería*, 3(5), 68-74. <https://doi.org/10.26507/rei.v3n5.151>
- Cheng, X., Nanjundan, P., & George, J. P. (2025). *Introduction to Natural Language Processing: Exploring Techniques, Applications, and Challenges* (1.^a ed.). River Publishers. <https://doi.org/10.1201/9781003660675>
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37-46. <https://doi.org/10.1177/001316446002000104>
- Cordero, M. (2024). Inteligencia Artificial en el aula: oportunidades y desafíos para la didáctica de la matemática y física universitaria. *Revista In-*

- ternacional de Pedagogía e Innovación Educativa*, 4(1), 193-207. <https://doi.org/10.51660/ripie.v4i1.154>
- Corrales, A. P., & Meneses, W. (2025). Interpretaciones de la primera ley de Newton: una revisión desde textos clásicos y la inteligencia artificial. En I. J. Idoyaga & J. E. Maeyoshimoto (Eds.), *Diálogo Interamericano sobre Inteligencia Artificial en Educación Científica* (pp. 55-64). Universidad de Buenos Aires. Instituto de Investigación en Educación Superior.
- Damayanti, N., Hartono, Subali, B., Nugroho, S. E., & Sureeporn, K. (2020). Items analysis of physics assessment based on cognitive level of high order thinking skills in bloom taxonomy. *Journal of Physics: Conference Series*, 1521(2), 022022. <https://doi.org/10.1088/1742-6596/1521/2/022022>
- Di Laccio, J. (2017). Los cinco problemas en la enseñanza de la física experimental. *Departamento de física del CENUR*, 1-8. http://les.edu.uy/fisica/doc/F%7B%5C'%7Bi%7D%7Dsica-Di-Laccio/Los_cinco_problemas_en_la_ensenanza_de_FEXP.pdf
- Di Laccio, J. (2024). *Aprendizaje de Física en cursos universitarios con TIC*. Universidad de la República. Comisión Sectorial de Enseñanza.
- Di Mauro, M. F., Furman, M., & Bravo, B. (2015). Las habilidades científicas en la escuela primaria: un estudio del nivel de desempeño en niños de 4to año. *Revista electrónica de investigación en educación en ciencias*, 10, 1-10.
- Dias, V. S., Marins, F. A. S., & Kitamura Filho, Y. (2021). Análise das questões de Física do Exame Nacional do Ensino Médio na perspectiva da taxonomia de Bloom revisada (2014–2019). *Revista de Enseñanza de la Física*, 33(2), 495-502.
- Ding, L. (2014). Seeking missing pieces in science concept assessments: Re-evaluating the Brief Electricity and Magnetism Assessment through Rasch analysis. *Phys. Rev. ST Phys. Educ. Res.*, 10, 010105. <https://doi.org/10.1103/PhysRevSTPER.10.010105>
- Editorial Reverté. (s. f.). Physical Science Study Committee (PSSC) [Recuperado el 15 de abril de 2026]. <https://www.reverte.com/coleccion/pssc/>
- Etkina, E., Brookes, D., Planinsic, G., Gentile, M., Buggé, D., Flakker, J., Brahmia, S., Blackman, M., Zou, X., & Van Heuvelen, A. (2025). *ISLE*

- Physics*. Consultado el 29 de octubre de 2025, desde <https://www.islephysics.net/>
- Etkina, E., Karelina, A., Ruibal-Villasenor, M., Rosengrant, D., Jordan, R., & Hmelo-Silver, C. E. (2010). Design and reflection help students develop scientific abilities: Learning in introductory physics laboratories. *Journal of the Learning Sciences*, 19(1), 54-98. <https://doi.org/10.1080/10508400903452876>
- Etkina, E., Van Heuvelen, A., White-Brahmia, S., Brookes, D. T., Gentile, M., Murthy, S., Rosengrant, D., & Warren, A. (2006). Scientific abilities and their assessment. *Physical Review Special Topics - Physics Education Research*, 2(2), Artículo 020103. <https://doi.org/10.1103/PhysRevSTPER.2.020103>
- Fadillah, M. A., Usmeldi, U., & Asrizal, A. (2024). the Role of Chatgpt and Higher-Order Thinking Skills As Predictors of Physics Inquiry. *Journal of Baltic Science Education*, 23, 1178-1192. <https://doi.org/10.33225/jbse/24.23.1178>
- Ferraz, A. P. d. C. M., & Belhot, R. V. (2010). Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. *Gestão and Produção*, 17(2), 421-431. <https://doi.org/10.1590/S0104-530X2010000200015>
- Fussell, R. K., Flynn, M., Damle, A., Fox, M. F. J., & Holmes, N. G. (2025). Comparing large language models for supervised analysis of students' lab notes. *Phys. Rev. Phys. Educ. Res.*, 21, 010128. <https://doi.org/10.1103/PhysRevPhysEducRes.21.010128>
- Fussell, R. K., Mazrui, A., & Holmes, N. G. (2022). Machine learning for automated content analysis: characteristics of training data impact reliability. *Physics Education Research Conference 2022*, 194-199.
- Gallardo Fernández, T., Salgueiro Finkelstein, G., & Rivas Masullo, M. (2024). *Evolución temporal del tratamiento periodístico de tópicos* [Tesis de grado]. Universidad de la República (Uruguay), Facultad de Ingeniería.
- Gamboa Solano, L., Guevara Mora, M. G., Mena, Á., & Umaña Mata, A. C. (2023). Taxonomia revisada de Bloom como apoyo para la redacción de resultados de aprendizaje y el alineamiento constructivo. *Innovaciones Educativas*, 25(38), 140-155. <https://doi.org/10.22458/ie.v25i38.4529>
- Gani, M. O., Ayyasamy, R. K., Sangodiah, A., et al. (2023). Clasificación de preguntas de examen basada en la taxonomía de Bloom: El resultado

- de CNN y la técnica óptima de incrustación de palabras preentrenada. *Education and Information Technologies*, 28, 15893-15914. <https://doi.org/10.1007/s10639-023-11842-1>
- Gemini Team, G. (2025). Gemini 2.5: Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next Generation Agentic Capabilities. <https://arxiv.org/abs/2507.06261>
- Gogus, A. (2012). Learning Objectives. En N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 1950-1954). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_144
- Google. (2017). *Google Colaboratory (Colab)*. Consultado el 24 de noviembre de 2025, desde <https://colab.google/>
- Google. (2026a). *Gemini 3 Developer Guide*. Google AI for Developers. <https://ai.google.dev/gemini-api/docs/gemini-3>
- Google. (2026b). *NotebookLM*. Google. Consultado el 3 de diciembre de 2025, desde <https://notebooklm.google/>
- Google AI. (2025). *Descripción general del modelo EmbeddingGemma*. Consultado el 24 de noviembre de 2025, desde <https://ai.google.dev/gemma/docs/embeddinggemma?hl=es-419>
- Google DeepMind. (2025). Gemini 2.5 Flash-Lite [Model card]. Consultado el 16 de noviembre de 2025, desde <https://deepmind.google/models/gemini/flash-lite/>
- Gui, J., Chen, T., Zhang, J., Cao, Q., Sun, Z., Luo, H., & Tao, D. (2024). A Survey on Self-Supervised Learning: Algorithms, Applications, and Future Trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12), 9052-9071. <https://doi.org/10.1109/TPAMI.2024.3415112>
- Gunawan, G., Nisrina, N., Y Suranti, N. M., Herayanti, L., & Rahmatiah, R. (2018). Virtual Laboratory to Improve Students' Conceptual Understanding in Physics Learning. *Journal of Physics: Conference Series*, 1108(1), 012049. <https://doi.org/10.1088/1742-6596/1108/1/012049>
- Hake, R. R. (1992). Socratic pedagogy in the introductory physics laboratory. *The Physics Teacher*, 30(9), 546-552. <https://doi.org/10.1119/1.2343637>
- Hofstein, A., & Lunetta, V. N. (2004). The Laboratory in Science Education: Foundations for the Twenty-First Century. *Science Education*, 88(1), 28-54. <https://doi.org/10.1002/sce.10106>

- Jurafsky, D., & Martin, J. H. (2009). *Speech and Language Processing (2nd Edition)*. Prentice-Hall, Inc.
- Jurafsky, D., & Martin, J. H. (2025a). *Large Language Models* [Chapter 7 of *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, draft of August 24, 2025]. Consultado en agosto de 2025, desde https://web.stanford.edu/~jurafsky/slp3/old_aug25/7.pdf
- Jurafsky, D., & Martin, J. H. (2025b). *LogisticRegression and Text Classification* [Chapter 4 of *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, draft of August 24, 2025]. Consultado en agosto de 2025, desde https://web.stanford.edu/~jurafsky/slp3/old_aug25/4.pdf
- Jurafsky, D., & Martin, J. H. (2025c). *Neural Networks* [Chapter 6 of *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, draft of August 24, 2025]. Consultado en agosto de 2025, desde https://web.stanford.edu/~jurafsky/slp3/old_aug25/6.pdf
- Jurafsky, D., & Martin, J. H. (2026a). *Embeddings* [Chapter 5 of *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, draft of January 6, 2026]. Consultado en abril de 2026, desde <https://web.stanford.edu/~jurafsky/slp3/5.pdf>
- Jurafsky, D., & Martin, J. H. (2026b). *N-gram Language Models* [Chapter 3 of *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, draft of January 6, 2026]. Consultado en marzo de 2026, desde <https://web.stanford.edu/~jurafsky/slp3/3.pdf>
- Jurafsky, D., & Martin, J. H. (2026c). *Naive Bayes, Text Classification, and Sentiment* [Appendix B of *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, draft of January 6, 2026]. Consultado en marzo de 2026, desde <https://web.stanford.edu/~jurafsky/slp3/B.pdf>

- Kilde-Westberg, S., Johansson, A., & Enger, J. (2025). Generative AI as a lab partner: A case study. *Phys. Rev. Phys. Educ. Res.*, 21, 020119. <https://doi.org/10.1103/ggy1-3kjk>
- Kotsis, K. T. (2025). Artificial Intelligence and the Scientific Process: A Review of ChatGPT's Role to Foster Experimental Thinking in Physics Education. *European Journal of Contemporary Education and E-Learning*, 3, 183-198. [https://doi.org/10.59324/ejceel.2025.3\(3\).14](https://doi.org/10.59324/ejceel.2025.3(3).14)
- Kozminski, J., Lewandowski, H., Beverly, N., Lindaas, S., Deardorff, D., Reagan, A., Dietz, R., Tagg, R., Eblen-Zayas, M., Williams, J., Hobbs, R., & Zwickl, B. (2014). *AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum*. American Association of Physics Teachers. https://www.aapt.org/resources/upload/labguidelinesdocument_ebendorsed_nov10.pdf
- Küchemann, S., & Kuhn, J. (2026). Editorial: Focused Collection: Artificial Intelligence Tools in Physics Teaching and Physics Education Research. *Phys. Rev. Phys. Educ. Res.*, 22, 010001. <https://doi.org/10.1103/PhysRevPhysEducRes.22.010001>
- Kuncheva, L. I. (2004). *Combining Pattern Classifiers: Methods and Algorithms*. John Wiley & Sons, Inc.
- Kyritsis, K., Liapis, C. M., Perikos, I., Paraskevas, M., & Kapoulas, V. (2025). From Transformers to Voting Ensembles for Interpretable Sentiment Classification: A Comprehensive Comparison. *Computers*, 14(5), 167. <https://doi.org/10.3390/computers14050167>
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174.
- Lane, H., & Dyshel, M. (2025). *Natural Language Processing in Action* (2.^a ed.). Manning.
- Latif, E., Parasuraman, R., & Zhai, X. (2024). PhysicsAssistant: An LLM-Powered Interactive Learning Robot for Physics Lab Investigations. <https://arxiv.org/abs/2403.18721>
- Lazarte, G., Pérez, L. A. L., & Díaz Dávila, L. C. (2024). Inteligencia Artificial para la Enseñanza de la Física: exploración de BERT en el contexto de experimentos físicos. *JAIIO, Jornadas Argentinas de Informática*, 10(1), 133-136. <https://revistas.unlp.edu.ar/JAIIO/article/view/17915>

- Li, H., Li, X., Caragea, D., & Caragea, C. (2018). Comparison of word embeddings and sentence encodings as generalized representations for crisis tweet classification tasks. *Proceedings of ISCRAM Asia Pacific*.
- Liapis, C. M., Kyritsis, K., Perikos, I., Spatiotis, N., & Paraskevas, M. (2024). A Hybrid Ensemble Approach for Greek Text Classification Based on Multilingual Models. <https://doi.org/10.3390/bdcc8100137>
- Liu, X. (2025). An Evaluation of the Pedagogical Soundness and Usability of AI-Generated Lesson Plans Across Different Models and Prompt Frameworks in High-School Physics. https://doi.org/10.35542/osf.io/r3xkt_v1
- Luo, H., Weng, J., Lu, X., & Yu, Z. (2024). BPE: Exploring the Prompt Framework of Physics Exercises Generated from Bloom's Taxonomy in LLM. *2024 5th International Conference on Machine Learning and Computer Application (ICMLCA)*, 26-29. <https://doi.org/10.1109/ICMLCA63499.2024.10754020>
- Mahligawati, F., Allanas, E., Butarbutar, M. H., & Nordin, N. A. N. (2023). Artificial intelligence in Physics Education: a comprehensive literature review. *Journal of Physics: Conference Series*, 2596(1), 012080. <https://doi.org/10.1088/1742-6596/2596/1/012080>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- May, J. M. (2023). Historical analysis of innovation and research in physics instructional laboratories: Recurring themes and future directions. *Phys. Rev. Phys. Educ. Res.*, 19, 020168. <https://doi.org/10.1103/PhysRevPhysEducRes.19.020168>
- Meher, J. P., & Mall, R. (2023). BloomBERT: A Deep Learning-Based Cognitive Complexity Classifier of Assessment Questions. *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)*, 318-322. <https://doi.org/10.1109/ICALT58122.2023.00099>
- Mestre, J. P., & Docktor, J. L. (2021). *The Science of Learning Physics: Cognitive Strategies for Improving Instruction*. World Scientific Publishing Co. Pte. Ltd.
- Meta, A. (2024). Llama 3 Model Card. https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md
- Mills, E., Mizouri, A., & Peach, A. (2025). Prompting Better Feedback: A Study of Custom GPT for Formative Assessment in Undergradua-

- te Physics. *Education Sciences*, 15(8). <https://doi.org/10.3390/educsci15081058>
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Mohammed, M., & Omar, N. (2020). Clasificación de preguntas basada en el dominio cognitivo de la taxonomía de Bloom mediante TF-IDF modificado y word2vec. *PLOS ONE*, 15(3), e0230442. <https://doi.org/10.1371/journal.pone.0230442>
- Monteiro, M., Stari, C., Cabeza, C., & Marti, A. C. (2019). Sensors based on micromechanical devices: mobile labs at the service of experimental science teaching. <http://arxiv.org/abs/1901.05079>
- Monteiro, M., Martí, A., & Cabeza, C. (2014). Con la Física a todas partes: experiencias utilizando el teléfono inteligente. *Revista Educación en Física*, 8, 17.
- Monteiro, M., Stari, C., & Marti, A. C. (2023). Los sensores de los dispositivos móviles: una herramienta innovadora en la enseñanza de las ciencias físicas. *Modelling in Science Education and Learning*, 16(2), 37-42. <https://doi.org/10.4995/msel.2023.19104>
- Nurlailiyah, A., Deta, U. A., Ain, T. N., Haq, M. S., Lestari, N. A., & Yantidewi, M. (2019). Analysis of High School Physics National Examination questions based on Bloom Taxonomy and National Examination Question Standard in 2017/2018. *Journal of Physics: Conference Series*, 1171(1), 012041. <https://doi.org/10.1088/1742-6596/1171/1/012041>
- Olivera, N. (2023). *Estudio de la incidencia del planteo de prácticas en cursos de física experimental* [Monografía del curso “Iniciación en la investigación”. Universidad de la República (Uruguay), Facultad de Ciencias].
- Omopekunola, M. O., & Kardanova, E. Y. (2026). Can Large Language Models Develop High-Stakes Physics Exam Items? A Comprehensive Study of Cognitive and Psychometric Efficacy. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-025-10289-y>
- OpenAI. (2025a). *GPT-5.2 Model*. OpenAI. Consultado el 22 de febrero de 2026, desde <https://developers.openai.com/api/docs/models/gpt-5.2>
- OpenAI. (2025b). *Presentación de gpt-oss*. Consultado el 24 de noviembre de 2025, desde <https://openai.com/es-419/index/introducing-gpt-oss/>
- OpenAI, : Agarwal, S., Ahmad, L., Ai, J., Altman, S., Applebaum, A., Arbus, E., Arora, R. K., Bai, Y., Baker, B., Bao, H., Barak, B., Bennett, A., Bertao, T., Brett, N., Brevdo, E., Brockman, G., Bubeck, S., ... Zhao,

- S. (2025). gpt-oss-120b and gpt-oss-20b Model Card. <https://arxiv.org/abs/2508.10925>
- Ortigoza, L. d. V., Contini, L. E., & Lessa de Oliveira, M. d. S. (2016). Actividades en laboratorios de enseñanza de física: percepciones de estudiantes de Brasil y Argentina: tema de reflexión docente. *Revista Inter-Ação*, 41 (3), 651-670. <https://doi.org/10.5216/ia.v41i3.41900>
- Pereira, M. V., & Moreira, M. C. d. A. (2017). Atividades prático-experimentais no ensino de Física. *Caderno Brasileiro de Ensino de Física*, 34 (1), 265-277. <https://doi.org/10.5007/2175-7941.2017v34n1p265>
- Petrucci, D., Ure, J., & Salomone, H. D. (2006). Cómo ven a los trabajos prácticos de laboratorio de física los estudiantes universitarios. *Revista de Enseñanza de la Física*, 19(1), 7-19. <https://doi.org/10.55767/2451.6007.v19.n1.8060>
- PhysPort. (2025). *Force Concept Inventory (FCI)*. Consultado el 29 de octubre de 2025, desde <https://www.physport.org/assessments/assessment.cfm?A=FCI>
- Pollard, B., Hobbs, R., Henderson, R., Caballero, M. D., & Lewandowski, H. J. (2021). Introductory physics lab instructors' perspectives on measurement uncertainty. *Phys. Rev. Phys. Educ. Res.*, 17, 010133. <https://doi.org/10.1103/PhysRevPhysEducRes.17.010133>
- Poluakan, C., Tilaar, A. F., Tuerah, P., & Mondolang, A. (2019). Implementation of the revised bloom taxonomy in assessment of physics learning. *Proceedings of the 1st international conference on education, science and technology (ICESTech)*, 1-6.
- Pustejovsky, J., & Stubbs, A. (2014). *Natural Language Annotation for Machine Learning* (J. Steele & M. Blanchette, Eds.). O'Reilly Media, Inc.
- Qadar, R., Syam, M., & Mahdiannur, M. A. (2025). Analyzing high school physics teachers' understanding of cognitive process and knowledge dimensions in assessment design using the revised Bloom's taxonomy. *Discover Education*.
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf

- Ribeiro, M. S., Freitas, D. d. S., & Miranda, D. E. d. (1997). O ensino de laboratório de física na UEFS: considerações teórico-pedagógicas. *Sitientibus*, (16), 123-130. <https://doi.org/10.13102/sitientibus.vi16.9730>
- Riveros, H. G. (2017). Investigación en enseñanza de la física experimental en el siglo XXI [Recuperado el 16 de abril de 2026]. *Revista mexicana de física E*, 63(1), 68-75. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-35422017000100068&lng=es&tlng=es
- Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach* (3.^a ed.). Prentice Hall.
- Russell, S. J., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4.^a ed.) [Global Edition]. Pearson.
- Samuelsson, C. R., Polverini, G., & Gregorcic, B. (2025). Embodied AI?: Exploring GenAI's role in physics lab design. *The 16th Conference of The European Science Education Research Association (ESERA 2025), Copenhagen, Denmark, August 25-29, 2025*.
- Santos, A., Aquines, O., & Martínez, H. (2025). Impacto de la inteligencia artificial en la enseñanza de la Física. *Ixaya. Revista Universitaria de Desarrollo Social*, 15(29), 167-196. <https://revistaixaya.cucsh.udg.mx/index.php/ixa/article/view/7842>
- Scikit-learn developers. (s.f.). Supervised learning — scikit-learn: Machine Learning in Python. Consultado el 11 de noviembre de 2025, desde https://scikit-learn.org/stable/supervised_learning.html
- Sebastià, J. M. (2006). ¿Qué se pretende en los laboratorios de física universitaria? *Enseñanza de las Ciencias: revista de investigación y experiencias didácticas*, 5(3), 196-204. <https://raco.cat/index.php/Ensenanza/article/view/50999>
- Sepúlveda Navarro, N., & Carrasco Maturana, J. (2024). Física experimental para ingeniería, el desafío de implementar la retroalimentación en el laboratorio. *Revista de Enseñanza de la Física*, 36, 317-323. <https://doi.org/10.55767/2451.6007.v36.n.47270>
- Silva, V. A. d., & Martins, M. I. (2014). Análise de questões de Física do ENEM pela Taxonomia de Bloom Revisada. *Ensaio: Pesquisa em Educação em Ciências (Belo Horizonte)*, 16(3), 189-202. <https://doi.org/10.1590/1983-21172014160309>
- Silva Dias, V., Silva, F. A., & Kitamura Filho, Y. (2021). Análise das questões de Física do Exame Nacional do Ensino Médio na perspectiva da taxo-

- nomia de Bloom revisada (2014-2019). *Revista de Enseñanza de la Física*, 33(2), 495-502. <https://doi.org/10.55767/2451.6007.v33.n2.35304>
- Sokoloff, D. R., Laws, P. W., & Thornton, R. K. (2007). RealTime Physics: active learning labs transforming the introductory laboratory. *European Journal of Physics*, 28(3), S83. <https://doi.org/10.1088/0143-0807/28/3/S08>
- Stanford NLP Group. (2020). Stanza: A Python Natural Language Processing Toolkit for Many Human Languages.
- Suárez, Á., & Martí, A. C. (2024). El ChatGPT como una herramienta en los cursos de física. En *Inteligencia artificial en educación: contribuciones desde la teoría, las prácticas pedagógicas y la investigación* (pp. 209-221). Magro.
- Suleimenova, L., Umarova, Z., Berkimbayev, K., Adylbekova, E., Zhetpisbayeva, G., & Zhamalova, K. (2025). Leveraging Bloom's Cognitive Framework and Machine Learning Methods for Academic Performance Prediction. *2025 10th International Conference on Computer Science and Engineering (UBMK)*, 108-113. <https://doi.org/10.1109/UBMK67458.2025.11207066>
- Thacker, B. (2023). Inquiry-based experimental physics: Twenty years of an evidence-based, laboratory-based physics course for algebra-based physics students. *Phys. Rev. Phys. Educ. Res.*, 19, 020116. <https://doi.org/10.1103/PhysRevPhysEducRes.19.020116>
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). LLaMA: Open and Efficient Foundation Language Models. arXiv preprint arXiv:2302.13971. <https://doi.org/10.48550/arXiv.2302.13971>
- Trumper, R. (2003). The Physics Laboratory. A Historical Overview and Future Perspectives. *Science and Education*, 12, 645-670. <https://doi.org/10.1023/A:1025692409001>
- Vera, H. S., Dua, S., Zhang, B., Salz, D., Mullins, R., Panyam, S. R., Smoot, S., Naim, I., Zou, J., Chen, F., Cer, D., Lisak, A., Choi, M., Gonzalez, L., Sanseviero, O., Cameron, G., Ballantyne, I., Black, K., Chen, K., ... Seyedhosseini, M. (2025). EmbeddingGemma: Powerful and Lightweight Text Representations. <https://arxiv.org/abs/2509.20354>
- VERBI Software. (2021). MAXQDA 2022. <https://www.maxqda.com>

- Villani, A., Dias, V. S., & Valadares, J. M. (2010). The Development of Science Education Research in Brazil and Contributions from the History and Philosophy of Science. *International Journal of Science Education*, 32(7), 907-937. <https://doi.org/10.1080/09500690902855711>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes* (M. Cole, V. John-Steiner, S. Scribner & E. Souberman, Eds.). Harvard University Press.
- Wang, L., Yang, N., Huang, X., Yang, L., Majumder, R., & Wei, F. (2024). Multilingual E5 Text Embeddings: A Technical Report. <https://arxiv.org/abs/2402.05672>
- Wei, Y., Wei, X., & Chen, Z. (2026). Trends and themes of twenty years of physics education research: A bibliometric analysis of PRPER publications. *Phys. Rev. Phys. Educ. Res.*, 22, 010104. <https://doi.org/10.1103/384p-1hfv>
- Werner da Rosa, C., & Becker da Rosa, A. (2007). O ensino de física na universidade de passo fundo: uma investigação nos objetivos das atividades experimentais [Recuperado el 16 de abril de 2026]. *Educere*, 11(37), 327-332. http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1316-49102007000200019&lng=es&tlng=pt
- Wijesinghe, P., Seneviratne, V. A., Medsker, L. R., Giles, C. D., & Ottati, M. (2025). Using learning maps and bloom's taxonomy to develop a new instrument to assess knowledge transfer from physics to statics courses. *Proceedings of the 2025 ASEE Annual Conference & Exposition*. <https://doi.org/10.18260/1-2--57307>
- Wilcox, B. R., & Lewandowski, H. J. (2017). Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics. *Phys. Rev. Phys. Educ. Res.*, 13, 010108. <https://doi.org/10.1103/PhysRevPhysEducRes.13.010108>
- Yahya, A. A., Toukal, Z., & Osman, A. (2012). Bloom's Taxonomy-Based Classification for Item Bank Questions Using Support Vector Machines. En W. Ding, H. Jiang, M. Ali & M. Li (Eds.), *Modern Advances in Intelligent Systems and Tools* (pp. 135-140). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30732-4_17

Publicaciones y presentaciones en congresos derivadas de la tesis

A partir de esta tesis surgió la siguiente publicación:

- Pagella, L., Stari, C., & Rosá, A. (2026). Hacia una metodología automatizada para identificar habilidades científicas en informes de laboratorio de Física. *Revista de Enseñanza de la Física*, 38(1), 229–235. DOI: [10.55767/2451.6007.v38.n1.52388](https://doi.org/10.55767/2451.6007.v38.n1.52388)

También se realizaron las siguientes publicaciones en congresos nacionales e internacionales, indicándose con un asterisco la persona que realizó la presentación:

- Lorena Pagella*, Cecilia Stari & Aiala Rosá. *Hacia una metodología automatizada para identificar habilidades científicas en informes de laboratorio de Física*. Trabajo presentado en la XVI Conferencia Interamericana sobre Educación en Física (CIAEF), Valparaíso, Chile, 2025.
- Lorena Pagella*, Cecilia Stari & Aiala Rosá. *¿Qué habilidades científicas se desarrollan en un curso de Física experimental? Estudio de informes de laboratorio*. Póster presentado en el XVIII Encuentro de la Sociedad Uruguaya de Física (SUF), Hotel del Lago, Punta del Este, Uruguay, 18 y 19 de octubre de 2024. Disponible en: <https://sites.google.com/view/suf2024/inicio/poster-resumenes?authuser=0>

Glosario

Accuracy En español *Exactitud*. Métrica en PLN que mide el porcentaje de casos que el algoritmo acierta para todas las clases.

Anotación Proceso mediante el cual se asignan etiquetas a un documento o a segmentos de este (oraciones, enunciados, párrafos), de acuerdo con un esquema predefinido, para hacer el texto utilizable en tareas de Procesamiento del Lenguaje Natural.

Aprendizaje Automático En inglés, Machine Learning. Es un subcampo de la IA en el que los sistemas informáticos son capaces de comprender y aprender patrones para resolver problemas sin reprogramación externa.

Corpus Conjunto de datos para utilizado en PLN

Corpus anotado Conjunto de datos anotado bajo una especificación útil para el entrenamiento de algoritmos de aprendizaje automático.

F1-score Métrica en PLN que se define como la media armónica entre *Precision* y *Recall*.

Guía de anotación Documento que especifica el conjunto de etiquetas y criterios para aplicar el esquema de anotación a uno o varios documentos.

Inteligencia Artificial Es un campo de la ciencia dedicado al estudio, diseño y desarrollo de sistemas o agentes capaces de percibir información del entorno, aprender o adaptar su comportamiento y actuar de manera orientada al logro de determinados objetivos.

Precision En español, *Precisión*. Métrica que refiere a cuántas de las instancias que han sido clasificadas como positivas por un algoritmo de aprendizaje automática, realmente son ejemplos de la clase positiva.

Procesamiento del Lenguaje Natural Es una subdisciplina de la IA que intenta resolver con computadoras tareas vinculadas al lenguaje humano.

Prompt Es una indicación o instrucción que un usuario introduce en un modelo de lenguaje para conseguir que haga algo.

Recall En español, *recuperación* o *exhaustividad*. Métrica en PLN que mide la proporción de los casos positivos reales el algoritmo es capaz de detectar.

APÉNDICES

Apéndice A

Lisado de palabras pertenecientes
a la categorías de la TBA en
Física.

Tabla A.1: Palabras asociadas a las categorías de la TBA utilizadas en los informes de laboratorio. Fuente: Elaboración propia.

Recordar	Comprender	Aplicar	Analizar	Evaluar	Crear
recordar	abstraer	aplicar	analizar	apoyar	adaptar
denominar	categorizar	conectar	atribuir	argumentar	añadir
identificar	clarificar	desarrollar	categorizar	concluir	componer
localizar	clasificar	ejecutar	comparar	corroborar	construir
memorizar	comprender	elegir	considerar	criticar	corregir
mostrar	destacar	emplear	contrastar	debatir	crear
reconocer	ejemplificar	implementar	decidir	defender	diseñar
resaltar	explicar	medir	delinear	demostrar	elaborar
listar	extrapolar	planear	descomponer	detectar	formular
contar	generalizar	registrar	despreciar	establecer	generar
definir	ilustrar	usar	destacar	estimar	idear
	inferir	utilizar	diferenciar	evaluar	innovar
	informar	calcular	distinguir	formular	inventar
	interpretar	determinar	examinar	justificar	mejorar
	notar	despejar	idealizar	juzgar	modificar
	parafrasear	hallar	inspeccionar	probar	producir
	predecir		interpretar	proponer	transformar
	resumir		investigar	recomendar	trazar
	revisar		modelar	respaldar	hipótesis
	observar		organizar	revisar	
			preguntar	validar	
			priorizar	valorar	
			razonar		
			seleccionar		
			suponer		

Apéndice B

Palabras y secuencias de palabras asociadas con las diferentes categorías de la TBA.

COMPRENDER		
Condición o suposición	Discutiendo o planificando	Confirmando
al ser si este fuera el caso suponemos que ya que de no ser así ya que al hacer esto ya que por (el teorema)	discutimos diremos que objetivo para poder hacer Buscando ver con el fin de de manera de poder considerar	por lo que confirmar se precisa esperado
APLICAR		
Calculando	Describiendo acciones	Aplicando conceptos
estudiar despejar fórmula hallar promedio incertidumbre	tomar manipular minimizar reducir	a partir de fundamentos teóricos cantidad de muestras es grande siguiendo (un criterio)
ANALIZAR		
Analizando	Modelando	Condicional o causal
concordante diferencia eliminar ajustar observación observar constatar despreciar asumir variación variar	ideal inextensible insignificante puntual los datos siguen	lo que indica / lo que muestra se puede ver notar, se nota, notamos verificamos que quiere decir podemos deducir que por lo tanto / por ende nos lleva a que podemos identificar podemos dar cuenta nos dimos cuenta que se debe / deber a / se deba a causa de puesto que posiblemente por aunque
EVALUAR		
Comparando	Concluyendo	Decidiendo
la gran diferencia entre nos parece (más preciso, más adecuado)	implica concluir podemos decir que creemos que	deberíamos decidir elegir

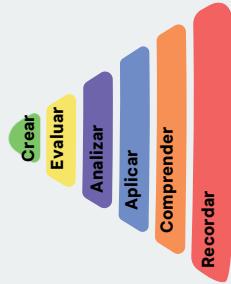
Apéndice C

Descripción de las categorías y
procesos cognitivos de la TBA
aplicado a la Física Experimental.

DIMENSIONES DEL PROCESO COGNITIVO

Taxonomía de Bloom adaptada a Física Experimental

Descripción de las categorías de la Taxonomía "revisada" de Bloom (Anderson y Krathwohl, 2001)

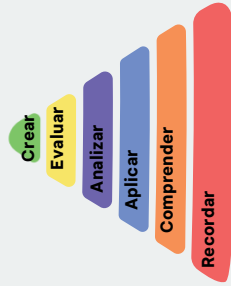


CATEGORÍA	DESCRIPCIÓN	SUBCATEGORÍA	DESCRIPCIÓN	EJEMPLOS
RECORDAR	Recupera conocimientos relevantes de la memoria a largo plazo.	Reconocer o Identificar.	Reconoce materiales utilizados describiendo el dispositivo experimental.	"El dispositivo experimental está compuesto por un soporte vertical de aluminio, con una tanza atada al cual se unieron dos semicírculos en la parte superior mediante una cinta adhesiva".
		Recuperar información. Recordar.	Identifica materiales utilizados mientras describe el proceso experimental.	"Se soltó un bloque macizo de aluminio desde la parte superior del sistema, deslizando sobre la superficie de madera y terminando su recorrido sobre el amortiguamiento."
COMPRENDER			Recupera información de la memoria a largo plazo citando, reconociendo o reconociendo conceptos teóricos importantes.	"La ley de los grandes números indica que para un número grande de medidas el promedio tiende al valor esperado μ y la desviación de los datos S_n tiende a la desviación standard."
		Interpretar.	Convierte información de una forma de representación a otra.	"En este gráfico, se puede ver claramente como para ángulos menores o iguales que 10° el período permanece casi constante..."
		Ejemplificar	Ir más allá de lo superficial explicando un concepto que ha de aplicarse a la actividad experimental.	"Siendo ℓ el largo de la tanza más el radio de la masa y T el período determinado anteriormente."
		Clasificar	Incluir algo dentro de un grupo más grande reconociendo que pertenece a cierta categoría.	Se da un ejemplo específico de un concepto.
		Resumir	Parafrasear, seleccionar puntos importantes.	
				"Se calcularán los valores del coeficiente de rozamiento cinético, con sus respectivas incertidumbres, a partir de dos métodos distintos"

DIMENSIONES DEL PROCESO COGNITIVO

Taxonomía de Bloom adaptada a Física Experimental

Descripción de las categorías de la Taxonomía "revisada" de Bloom (Anderson y Krathwohl, 2001)

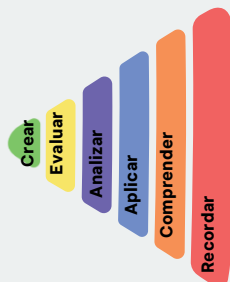


CATEGORÍA	DESCRIPCIÓN	SUBCATEGORÍA	DESCRIPCIÓN	EJEMPLOS
COMPRENDER	Construye significado a partir de instrucciones, incluida la comunicación oral, escrita o gráfica.	• Inferir, predecir, concluir.	Concluye, extrapola, predice un resultado esperado a partir de la aplicación de ciertos conocimientos previos.	"En cuanto a la forma de ambos, por la Ley de los grandes números, se espera que los datos sigan una distribución normal."
	Construye conexiones entre el "nuevo" conocimiento y el previo.	• Comparar	Detectar correspondencias entre dos objetos, ideas o similares. Detectar similitudes y diferencias con un conocimiento familiar. Comparar por analogía.	"En cuanto a los valores de las aceleraciones gravitatorias calculadas, ambos métodos dieron resultados parecidos."
		• Explicar	Construir un modelo de causa y efecto de un sistema. Realizar consideraciones experimentales mientras se construye un modelo.	"Habiendo obtenido el período por medio de mediciones realizadas con el cronómetro y con el sensor, podemos comparar el valor experimental con el valor de referencia que fue calculado teóricamente"
APLICAR				"Para esto se suelta el cuerpo de forma que las únicas fuerzas que actúen en el proceso sean el peso, la normal y la fricción"
				"La frecuencia angular depende únicamente del valor de la aceleración gravitatoria y el largo del hilo utilizado en el péndulo, ambos constantes durante la práctica."
				"En esta oportunidad solamente se utilizó el fotosensor para determinar el período de oscilación y se extrajo un total de 50 tomas del período."
				"Por lo que al sustituir el valor de 'a' en la ecuación 3 se obtiene el valor del coeficiente de rozamiento cinético..."
		• Ejecutar	Aplicar un procedimiento a una tarea familiar. Medir, graficar, calcular, determinar la incertidumbre.	"Posterior a la medición, se realizó un análisis estadístico de los datos obtenidos por el cronómetro y el fotosensor y se eliminaron algunos de los datos obtenidos."
				"Se volvió a realizar el histograma con su correspondiente gaussiana luego de descartar los valores."

DIMENSIONES DEL PROCESO COGNITIVO

Taxonomía de Bloom adaptada a Física Experimental

Descripción de las categorías de la Taxonomía "revisada" de Bloom (Anderson y Krathwohl, 2001)

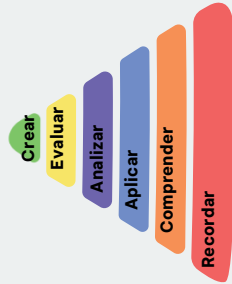


CATEGORÍA	DESCRIPCIÓN	SUBCATEGORÍA	DESCRIPCIÓN	EJEMPLOS
APLICAR	Realizar o usar un procedimiento en una determinada situación.	Implementar	Aplicar un procedimiento a una tarea desconocida, utilizar un método nuevo (no habitual) de registro de mediciones que impliquen un aprendizaje o seleccionar un procedimiento para una situación desconocida.	"Por lo tanto, a partir de la pendiente de la recta que mejor ajusta los valores, puede obtenerse el valor de la velocidad del sonido en aire."
		Diferenciar	Destaca diferencias entre procedimientos, ventajas y desventajas.	"La elección de temperaturas iniciales del agua para estos cálculos fueron elegidas de forma arbitraria, intentando cubrir un rango de temperaturas que en principio consideramos suficiente."
ANALIZAR	Dividir el material en sus partes constituyentes y determinar cómo se relacionan las partes entre sí y con una estructura o propósito general.			"Esta diferencia se la podemos atribuir a varios factores, entre ellos, el hecho de que el vaso térmico utilizado no sea lo suficientemente aislante como para mantener el intercambio de calor nulo entre nuestro sistema y su entorno."
			Encuentra cómo incide el procedimiento en los resultados.	"A tales efectos, está claro que constantemente podía ingresar o salir aire de nuestro sistema, que supusimos aislado."
		Organizar	Buscar hallazgos. Determinar cuáles elementos encajan, ajustan o funcionan dentro de una estructura o modelo.	"Se observa que para los ángulos comprendidos en el rango [5º;10º] se cumple tal hipótesis de forma que el periodo de oscilación es independiente del ángulo; a partir de los 10 grados comienza a aumentar el periodo, marcándose ahí un límite de este modelo propuesto.."
			Suposiciones realizadas para que ajusten al modelo.	"En el Gráfico 1 se puede observar que la hipótesis de pequeñas oscilaciones para los ángulos que toman valor 5º y 10º es correcta ya que el comportamiento del periodo en función del ángulo dentro de los primeros valores de theta es lineal."
		Atribuir	Expone motivos para seguir cierto procedimiento fundamentando la incidencia en los resultados.	"Para este análisis no se consideraron las muestras tomadas por el barrido de 50 a 800 Hz debido a que estas presentaban un mayor número de irregularidades."

DIMENSIONES DEL PROCESO COGNITIVO

Taxonomía de Bloom adaptada a Física Experimental

Descripción de las categorías de la Taxonomía "revisada" de Bloom (Anderson y Krathwohl, 2001)



CATEGORÍA	DESCRIPCIÓN	SUBCATEGORÍA	DESCRIPCIÓN	EJEMPLOS
EVALUAR	Emitir juicios basados en criterios y estándares.	• Comprobando	Detecta inconsistencias dentro del proceso o resultados.	"Bajo este criterio, como muestra la fig.5, todas las medidas obtenidas son confiables."
			Evalúa la efectividad de un procedimiento mientras que se implementa.	"Esto parece respaldar la hipótesis que supone que las condiciones bajo las cuales se realizó el experimento se mantuvieron constantes y que cada medida fue independiente."
			Determina si los resultados apoyan o refutan una hipótesis o si el material presentado tiene aspectos que se contradicen entre sí.	"De todas formas, el valor de referencia puede variar sustancialmente con respecto los resultados obtenidos posiblemente a la diferencia en las condiciones de las medidas experimentales de los coeficientes, como por ejemplo la humedad."
		• Criticando	Emite un juicio sobre el valor o calidad de cierto procedimiento, resultado o idea.	"De esta experiencia, encontramos que la definición de la temperatura de equilibrio es una dificultad al momento de llevar a cabo el experimento."
			Identifica puntos fuertes y débiles, realiza sugerencias de mejora.	"La implementación de termómetros digitales puede mejorar significativamente la precisión de las mediciones de temperatura."
CREAR	Juntar elementos para formar un todo coherente o funcional, reorganizar elementos en un nuevo patrón o estructura.	• Generar	Proponer hipótesis alternativas.	
		• Planificar	Diseñar un procedimiento para ejecutar una tarea.	
		• Producir	Inventar un producto	

Apéndice D

Desglose de las áreas clave en los cursos de laboratorio y códigos asociados.

Código	Construcción del conocimiento	Código	Modelado	Código	Diseño de experimentos	Código	Desarrollo de habilidades técnicas y prácticas en el laboratorio	Código	Análisis y visualización de datos	Código	Comunicación de la Física
CC1	Generar preguntas científicas que deseen explorar, diseñando preguntas y experimentos apropiados.	M1	Comprender cualitativa y cuantitativa del sistema a modelar.	DE1	Diseñar un procedimiento para probar un modelo o hipótesis, o realizar la medición de algo desconocido.	TPL1	Utilizar dispositivos de medición y realizar medidas coherentes.	AVD1	Usar métodos estadísticos para analizar datos.	CF1	Presentar argumentos razonados y respaldados por evidencia experimental.
CC2	Revisar preguntas que estén mal diseñadas o no sean comprobables.	M2	Elegir el marco conceptual adecuado para la situación física que se está modelando.	DE2	Realizar una revisión bibliográfica del modelo científico para mejorar el diseño experimental.	TPL2	Comprender/identificar las limitaciones e incertidumbres asociadas a los dispositivos de medición.	AVD2	Seleccionar métodos de representación de datos de forma adecuada para extraer cantidades físicas de los parámetros de ajuste.	CF2	Utilizar varios elementos estándar de comunicación técnica como gráficos, tablas, resultados numéricos con incertidumbres, diagramas.
CC3	Idear modelos o hipótesis falsables para explicar características observables en la naturaleza.	M3	Utilizar múltiples representaciones de modelos utilizando descripciones verbales, escritas, modelos matemáticos, computacionales, gráficos o diagramas.	DE3	Construir y/o montar el dispositivo experimental para realizar mediciones.	TPL3	Elegir un dispositivo apropiado para realizar la medición.	AVD3	Extraer información de los gráficos interpretando críticamente la validez y limitaciones de los datos presentados.	CF3	Utilizar vocabulario técnico apropiado para el contenido y los materiales o instrumentos de laboratorio.
CC4	Describir observaciones de manera clara, precisa y concisa.	M4	Reconocer las limitaciones, suposiciones y limitaciones del modelo, así como la incertidumbre que puede introducirse debido a esto.	DE4	Evaluar y/o mejorar el dispositivo experimental. Sugerir formas de mejora del diseño experimental.	TPL4	Cuantificar la incertidumbre tanto de dispositivos analógicos como digitales.	AVD4	Representar información de forma adecuada en un gráfico y extraer información del mismo realizando los ajustes apropiados.	CF4	Utilizar métodos de escritura y presentación científica pero no necesariamente al estilo de revista o presentaciones en conferencias.
CC5	Identificar los conceptos de física más importantes de un experimento.	M5	Calibrar de los instrumentos para su uso.	DE5	Reconocer las limitaciones del diseño experimental como calidad de datos, costo, tiempo, equipo disponible y fuentes de incertidumbre.	TPL5	Desarrollar de habilidades prácticas manuales, nivelar, alinear, organizar, seleccionar pesos, tensiones adecuados, teniendo en cuenta el impacto en la recolección de datos.	AVD5	Cuantificar las incertidumbres de sus resultados y realizar la propagación de incertidumbres en cálculos.	CF5	Comunicar resultados de manera ética y efectiva en formas orales o escritas.
CC6	Elaborar argumentos e identificar tendencias basadas en datos experimentales.	M6	Identificar errores sistemáticos o sesgos producidos por una investigación experimental.	DE6	Diagnosticar y solucionar problemas básicos durante el montaje experimental.	TPL6	Utilizar herramientas comunes como software de recolección de datos.	AVD6	Comparar sus resultados con modelos matemáticos, computacionales, o simulaciones.	CF6	Identificar las conexiones lógicas entre el contexto teórico, la evidencia experimental y sus argumentos.
CC7	Transferir conocimiento entre diferentes contextos, aplicando los mismos en diferentes situaciones.	M7	Utilizar modelos para predecir resultados.	DE7	Trabajar en fomar colaborativa para diseñar y construir el experimento.	TPL7	Registrar observaciones llevando un cuaderno de laboratorio.			CF7	Interpretar componentes de comunicación técnica como vocabulario, tablas, figuras, resultados numéricos, gráficos.
CC8	Divulgar los resultados de manera profesional y ética como científicos, teniendo una conducta responsable (evitar plagio).	M8	Evaluar si determinada cantidad física es razonable dado el sistema que se está modelando.			TPL8	Realizar una descripción del aparato y el procedimiento de medición.			CF8	Críticar sus propias presentaciones, tanto por la calidad de argumentos científicos como por el estilo de laboratorio como un registro para explicar los detalles de su trabajo, en cualquier resumen escrito.
						TPL9	Registrar datos y analizar los mismos.			CF9	Usar su cuaderno de laboratorio como un registro para explicar los detalles de su trabajo, en cualquier resumen escrito.
						TPL10	Aplicar la física a sistemas familiares, dispositivos de consumo, médicos o procesos industriales.			CF10	Discutir las ideas en pequeños grupos como parte del proceso científico en general.

Figura D.1: Desglose de objetivos dentro de cada una de las áreas clave en los cursos de laboratorio según las recomendaciones de la AAPT y códigos asociados.

Apéndice E

Guía de anotación (segunda versión).

NOTA: Se codifican enunciados de las secciones Metodología, Resultados y Análisis y Conclusiones. No se codifica pie de página, leyendas de figuras, ecuaciones y/o cálculos, resultados numéricos.

RECORDAR				
Nombre de la categoría	Descripción de la categoría			
	Recupera conocimientos relevantes de la memoria a largo plazo.			
	Subcategoría	Código HC	Descripción	Comentarios
Criterio de inclusión de la categoría. Se codifica si está presente alguno de los siguientes aspectos.	Describe dispositivo sin toma de medidas.	TPL8	Realiza una descripción de los dispositivos utilizados y/o cómo se ejecutó el armado o trabajo con el mismo.	El estudiante puede realizar un esbozo del dispositivo experimental y/o describirlo con palabras.
	Enuncia teoría	CC5	Identifica los conceptos de física más importantes en un experimento.	Recupera información acerca de un concepto teórico, expone o declara un concepto.
Ejemplos de inclusión	Describe dispositivo sin toma de medidas.	"El segundo era un cronómetro utilizado con el mismo objetivo que el sensor, medir el período de pequeñas oscilaciones"		
		"Con un ángulo pequeño, se dejó oscilar el péndulo libremente durante 50 períodos consecutivos".		
	Enuncia teoría	"La incertidumbre del tipo A, que son aquellas que son estudiadas mediante análisis estadístico, y por otro lado las del tipo B, que son las que se asocian a las demás cosas, como la incertidumbre de un instrumento debido a su apreciación".		
Criterio de exclusión de la categoría. La categoría no se codifica si...	Toma de medidas	Si en el enunciado se evidencia la toma de medidas entonces revisar la categoría Aplicar.		
Ejemplos de exclusión	"Las mediciones se realizaron con el fotosensor y el cronómetro manual en simultáneo"			Código: TPL1 – Utiliza instrumentos de medición para realizar medidas coherentes.
	"Se colocó el sensor en un mástil con forma de "L", firmemente sujetado a la mesa con l ayuda de una morsa, y de este se colgó la masa utilizando tanza de nylon y sujetando con cinta de papel para evitar cualquier tipo de deslizamiento relativo entre la tanza y el mástil".			Código: TPL5 – Desarrolla algunas habilidades prácticas de laboratorio manuales, se evidencia al incluir la sección "para evitar cualquier tipo de deslizamiento relativo", es decir, que tiene en cuenta el posible impacto que puede existir en la recolección de datos.

Nombre de la categoría	COMPRENDER				
Descripción de la categoría	Construye significado a partir de instrucciones, incluyendo la comunicación oral , escrita o gráfica. Establece conexiones entre el “nuevo” conocimiento y el previo.				
Criterio de inclusión de la categoría. Se codifica si está presente alguno de los siguientes aspectos.	Subcategoría	Código HC	Descripción	Comentarios	
	Modelado Teórico-Experimental	M1	Vincula la teoría y el experimento, logra una comprensión cualitativa y cuantitativa de un sistema.		
		M2	Seleccionan un marco conceptual adecuado a la situación que se está modelando.	Los estudiantes pueden usar la conservación de la energía o las Leyes de Newton como un marco conceptual para describir el movimiento de un objeto deslizándose por una pendiente.	
	Identificación de Restricciones en Modelos y dispositivos de medición.	M4	Reconoce o comprende las suposiciones, limitaciones y simplificaciones del sistema a modelar.	Un estudiante que modele el movimiento de un objeto sin fricción debe ser capaz de identificar que la fricción ha sido omitida y comprender cuáles son las consecuencias de hacer esta simplificación.	
		M6	Reconoce las limitaciones de un modelo, incluidas las consideraciones de las incertidumbres en las mediciones y las limitaciones de los dispositivos de medición.	Sólo identifica limitaciones, si realiza juicios de valor o críticas revisar la categoría evaluar.	
	Expresión y representación de la Física	CC4	Describe las observaciones experimentales de manera clara, precisa y concisa.	Describe lo que se observa en un gráfico, es decir, describe la figura	
		CF7	Interpreta componentes estándar de la comunicación técnica y la Física como el vocabulario, los resultados numéricos con incertidumbre, las tablas y los gráficos.	Realiza una descripción de lo que se presenta en un gráfico o se observa en tablas, si analiza los resultados revisar la categoría analizar.	
Ejemplos de inclusión	Ejemplifica	“Esto es importante ya que, por ejemplo, para que las medidas sigan una distribución gaussiana, la medición debe realizarse en iguales condiciones experimentales cada una de las veces.”			
	Infiere	“ .. al observar el histograma con la curva teórica superpuesta podemos inferir de forma cualitativa que los datos se distribuyen de manera normal”			
	Interpreta	“Al observar la figura se ve que el período no decae con el tiempo sino que más bien parece oscilar en torno al valor XX”			
Criterio de exclusión de la categoría. La categoría no se codifica si..	Es posible confundir la comprensión con la evaluación de la situación, revisar la categoría evaluación.				
Ejemplos de exclusión	“Por otro lado, tomar un hilo muy largo provocaba que el péndulo dejará de oscilar en un plano y comenzará a desplazarse en otra dirección”.			Código: DE5 - Reconoce limitaciones del diseño experimental como calidad de datos, costo, tiempo, equipo disponible y condiciones de seguridad.	

Nombre de la categoría	APLICAR				
Descripción de la categoría	Realiza o ejecuta un procedimiento en determinada situación.				
Criterio de inclusión de la categoría. Se codifica si está presente alguno de los siguientes aspectos.	Subcategoría	Código HC	Descripción	Comentarios	
	Cuantifica incertidumbre	TPL4	Cuantifica la incertidumbre tanto de dispositivos analógicos como digitales.	Los estudiantes pueden consultar la incertidumbre de un instrumento particular en el manual o en las especificaciones en línea.	
		AVD5	Cuantifica o explica cómo puede cuantificar la incertidumbre de sus resultados incluyendo el estudio de la propagación de la incertidumbre.	Los estudiantes pueden realizar el cálculo, explicar cómo lo realizaría, explicar qué incertidumbres debe considerar para realizar la propagación de la incertidumbre. Calcula la desviación estándar.	
	Habilidades de Laboratorio y Medición	TPL1	Utiliza instrumentos de medición para realizar medidas coherentes.	Los estudiantes deberían, como mínimo, ser capaces de medir el tiempo, la distancia, la masa, la temperatura, la diferencia de potencial y la intensidad de corriente.	
		TPL5	Desarrolla algunas habilidades prácticas de laboratorio manuales.	Los estudiantes pueden construir y analizar circuitos simples, nivelar o alinear aparatos, y organizar abrazaderas o soportes que sean adecuados para los pesos o tensiones involucradas, todo teniendo en cuenta el posible impacto que estas elecciones pueden tener en la recolección de datos.	
		TPL6	Utiliza herramientas comunes de recolección de datos, como por ejemplo videos, interfaz o software para extraer datos físicos.	Los estudiantes podrían usar software de seguimiento de movimiento en video para estudiar un modelo del salto alto de un atleta.	
		M5	Calibra instrumentos para su uso.	Por ejemplo escalar un video es una calibración (en este caso, escalar los píxeles a una unidad de longitud).	
	Comunicación científica	M3	Utiliza múltiples representaciones de modelos.	En estas representaciones se incluyen descripciones verbales y escritas, modelos analíticos (matemáticos, cálculos), modelos computacionales, modelos gráficos o diagramas.	
		CF3	Utiliza elementos estándar de comunicación científica.	Usa vocabulario técnico apropiado para el contenido de física y el aparato utilizado en el laboratorio introductorio. Expresa datos de medición y análisis con cifras significativas e incertidumbre. Presentar datos en tablas y gráficos. Realizan bocetos o diagramas básicos del aparato o sistema.	
	Ejemplos de inclusión	Calcula	"Luego, al despejar la longitud de la expresión de T , se obtuvo que el largo del hilo debe ser de aproximadamente un metro". Código: M3.		
Cuantifica		"Se calculó el promedio y la desviación estándar obteniéndose..."			
Hab. laboratorio		"Las mediciones se realizaron con el fotosensor y cronómetro manual en simultáneo." Se codifica como TPL1			
Criterio de exclusión de la categoría. La categoría no se codifica si...	No se codifica como aplicar si únicamente describe procedimiento sin realizar toma de medidas.				
Ejemplos de exclusión	"Con un ángulo pequeño, se dejó oscilar el péndulo libremente durante 50 períodos consecutivos".			Se codifica como TPL8 - Realiza una descripción de los dispositivos utilizados y/o cómo se ejecutó el armado o trabajo con el mismo.	

ANALIZAR					
Nombre de la categoría		Divide el material en sus partes constituyentes y determina cómo se relacionan las partes entre sí, con una estructura o un propósito general.			
Descripción de la categoría		Subcategoría	Código HC	Descripción	Comentarios
Criterio de inclusión de la categoría. Se codifica si está presente alguno de los siguientes aspectos.	158	Diferencia, compara. Determina cuáles elementos encajan, ajustan o funcionan dentro de una estructura o modelo.	AVD3	Extrae información de los gráficos analizando críticamente la validez y limitaciones de los datos presentados.	
			AVD1	Analiza datos utilizando un software o realiza un análisis estadístico.	Los estudiantes deberían ser capaces de realizar un análisis estadístico básico, como por ejemplo de los resultados de la media y la desviación estándar.
			AVD2	Representa datos de forma adecuada para extraer cantidades físicas de los parámetros de ajuste.	Realiza ajustes de curvas básicos y relacionar los parámetros de ajuste con cantidades físicas.
			AVD6	Compara sus resultados con modelos matemáticos, computacionales o simulaciones.	
		Determina relación entre las partes constituyentes	CF6	Realiza conexiones lógicas entre el contexto teórico, la evidencia experimental y sus explicaciones, fundamentos o razonamientos.	
		Analiza y predice	M7	Utiliza modelos para predecir resultados.	
Ejemplos de inclusión		Código CF6	"Esto es concordante con la hipótesis de realizar las medidas bajo las mismas condiciones experimentales"		
		Código AVD1	"El criterio empleado para el descarte de medidas fue que los datos disten más de dos desviaciones estándar del promedio muestral".		
Criterio de exclusión de la categoría. La categoría no se codifica si...		Además de analizar en el enunciado se puede realizar una evaluación de los resultados por ejemplo afirmando que una medida es más precisa que otra. Elaborar una conclusión es característico de evaluar.			
Ejemplos de exclusión		"Además comparando la exactitud medida en base al error relativo en la medida, también se puede concluir, en este caso, que la medida con el sensor es más precisa y lo mismo para la exactitud"			Si el enunciado hubiera sido "Se encontró que la medición con el sensor presenta menor error relativo." entonces sería Analizar porque no hay un juicio crítico.

Nombre de la categoría	EVALUAR			
Descripción de la categoría	Emitir juicios basados en criterios o estándares.			
	Subcategoría	Código HC	Descripción	Comentarios:
Criterio de inclusión de la categoría. Se codifica si está presente alguno de los siguientes aspectos.	Evalúa sobre el dispositivo experimental.	TPL3	Elige un dispositivo apropiado para realizar la medición comprendiendo las limitaciones y/o incertidumbres asociadas a éste.	Al realizar una medición de longitud, los estudiantes eligen un dispositivo apropiado (por ejemplo, regla, calibre o micrómetro).
		DE4	Diagnostica, realiza o sugiere mejoras al dispositivo experimental.	Sugieren mejoras a futuro, posteriormente a la realización del experimento.
		DE5	Reconoce limitaciones del diseño experimental como por ejemplo calidad de datos, costo, tiempo, equipo disponible y condiciones de seguridad.	
		DE6	Soluciona problemas durante el montaje experimental.	
	Evalúa resultados	M8	Evalúa si una cantidad física es razonable dado el sistema que se está modelando.	Estima, utilizando las unidades de medida apropiadas y los parámetros de entrada los parámetros de salida de un modelo.
	Concluye o critica con base en datos experimental es.	CC6	Sintetiza la información disponible, incluidos los resultados e incertidumbres, para llegar a una conclusión coherente y basada en los datos.	Construye argumentos e identifica tendencias basadas en observaciones controladas experimentalmente.
		CF1	Utilizar la síntesis de la información y los datos analizados para apoyar o refutar la hipótesis inicial. Evaluar la evidencia y desarrollar las conclusiones basadas en datos.	
Ejemplos de inclusión	Código M8	"Esto hace que el valor medido con el cronómetro y el fotosensor sean concordantes entre sí, y también con el valor de referencia por la aceleración gravitatoria."		
	Código DE5	"Encontramos que la definición de la temperatura de equilibrio es una dificultad al momento de llevar a cabo el experimento".		
Criterio de exclusión de la categoría. La categoría no se codifica si...	La evaluación implica emitir juicios basados en criterios y estándares, en ocasiones puede confundirse con comprender o analizar.			
Ejemplos de exclusión	"Con ángulos pequeños aproximadamente iguales se obtuvieron cincuenta medidas, ya que, debido al estudio estadístico de los datos, fue necesario un muestreo extenso sobre la población".			No se evidencia una crítica o juicio de valor acerca de los resultados o el experimento, se evidencia la comprensión del motivo por el cual debe realizar la toma de múltiples medidas. Si describe y conecta ideas en un nivel explicativo, es Comprender .

Nombre de la categoría	CREAR			
Descripción de la categoría	Juntar elementos para formar un todo coherente o funcional, reorganizar elementos en un nuevo patrón o estructura.			
	Subcategoría	Código HC	Descripción	Comentarios:
Criterio de inclusión de la categoría. Se codifica si está presente alguno de los siguientes aspectos.	Diseña, genera, construye dispositivos o procedimientos.	DE1	Diseña un procedimiento para probar un modelo o hipótesis, o para medir algo desconocido teniendo en cuenta la cantidad y precisión de los datos para obtener resultados reproducible y precisos.	
		DE3	Diseña y construye un dispositivo para realizar mediciones y probar un modelo o hipótesis.	
	Genera o desarrolla preguntas o hipótesis.	CC1	Genera preguntas científicas que deseen explorar diseñando preguntas y experimentos apropiados.	
		CC3	Desarrolla modelos o hipótesis falsables para explicar características observables en la naturaleza.	
Ejemplos de inclusión				
Criterio de exclusión de la categoría. La categoría no se codifica si...				
Ejemplos de exclusión				

Apéndice F

Guía de anotación, versión final.

Código	Descripción breve	Criterios de inclusión	Ejemplo de inclusión.	Observaciones	Ejemplo de exclusión.	Hab. AAPT
Aplicar: Realizar o usar un procedimiento en determinada situación. La categoría aplicar comienza con la ejecución estrecha y estructurada en la que el conocimiento procedimental se aplica casi de forma rutinaria. Luego continúa con una implementación cada vez menos estructurada, en la que debe seleccionarse el procedimiento para adaptarse a la situación.						
APL	Desarrolla habilidades propias de la medición en laboratorio. Nivel TBA: Aplicar	Describe instrumentos de medición utilizados o el procedimiento necesario para realizar medidas coherentes.	"Las mediciones se realizaron con el fotosensor y cronómetro manual en simultáneo. "			TPL1
			"Se realizó un histograma con los promedios de los periodos de 5 oscilaciones de 30 medidas diferentes en el mismo ángulo inicial, con el mismo instrumento de medida, en esta caso el Fotosensor. "			
		Se evidencia el desarrollo de algunas habilidades prácticas de laboratorio manuales.	"...se montó el soporte sobre la mesa de forma que queda lo mas vertical y firme posible... "			TPL5
		Utiliza herramientas comunes de recolección de datos, como por ejemplo videos, interfaz o software para extraer datos físicos.	"Para la recolección y el análisis de datos se utilizaron los programas LoggerPro y SciDAVis"			TPL6
		Calibra instrumentos para su uso.				M5
Analizar: Dividir el material en sus partes constituyentes y determinar relaciones entre las partes y/o con una estructura o propósito general.						
ANG	Analiza gráficos de medidas experimentales. Nivel TBA: Analizar	Analiza gráficos extrayendo información de los mismos.	"Se evidencia un patrón casi lineal en las mediciones, lo cual confirma nuestras expectativas al emplear el RTD"	Si bien en el enunciado no queda explícito que refiere a un gráfico, el enunciado anterior era: "En la siguiente imagen se puede ver el gráfico de la resistencia en función de la temperatura con los datos obtenidos por el RTD."		AVD3
		Analiza críticamente la validez y limitaciones de los datos presentados en el gráfico.	"El criterio empleado para el descarte de medidas fue que los datos disten más de dos desviaciones estándar del promedio muestral"	Es posible que en el enunciado no se mencione el gráfico, pero en la lectura de los enunciados en forma consecutiva se evidencia que se realizó previamente el mismo para llegar a realizar un análisis estadístico.		AVD3
		Analiza datos utilizando un software o un análisis estadístico sobre los datos representados en forma gráfica. (Por análisis estadístico básico se entiende aplicar mínimos cuadrados, resultados de la media y la desviación estándar.)	"Después procedemos a linealizar para aplicar el método de mínimos cuadrados , en caso del RTD ya tenemos una gráfica lineal , por lo que directamente aplicamos mínimos cuadrados:"	Extrae información de un gráfico, pero además analiza datos, este enunciado se codifica ANG y AND		AVD1
		Representa datos de forma adecuada para extraer cantidades físicas de los parámetros de ajuste.	"Ahora, para linealizar el gráfico del termistor representado tal que $Y = AX + D$ debemos aplicar el siguiente cambio de variable: "	Refiere a la información que puede extraer a partir del gráfico, pero también analiza datos al aplicar un cambio de variable. Se codifica ANG y AND		AVD2
		Realiza ajustes de curvas básicos y relaciona los parámetros de ajuste con cantidades físicas.	De la pendiente de la recta (M) se obtiene el valor de k1 según se puede ver en la Ec. 4.	Relaciona los parámetros de ajuste con cantidades física según el experimento realizado.		AVD2
		Compara sus resultados con modelos matemáticos, computacionales o simulaciones.	De este gráfico obtuvimos la pendiente $A = [4008, 13 \pm 36, 08]$ y término independiente $D = [2, 62 \pm 0, 01]$.	Compara los resultados por medio de un ajuste lineal. Cuando se refieren a la "pendiente" queda implícito que es de un gráfico ANG, además como indica lo valores (datos) se codifica AND.		AVD6

AND	Analiza datos o resultados experimentales. Nivel TBA: Analizar	Compara sus resultados con modelos matemáticos, computacionales o simulaciones.	El coeficiente de correlación R , lo que indica que el modelo lineal es bueno.	Cuando utiliza el coeficiente de correlación para valorar un ajuste entonces se codifica AND y ANG		AVD6
		Compara sus resultados con modelos matemáticos, computacionales o simulaciones.	"Posteriormente, para determinar el valor de b , se realizó un análisis de la pendiente del ajuste lineal del gráfico presentado en la Figura 6, que arroja el valor de según la Ec.6, operando se llega al valor de b expresado en la Tabla 3"	Cuando se refiere a la "pendiente" si se considera que la misma es extraída de un gráfico entonces se codifica como AND y ANG.		AVD6
		Analiza datos utilizando un software o análisis estadístico. Analiza para obtener por ejemplo la media, la desviación estándar.	Nuevamente ningún dato se encontraba fuera de la región definida por lo tanto no hubo que realizar un nuevo histograma.	Se está refiriendo al análisis estadístico de datos, sin analizar directamente un gráfico.		AVD6
		Analiza datos utilizando un software o análisis estadístico. Analiza para obtener por ejemplo la media, la desviación estándar.	"Linealizando los datos por método de mínimos cuadrados resulta que $a = (0, 811 \pm 0, 003)$ y $b = (5, 5 \pm 0, 2)C$."	Utilizar "mínimos cuadrados" se codifica como análisis de datos (AND), en caso de que en el enunciado se refiere a un gráfico, puede codificarse en ANG también.		AVD6
		Realiza conexiones lógicas entre el contexto teórico, la evidencia experimental y sus explicaciones, fundamentos o razonamientos.	"La principal observación es que los valores obtenidos en la realización de la experiencia no concuerdan exactamente con el valor real de la aceleración gravitatoria calculado teóricamente. "	Si habla de concordancia entre resultados o con un valor de referencia pero no hace referencia a la incertidumbre se codifica como AND y no como EIM.		CF6
		Utiliza modelos para predecir resultados.				M7

Evaluar: Emitir juicios basados en criterios y estándares; detectar la efectividad de un procedimiento a medida que se implementa; detectar lo apropiado de un procedimiento para un problema dado

EDE	Evalúa el diseño (procedimiento o montaje) experimental. Nivel TBA: Evaluar	Evalúa el diseño experimental, reconoce aciertos, ventajas y/o limitaciones en el diseño (procedimiento o montaje) experimental como por ejemplo calidad de datos, costo, tiempo, equipo disponible y condiciones de seguridad.	"Los valores experimentales de B , R_0 y α , que han sido obtenidos parecen estar lo suficientemente próximos a sus valores de referencia, indicando así que el procedimiento experimental utilizado es probablemente preciso. "	Realiza una evaluación positiva sobre el procedimiento experimental utilizado.		DE5
		Diagnostica, realiza o sugiere mejoras al dispositivo experimental y/o procedimiento. Sugiere mejoras a futuro, posteriormente a la realización	"Además, se sugiere que se podría mejorar la precisión de las mediciones de temperatura utilizando un termómetro más exacto para calibrar los instrumentos."	Sugiere mejoras a futuro.		DE4
		Identifica, propone y/o soluciona problemas durante el procedimiento experimental.	"Con la finalidad de cuantificar esta posible fuente de error se ideó un experimento sencillo que utiliza las relaciones que nos da la primera ley de la termodinámica."	El enunciado previo era "La principal fuente de error en la práctica fue debido al poco aislamiento térmico de nuestro calorímetro." Si bien no modifican el dispositivo, reconocen un problema y		DE6
		Evalúa el dispositivo utilizado y propone un dispositivo apropiado para realizar la medición comprendiendo las limitaciones y/o incertidumbres asociadas a estos.		Proponer un dispositivo implica más que tomar materiales dados para la realización de la práctica y utilizarlos.	"Como lo representa la Figura 2 primero se montó el soporte sobre la mesa de forma que queda lo mas vertical y firme posible, luego se ató un extremo de la tanza al soporte y el otro extremo al objeto esférico." Este enunciado se codifica como APL pero no realiza una evaluación y está	TPL3

EIM	Evalúa el impacto de la incertidumbre en la medición Nivel TBA: Evaluar	Sintetiza la información disponible, basándose en los resultados e incertidumbres, para llegar a una conclusión coherente.	Luego, se concluyó que k_2 y k_2 , estadístico son concordantes entre sí porque 27,63 se encuentra entre 27,00 y 28,40, entonces se puede afirmar que ambas formas de obtener la constante elástica son adecuadas en esta práctica.	Para este código es necesario que el enunciado se refiera a la incertidumbre en las medidas o el proceso de medición, no a la incertidumbre propia del instrumento.	"A la hora de ensamblar el dispositivo es imprescindible que las puntas de los tres sensores no estén en contacto entre sí ya que esto provocaría una medición errónea." Se refiere a los defectos del montaje experimental, no a la incertidumbre en las medidas realizadas. Se codifica EDE.	CC6
		Evalúa los resultados y emite un juicio teniendo presente la incertidumbre en la medición	En cuanto a R_0 ; por un lado, tenemos el valor de referencia para el RTD, que sería de $R_0 = 1000 \Omega$, comparándolo con el obtenido en la práctica ($R(T_0) = [1079 \pm 15]$) podemos ver como este puede ser considerado como concordante.	Si bien no aparece explícita la mención a la incertidumbre, se considera que al incluir la misma en el resultado experimental se está tomando en cuenta a la hora de emitir el juicio.		CF1
Nota:		Cada enunciado puede tener más de un código o ninguno.				
Ejemplos			Códigos encontrados.		Comentarios	
También se elaboró un histograma del cual se extrajeron valores de promedio y desviación estándar, estos se utilizaron como valores de constante elástica k_2 estadístico e incertidumbre respectivamente concluyendo que es una medida confiable por tener confiabilidad de 2,5%.			ANG AND EIM		Se indica con cada formato cómo se indentifica cada código en un enunciado.	
Reescribiendo la ecuación (2) podemos llegar a algo del tipo $Ax+B$ siendo A la pendiente de la recta y B la constante.			ANG		Se extraen datos de un gráfico y se analizan en función de un modelo matemático.	
Además, calculando el coeficiente de correlación a la recta obtenida, resulta que, $r = 0,9997789$, de este número podemos concluir que la relación lineal de esta recta es fuerte.			ANG AND		Analiza el gráfico, para obtener un "dato" que le permite concluir acerca de que tan bueno es el ajuste realizado.	
Finalmente, la práctica no solo cumplió con los objetivos establecidos, sino que también nos permitió tener un mejor entendimiento del fenómeno físico de la fricción aplicada a un plano inclinado.			No se encuentran códigos que apliquen a este enunciado		Se realiza una afirmación, pero no se detecta el desarrollo de habilidades prácticas de laboratorio, no se analizan datos ni gráficos, ni se realiza una evaluación de la incertidumbre o dispositivo experimental.	
Se realizaron 30 repeticiones para luego hacer un analisis estadístico tomando en cuenta las incertidumbres de cada instrumento de medida.			APL		Utiliza instrumentos de medición para realizar medidas coherentes, pero no corresponde el código EIM porque no realiza una evaluación acerca del impacto de la incertidumbre en la medición, solo menciona que la toma en cuenta.	
Con los datos obtenidos en la primera practica se graficó el periodo en función del ángulo inicial al cual se dejaba caer el péndulo.			No se encuentran códigos que apliquen a este enunciado		Si bien menciona un gráfico no realiza un análisis del mismo.	

Apéndice G

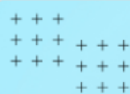
Lista de *Stop Words* en español.

Tabla G.1: Lista de *stopwords* en español

a	ahí	ahí	al
algo	alguien	alli	allá
allí	ante	aquel	aquí
aquél	aquí	asi	así
aun	aún	bajo	bien
cada	como	con	cual
cuál	cuán	cómo	de
del	demás	día	días
donde	dos	día	días
el	ella	ellas	ellos
en	entre	era	eran
eres	es	esa	esas
ese	eso	esos	esta
estas	este	esto	estos
está	fue	fui	hace
hay	hoy	la	las
le	les	lo	los
mal	mas	me	mi
mia	mias	mio	mios
mis	muy	más	mí
mía	mías	mío	míos
ni	no	nos	nuestra
nuestro	o	os	otra
otro	para	pero	poco
por	que	quien	quién
qué	se	sea	ser
si	sin	sino	son
su	sus	sí	tal
te	ti	tiene	tu
tus	tú	un	una
uno	vos	y	ya
yo	él	ésa	ésas
ése	esos	ésta	éestas
éste	éstos		

Apéndice H

Encuesta electrónica realizada a los docentes.



¿Qué habilidades científicas se evidencian en los informes de laboratorio de Física Experimental?



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



Habilidades Científicas en Física Experimental

El siguiente cuestionario está dirigido a docentes que dictan o han dictado el curso "Física Experimental 1" en la Facultad de Ingeniería. El objetivo es conocer su perspectiva sobre las habilidades científicas que los estudiantes desarrollan durante el curso, en particular aquellas que se evidencian en los informes de laboratorio. Este cuestionario forma parte de la tesis de Maestría en Física de la Prof. Lorena Pagella, bajo la tutoría de la Dra. Cecilia Stari (Instituto de Física) y la Dra. Aíala Rosá (Instituto de Computación). Su participación es muy valiosa para seguir profundizando esta investigación y no le tomará más de 15 minutos. ¡Muchas gracias por su colaboración!



Título

Nota: En este formulario se utiliza el término "los estudiantes" para hacer referencia a los y las estudiantes, sin intención de establecer una perspectiva de género. La elección responde únicamente a motivos de simplificación en la redacción.

A continuación, le solicitamos que ingrese su **dirección de correo electrónico**, en caso de que ^{*} sea necesario contactarlo/a para aclarar o ampliar alguna información proporcionada. *Esta dirección será utilizada exclusivamente con fines relacionados con la presente investigación y no será compartida con terceros bajo ninguna circunstancia. Todos los datos aportados serán debidamente anonimizados y se utilizarán únicamente en el marco del desarrollo de esta investigación.*

Texto de respuesta corta



1. ¿Cuántas veces ha dictado el curso Física Experimental 1 en FING? ^{*}

- ☐ 1 vez
- ☐ 2 o 3 veces
- ☐ entre 4 y 6 veces
- ☐ más de 6 veces

2. Indique si ha dictado el curso "Física Experimental 1" en los siguientes años : *

Seleccione todas las opciones que correspondan.

- ☐ 2023
- ☐ 2024
- ☐ 2025
- ☐ No he dictado cursos en esos años

3. Identifique y describa los elementos del curso que, desde su perspectiva, favorecen la motivación de los estudiantes. *

Texto de respuesta larga

4. ¿Qué estrategias de enseñanza utiliza habitualmente durante el desarrollo de una clase de Física Experimental I? *

Seleccione todas las que correspondan.

- ☐ Introducción teórica al comienzo de la clase
- ☐ Discusión del procedimiento experimental con todo el grupo
- ☐ Discusión del procedimiento en pequeños grupos
- ☐ Intervenciones del docente durante la práctica (acompañamiento, retroalimentación, sugerencias, etc.)
- ☐ Trabajo en grupos sin intervención del docente durante la práctica
- ☐ Puesta en común de los resultados al final de la clase
- ☐ Puesta en común de los resultados en la clase siguiente
- ☐ Otro:

...

5. ¿En qué momento considera más adecuado realizar la discusión acerca de cómo analizar los resultados? *

Seleccione una sola opción.

- ☐ Al comienzo de la clase, junto con la discusión teórica y del procedimiento experimental
- ☐ Luego del armado del experimento, antes de comenzar a medir
- ☐ No realizo esta discusión porque el protocolo ya lo explica
- ☐ Otro:

...

6. ¿En qué momento suele realizar la puesta en común de los resultados entre los grupos o con toda la clase? *

- ☐ Al final de la clase
- ☐ Al comienzo de la siguiente clase
- ☐ No realizo puesta en común de los resultados
- ☐ Otro:

...

7. De acuerdo a su percepción, seleccione los seis (6) objetivos de aprendizaje que considera más importantes en un curso de Física Experimental I. *

Los objetivos que se presentan a continuación están tomados del programa oficial del curso.

Seleccione hasta seis opciones.

- ☐ Utilizar instrumentos de medición y reconocer las características de su funcionamiento
- ☐ Implementar el montaje experimental siguiendo los lineamientos
- ☐ Aplicar una metodología para medir magnitudes físicas
- ☐ Seleccionar parámetros asociados a un sistema físico de interés.
- ☐ Utilizar herramientas de análisis de datos experimentales
- ☐ Interpretar resultados utilizando herramientas estadísticas o de análisis de datos.
- ☐ Realizar gráficos

- ☐ Analizar gráfico de medidas experimentales y modelos teóricos
- ☐ Analizar los resultados experimentales e interpretarlos en comparación con el modelo físico propuesto.
- ☐ Establecer conclusiones sobre el diseño experimental y sus resultados
- ☐ Analizar los puntos débiles del montaje y/o el procedimiento experimental.
- ☐ Definir mejoras para el montaje y/o procedimiento experimental.
- ☐ Identificar las fuentes de incertidumbre en la realización de medidas
- ☐ Evaluar el impacto de la incertidumbre en la medición.
- ☐ Presentar resultados de forma escrita en un formato preestablecido.
- ☐ Presentar y exponer de forma oral los resultados obtenidos.

8. Si considera que hay algún otro objetivo relevante que no está incluido en los ítems anteriores, por favor escríbalo a continuación:

*

Si no lo hay, escriba 'No'

Texto de respuesta larga

9. Seleccione cinco (5) de las siguientes habilidades que, a su juicio, son prioritarias para que los estudiantes desarrollen durante el curso de Física Experimental I. *

- ☐ Seleccionar el marco conceptual para la situación física que se está modelando, vincular la teoría y el experimento comprendiendo las suposiciones, limitaciones y simplificaciones propias del modelo.
- ☐ Montar o construir un dispositivo experimental para realizar mediciones.
- ☐ Utilizar correctamente instrumentos de medición y herramientas de recolección de datos para realizar mediciones coherentes.
- ☐ Cuantificar la incertidumbre y expresarla correctamente.
- ☐ Analizar datos, lo que implica el análisis estadístico básico y el uso de diversas herramientas informáticas para el análisis
- ☐ Representar datos en forma gráfica para analizarlos y extraer información.
- ☐ Evaluar el diseño experimental, identificar posibles problemas y proponer mejoras
- ☐ Evaluar cómo la incertidumbre influye en la validez de los resultados obtenidos identificando las fuentes de error
- ☐ Redactar informes utilizando lenguaje técnico adecuado y específico de la Física, expresar las medidas con cifras significativas e incertidumbre, elaborar diagramas de los dispositivos utilizados y presentar datos en tablas y gráficos de manera clara.
- ☐ Evaluar y reflexionar sobre su propio trabajo, criticar sus propias presentaciones, entregas o informes, tanto por la calidad de los argumentos como por el estilo.

10. ¿Considera que en este fragmento de informe se pueden identificar algunas de las siguientes habilidades?

A continuación, encontrará una posible conclusión dividida en enunciados, junto con cinco habilidades. Si cree que algún enunciado refleja una de estas habilidades, marque la casilla correspondiente en la columna indicada. En caso de no identificar ninguna, deje las casillas vacías.

En esta práctica medimos el período del péndulo utilizando dos instrumentos: un cronómetro manual y un fotosensor. Los resultados mostraron que ninguno de los valores de aceleración calculados coincidió con el valor teórico ni entre sí. Sin embargo, el valor obtenido con el fotosensor fue el más preciso, ya que minimiza el error humano presente al usar el cronómetro. Además, verificamos que hasta los 25° de amplitud se cumple la hipótesis de ángulos pequeños: el período se mantuvo prácticamente constante. A partir de ese valor, el período comenzó a variar, evidenciando una desviación respecto al modelo ideal del péndulo simple. Por último, el análisis del histograma mostró un buen ajuste a una curva gaussiana, lo que indica que las mediciones siguen una distribución normal, como era esperable.

Usar
instrumentos
para realizar
medidas
coherentes.

Analizar
datos.

Analizar
gráficos.

Evaluar el
diseño
experimental
(fallas,
mejoras).

Evaluar la
incertidumbre
en los
resultados.

Usar
instrumentos
para realizar
medidas
coherentes.

Analizar
datos.

Analizar
gráficos.

Evaluar el
diseño
experimental
(fallas,
mejoras).

Evaluar la
incertidumbre
en los
resultados.

En esta práctica medimos el período del péndulo utilizando dos instrumentos: un cronómetro manual y un fotosensor.

☐
☐
☐
☐
☐

Los resultados mostraron que ninguno de los valores de aceleración calculados coincidió con el valor teórico ni entre sí.

☐
☐
☐
☐
☐

Sin embargo, el valor obtenido con el fotosensor fue el más preciso, ya que minimiza el error humano presente al usar el cronómetro.

☐☐☐☐☐

Además, verificamos que hasta los 25° de amplitud se cumple la hipótesis de ángulos pequeños: el período se mantuvo prácticamente constante.

☐☐☐☐☐

A partir de ese valor, el período comenzó a variar, evidenciando una desviación respecto al modelo ideal del péndulo simple.

☐☐☐☐☐

Por último, el análisis del histograma mostró un buen ajuste a una curva gaussiana, lo que indica que las mediciones siguen una distribución normal, como era esperable.

☐☐☐☐☐

11. ¿Considera que las herramientas de evaluación que utiliza en el curso están alineadas con los objetivos de aprendizaje propuestos para el curso? *

Es decir, ¿cree que permiten valorar adecuadamente los logros de aprendizaje alcanzados por los estudiantes al finalizar el curso?

Tu respuesta

12. ¿Considera necesario incorporar otras herramientas de evaluación para el curso de Física Experimental I? *

En caso afirmativo, indique cuáles considera relevantes. Puede seleccionar más de una opción.

- ☐ Intervenciones orales para el control de lectura
- ☐ Observación del trabajo grupal durante la práctica, utilizando una rúbrica
- ☐ Uso de una lista de cotejo centrada en el uso de instrumentos, registro de datos y trabajo colaborativo.
- ☐ Evaluación entre pares o coevaluación (de informes o presentaciones orales)
- ☐ Pruebas orales que incluyan la toma de medidas como parte de una práctica
- ☐ Diario de laboratorio o bitácora individual: registro personal de observaciones, decisiones y dificultades durante la práctica.
- ☐ Portafolio de prácticas: donde el/la estudiante documenta su proceso, reflexiona sobre errores y avances, y presenta evidencia de aprendizaje a lo largo del curso.
- ☐ No considero necesario incorporar nuevas herramientas
- ☐ Otro: _____

13. ¿Por qué considera necesaria (o no) la incorporación de nuevas herramientas de evaluación en el curso?

Explique brevemente su respuesta anterior.

Tu respuesta

14. ¿Ha observado recientemente un incremento en el uso de herramientas de inteligencia artificial (IA) por parte de los estudiantes para redactar los informes de laboratorio? *

Seleccione una opción:

- ☐ No logro detectar el uso del a IA
- ☐ Si, en algunas ocasiones
- ☐ Si, frecuentemente.

Si desea **compartir alguna observación, sugerencia o comentario** que considere relevante y que no haya sido contemplado en las preguntas anteriores, puede hacerlo en el siguiente espacio.

Tu respuesta

¡Muchas gracias por su tiempo y colaboración!

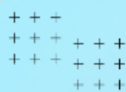
Ante cualquier duda o comentario, puede comunicarse al correo: lorena.pagella@pedeciba.edu.uy

Enviar

Borrar formulario

Apéndice I

Encuesta electrónica realizada a los estudiantes.



¿Qué habilidades científicas se evidencian en los informes de laboratorio de Física Experimental?



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY



Habilidades Científicas en Física Experimental

Estimado/a estudiante:

Esta encuesta forma parte de la tesis de *Maestría en Física de la Universidad de la República* de la Prof. Lorena Pagella, bajo la tutoría de la Dra. Cecilia Stari (Instituto de Física) y la Dra. Aiala Rosá (Instituto de Computación) de la Facultad de Ingeniería. El propósito del estudio es indagar sobre las habilidades científicas que se logran desarrollar en los cursos de "Física Experimental 1".

Su participación es voluntaria, anónima, confidencial y muy valiosa para seguir profundizando esta investigación y no le tomará más de 8 minutos. La información recabada a partir de la encuesta será utilizada únicamente con objetivos académicos.

Si está de acuerdo con participar de la investigación marque la opción "Leí y estoy de acuerdo" y podrá continuar con la encuesta.

Agradezco su colaboración y tiempo.

Por cualquier consulta adicional quedo a las órdenes y puede comunicarse conmigo a través del correo electrónico: lorena.pagella@pedeciba.edu.uy

Profa. Lorena Pagella

* Indica que la pregunta es obligatoria

¿Ha leído la información proporcionada y está de acuerdo en participar en la encuesta en las condiciones indicadas? *

☐ Leí y estoy de acuerdo.

Siguiente

Borrar formulario

Inicio de la encuesta

1. Indique si ha cursado "Física Experimental 1" en alguno de los siguientes años :

*

☐ 2023

☐ 2024

☐ 2025

☐ Otro:

2. Seleccione cinco (5) de las siguientes habilidades principales que, a su juicio, se buscan desarrollar durante el curso de Física Experimental I.

*

- ☐ Seleccionar el marco conceptual para la situación física que se está modelando, vincular la teoría y el experimento comprendiendo las suposiciones, limitaciones y simplificaciones propias del modelo.
- ☐ Montar o construir un dispositivo experimental para realizar mediciones.
- ☐ Utilizar correctamente instrumentos de medición y herramientas de recolección de datos para realizar mediciones coherentes.
- ☐ Cuantificar la incertidumbre y expresarla correctamente.
- ☐ Analizar datos, lo que implica el análisis estadístico básico y el uso de diversas herramientas informáticas para el análisis
- ☐ Representar datos en forma gráfica para analizarlos y extraer información.
- ☐ Evaluar el diseño experimental, identificar posibles problemas y proponer mejoras
- ☐ Evaluar cómo la incertidumbre influye en la validez de los resultados obtenidos identificando las fuentes de error
- ☐ Redactar informes utilizando lenguaje técnico adecuado y específico de la Física, expresar las medidas con cifras significativas e incertidumbre, elaborar diagramas de los dispositivos utilizados y presentar datos en tablas y gráficos de manera clara.
- ☐ Evaluar y reflexionar sobre su propio trabajo, criticar sus propias presentaciones, entregas o informes, tanto por la calidad de los argumentos como por el estilo.

3. ¿En qué medida cree que las siguientes habilidades se consideran en la evaluación del curso?

★

Nada

Poco

Algo

Bastante

Mucho

Redactar
informes con
formato y
lenguaje
científico
adecuados.

☐☐☐☐☐

Evaluar el efecto
de incertidumbre
en las medidas.

☐☐☐☐☐

Usar
correctamente
instrumentos
para obtener
medidas
coherentes.

☐☐☐☐☐

Analizar datos
usando
herramientas
informáticas

☐☐☐☐☐

4. ¿Qué sugerencias o cambios propondría en la evaluación del curso para que contemple mejor estas habilidades u otras que considere importantes?

Tu respuesta

5. ¿Ha utilizado herramientas de inteligencia artificial (IA) para mejorar los informes de laboratorio? *

	1	2	3	4	5	
Nada	☐	☐	☐	☐	☐	Mucho

6. Si ha usado IA, explique brevemente con qué fines la ha utilizando en los cursos de física experimental.

Tu respuesta

¡Muchas gracias por su tiempo y colaboración!

Ante cualquier duda o comentario, puede comunicarse al correo: lorena.pagella@pedeciba.edu.uy

[Atrás](#)

[Enviar](#)

[Borrar formulario](#)

ANEXOS

Anexo 1

Categorías de la Taxonomía de Bloom

Tabla 1.1: Descripción de las categorías de la TB. Elaboración propia. Fuente: Bloom (1990).

Categoría	Definición
Conocimiento	Incluye las situaciones en las que el estudiante debe recordar ideas, materiales, fenómenos. El aprendizaje consiste en que almacene y logre recordar cierta información. El conocimiento está asociado a recordar datos específicos, terminología, fechas, hechos, personas, lugares, fuentes de información. Refiere también al conocimiento de modos de organizar, presentar ideas en cierto campo con determinadas características, conocer procesos, procedimientos, metodologías, clasificaciones, categorías, teorías y estructuras.
Comprensión	Es entender lo que se transmite y hacer uso, de alguna manera, del material o las ideas proporcionadas. Un estudiante demuestra comprensión cuando puede traducir, es decir, logra expresar en otro lenguaje o forma de comunicación lo que ha recibido. También comprende cuando interpreta, reordenando ideas y creando una nueva configuración en su mente. Además, un estudiante puede comprender por extrapolación, lo que implica realizar estimaciones o predicciones basadas en tendencias o condiciones.
Aplicación	Se entiende por aplicar algo cuando esta acción exige la condición previa de la comprensión del método, teoría, principio o la abstracción que sea necesaria.
Análisis	Implica el fraccionamiento del material en sus partes constitutivas, determinar las relaciones entre las partes y comprender cómo están organizadas.
Síntesis	Es la reunión de los elementos y las partes para formar un todo. Los elementos y las partes deben ser combinados para constituir un esquema o estructura que antes no estaba presente con claridad. Implica la combinación de experiencias previas con materiales nuevos. Esta categoría brinda al estudiante la oportunidad de demostrar su capacidad productiva. El estudiante debe manejar elementos provenientes de diversas fuentes y reunirlos en una estructura o producto que sea más que la suma de materiales o partes.
Evaluación	Es la formación de juicios sobre el valor de ideas, obras, soluciones, métodos, materiales. Implica el uso de criterios y pautas para realizar la valoración. Los juicios pueden ser cuantitativos o cualitativos.

Anexo 2

Guía general para la elaboración de informes



Física Experimental 1

Guía general para la elaboración de informes



En este documento se sugieren pautas básicas para la elaboración de informes de laboratorio. El objetivo principal de un informe de laboratorio es describir la experiencia realizada y las conclusiones que se obtuvieron, eventualmente, qué enseñanzas quedan de su realización. Un informe debe ser redactado de forma tal que un lector no necesariamente informado pueda formarse un panorama general de los elementos que se desea transmitir. A continuación se detallan las distintas partes que debería tener un informe y que objetivo persigue cada una de ellas, destacando que pueden realizarse adaptaciones para cada informe en particular.

1. Objetivos

Esta sección del informe debe explicar claramente cual es la motivación del experimento realizado, el objetivo del mismo y la discusión involucrada que se quiere fomentar.

2. Fundamento teórico

El fundamento teórico debe ser breve y escrito con sus propias palabras. El mismo debe incluir solamente los conceptos teóricos más relevantes relacionados con la experiencia y con el análisis de datos que se pretende realizar. Es importante que se mencionen las hipótesis que se realizan en el modelo utilizado. Deben definirse el sistema físico de interés y todas las variables involucradas. Las ecuaciones que se coloquen deben estar numeradas y citadas en el texto cuando sea necesario.

Cuando los conceptos teóricos y deducciones sean obtenidos a partir de la lectura de material bibliográfico (libros, revistas, páginas web, etc.) deben citarse dichas referencias. Al final de este documento se muestra una posible forma de citar bibliografía.

3. Descripción de la experiencia

En esta sección se debe explicar de cómo fue realizado el experimento.

- ¿Qué magnitud se midió?
- ¿Cómo se midió?

- Es fundamental incluir un esquema o diagrama del experimento (imagen ilustrativa) y describir como realizaron las medidas.
- Hay que especificar los instrumentos de medida utilizados y sus características más relevantes incluida la incertidumbre.

Si hubiera ocurrido algún tipo de inconveniente durante la realización del experimento debe hacerse notar siempre y cuando sea de relevancia para el desarrollo del informe.

4. Resultados, análisis y discusión

En esta sección se deben presentar el trabajo realizado a partir de la adquisición de datos. Como todo el documento, debe seguir un hilo conductor definido tomando como base las secciones previas. Puede incluir sub-secciones para ordenar la información.

A modo de guía, es importante:

- Detallar los parámetros que definen el sistema experimental y sus incertidumbres.
- Cuidar el correcto uso de unidades. Toda magnitud que incluya en su informe debe ir acompañada de las unidades que le correspondan, con su incertidumbre asociada.
- Colocar los cálculos de las magnitudes obtenidas como resultado del análisis de datos con sus respectivas incertidumbres referenciando las ecuaciones del fundamento teórico que se están utilizando.
- El valor de la incertidumbre debe estar explicado y justificado.
- Discusiones y conclusiones que se obtienen de los datos, de los valores obtenidos, tablas y gráficas presentadas. Los resultados deben irse analizando a medida que son colocados en el informe.
- Tablas: Siempre con presentación y explicación.
- Gráficas explicativas para presentar los datos experimentales y el resultado de los análisis, ajustes o modelos teóricos. Deben poder leerse claramente, los ejes deben estar expresados con las magnitudes correctas y sus unidades.
- Tanto las tablas como las gráficas y figuras deben estar numeradas y ser citadas en el texto.

5. Conclusiones

En esta sección se deben resumir los resultados obtenidos y compararlos con lo esperado previo a la realización de la experiencia. Si alguna de las conclusiones ya fué comentada en la sección anterior, debe volver a incluirse en esta sección del documento brevemente. Las conclusiones finales incluyen por ejemplo:

- Comparar qué se midió con lo que se esperaba.
- ¿Qué conclusiones puede sacar de los resultados?
- Comparación los resultados obtenidos con referencias bibliográficas.
- ¿Concuerdan los valores dentro del rango de incertidumbre?
- Si no concuerdan, ¿Por qué no concuerdan?
- ¿Son razonables los resultados que obtuvo?
- ¿Qué evaluación hace de la experiencia realizada?
- ¿Se podría mejorar su experimento de alguna manera?

6. Bibliografía

6.1. ¿Cómo referenciar bibliografía?

Existen varias formas aceptadas para citar bibliografía en un texto. Por lo general la bibliografía se coloca al final del documento, donde las mismas aparecen listadas con un identificador. A continuación se muestra un ejemplo extraído de una práctica del curso.

Ejemplo:

Para citar bibliografía a lo largo de la redacción simplemente se incluye el identificador de la referencia ([Bev69] por ejemplo). Existen varias formas de construir el identificador, en este caso se utiliza las primeras tres iniciales del primer autor seguido del año de la publicación [Gil01]. El formato de las referencias también es variable, el formato utilizado aquí es: primero autores, luego título del trabajo en itálica y finalmente identificación de la publicación. Existen también otras formas de escribir las referencias [Res93].

Referencias:

- [Bev69] P. R. Bevington. Data reduction and error analysis for the physical sciences. Primera edición 1969.
- [Gil01] S. Gil, E. Rodríguez. Física re-Creativa, Experimentos de

Física usando nuevas tecnologías. Primera edición 2001. Actividad 50.
<http://www.fisicarecreativa.com>.

[Res93] R. Resnick, D. Halliday, K. S. Krane - Física Vol. 1. Tercera edición 1993, cuarta reimpresión 1996. Capítulo 25.

7. Apéndices

Se pueden colocar en los apéndices aquellos elementos que son necesarios para la elaboración y finalización de la experiencia pero que no representan un elemento sustancial, por ejemplo:

- Cálculos o desarrollos extensos.
- Códigos y scripts utilizados para el procesamiento de datos.
- Tablas de datos extensas. Respecto a este punto se analizará en cada situación si se colocan o no todos los datos obtenidos

8. Tips sobre la redacción general

A continuación se enumeran una serie de recomendaciones sobre la prolijidad general de un documento. Recuerde que a la hora de realizar un informe no sólo es importante el contenido de lo que se informa, sino también la forma en que se hace.

- Cuide rigurosamente las faltas de ortografía. Existen algunas palabras que los correctores ortográficos no son capaces de detectar.
- Cuide el correcto uso de unidades. Cualquier magnitud que incluya en su informe debe ir acompañada de las unidades que le corresponda.
- Todas las imágenes y tablas deben estar correctamente referenciadas, y debe estar redactado el por qué de esa imagen y que se pretende que el lector observe en ella.
- No incluya gran cantidad de elementos no relevantes en su redacción. Tampoco se sienta limitado a contar las experiencias realizadas, ponga el centro en lo que desea informar, y redacte todo aquello que crea que aporte a la discusión de resultados y explicación de la experiencia.
- Incluya numeración de página.

Anexo 3

Guías de Laboratorio

- 3.1. Guía de trabajo de la Práctica 1 - Determinación de la aceleración de la gravedad local.



Física Experimental 1



Práctica 1 - Determinación de la aceleración de la gravedad local

1. Objetivos

Con el fin de familiarizarnos con los procedimientos estadísticos para el tratamiento de datos expuestos en el repartido Medidas y tratamiento estadístico de datos, se propone determinar experimentalmente la aceleración local de la gravedad estudiando un péndulo simple.

2. Fundamento teórico

2.1. El péndulo simple

El péndulo simple es la idealización de un sistema físico real. Dicho sistema real es una masa m suspendida de un soporte mediante un hilo de masa muy pequeña comparada con m . La idealización es una masa puntual, suspendida de un soporte fijo, mediante un hilo flexible, inextensible y sin masa, que se mueve confinada a un plano, como se muestra en la Figura 1. Se desprecian las fuerzas de rozamiento entre el aire y la masa, y entre el hilo y el punto de soporte.

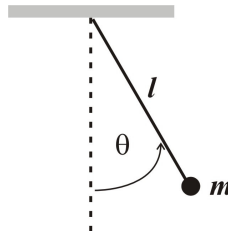


Figura 1: Esquema de un péndulo simple.

La ecuación de movimiento del péndulo simple es:

$$\ddot{\theta} + \frac{g}{l} \sin \theta = 0 \quad (1)$$

donde g es la aceleración gravitatoria, l la longitud del péndulo y θ el ángulo respecto a la vertical. La ecuación 1 es válida para todo valor de θ , pudiéndose linealizar en torno a la posición de equilibrio estable ($\theta = 0$), si θ es pequeño:

$$\ddot{\theta} + \frac{g}{l} \theta = 0 \quad (2)$$

De aquí en más nos referiremos a la ecuación 2 como la ecuación de pequeñas oscilaciones. Como observación, cuando θ deja de ser pequeño (mayor a 10°) sucede que $\sin(\theta) < \theta$, por lo que la ecuación 2 estaría sobrestimando la fuerza de restauración del péndulo. De resolver la ecuación 2 en forma analítica, se puede ver que si el movimiento se restringe al rango en que el ángulo formado por el hilo y la vertical es pequeño, el período se vuelve independiente de la amplitud de la oscilación, esto es, que para cualquier amplitud dentro de ese rango, es aproximadamente constante.

2.2. Ejercicio 1

- Deducir la ecuación 1 a partir del modelo físico correspondiente al péndulo simple.
- Demostrar que la expresión del período, considerando la ecuación linealizada es:

$$T = 2\pi\sqrt{\frac{l}{g}} \quad (3)$$

Nota : La expresión del período puede ser usada para la determinación del valor de g , siempre que se mide T y l con suficiente precisión.

2.3. La aceleración de la gravedad

La llamada fuerza de la gravedad ($\vec{F}$) es aquella que hace que dos cuerpos masivos sean atraídos uno hacia el otro. La expresión para dicha fuerza fue propuesta por Newton en 1666 y puede ser definida como:

$$\vec{F} = -G \frac{Mm}{R^2} \hat{r} \quad (4)$$

donde R es la distancia entre los objetos de masas m , M y $G = 6,67 \times 10^{-11} \text{ Nm}^2/\text{kg}^2$ es la constante de Newton y $\hat{r}$ es un versor que define la línea que conecta los centros de masa de los dos cuerpos.

Al aplicar dicha fórmula al problema del movimiento de los cuerpos cerca de la superficie de la Tierra, basta con tomar M y R como la masa y radio medio de la Tierra, siendo: $M_{\text{Tierra}} = 5,97 \times 10^{24} \text{ kg}$ y $R_{\text{Tierra}} = 6371 \text{ km}$ para obtener la forma del módulo de $\vec{F}$ en la ecuación:

$$F = mg \quad (5)$$

por lo tanto:

$$g = G \frac{M}{R^2} \cong 9,8163 \text{ m/s}^2 \quad (6)$$

donde m es la masa de un cuerpo cuyo movimiento se quiere describir. La ecuación 6 indica el valor esperado para la aceleración de la gravedad g cerca de la superficie de la Tierra.

Estrictamente hablando, g no es una constante, ya que depende de la distribución de masa y de la distancia al centro de la Tierra. Eso implica que aún imaginando un proceso de medida sin errores y con una muy baja incertidumbre, el valor medido de g debería depender al menos del punto geográfico en el cual ha sido tomado. Además, la presencia cercana de otros cuerpos astronómicos cambiaría el valor de la aceleración experimentada por el cuerpo de masa m . En el caso de la Tierra, las fases de la luna pueden afectar el valor medido de g para una alguna ubicación geográfica. Como veremos, las limitaciones impuestas por nuestros instrumentos de medida nos impedirán de distinguir entre cambios tan chicos, donde no podremos expresar el valor de g con más de 3 cifras significativas.

3. Procedimiento experimental

Se utilizarán materiales disponibles en el laboratorio para armar un péndulo simple (tanza, objetos esféricos, soportes). Para la medición del período de las pequeñas oscilaciones se dispondrá de un fotosensor y un cronómetro.

1. Antes de armar el dispositivo deben discutir en grupos acerca del modelo utilizado y las hipótesis que se realizan en el mismo. Analizar y definir el largo del péndulo que utilizarán para la experiencia.
2. Medir la longitud del péndulo l . Recordar que la masa que estamos usando no es puntual.

3. Discutir en grupos y luego con el docente: ¿Cómo se determinarán el período? ¿Cuál es la incertidumbre al medir el período?, ¿Cómo puedo disminuirla?.
4. Una vez elegida la técnica de medida, determine el período en función del ángulo inicial desde donde se suelta la masa. ¿Cuál es el comportamiento esperado? ¿En qué rango de ángulos iniciales es razonable medir?
5. Mida el período de oscilación con dos instrumentos diferentes (un cronómetro y un fotosensor infrarojo). Recuerde que es importante poner la masa en movimiento con un ángulo inicial que permita considerar pequeñas oscilaciones procurando que el movimiento esté restringido a un plano.
6. Repetir el punto ?? un número suficiente de veces, que permita realizar un análisis estadístico. Registre los valores de los períodos en un archivo de datos.
7. Analizar la incertidumbre en la medida del período al utilizar cada uno de los instrumentos y métodos.
8. Dejar oscilar el péndulo por un tiempo de algunos minutos y medir de forma consecutiva el periodo de oscilación (solo con el fotosensor infrarojo).

4. Tratamiento de datos

1. Graficar con barras de error el período en función del ángulo inicial.
2. Realice un histograma que muestre las frecuencias de ocurrencia de las mediciones del período. Esto debe realizarse para las medidas obtenidas con el fotosensor y con el cronómetro.
3. Calcule el promedio del período y su desviación estándar para los 2 experimentos realizados. Verifique si es necesario descartar medidas en los datos del período (recuerde el criterio de descarte de medidas). En ese caso, se deberá graficar un nuevo histograma, y calcular nuevamente el promedio y desviación estándar para el vector de tiempos resultante.
4. Representar sobre cada histograma la curva de Gauss correspondiente. Tenga precaución con el factor de normalización de la curva gaussiana.
5. Exprese el valor de la aceleración de la gravedad con la incertidumbre correspondiente, obtenida mediante las medidas de cada uno de los instrumentos utilizados. Comentar sobre la exactitud y precisión del valor de g obtenido.
6. Realizar un gráfico de las medidas del periodo realizadas de forma consecutiva (con el fotosensor).

3.2. Guía y pautas de corrección de informe: Péndulo Simple y Estadística.



Física Experimental 1



Guía y pautas de corrección del Informe - Péndulo simple y Estadística

Este material es una guía para la realización del informe. Tiene detallada información específica para esta práctica. Su lectura debe complementarse con la Guía para escribir un informe donde encontrarán información general válida para todos los informes del curso.

Formato general - max 1.5 puntos

Incluye: organización, contenido general del informe, redacción y ortografía. Las graficas deben estar correctamente presentadas, con pie de figura y citadas en el texto.

Objetivos - max 0.5 puntos

Detallen los objetivos de la práctica en un máximo de 5 líneas

Fundamento teórico - max 1 punto

Explique el fundamento teórico de la práctica, sintetizando los conceptos que son importantes para el desarrollo de la misma. Deben explicarse únicamente los conceptos y ecuaciones más importantes que sean necesarias para el desarrollo de la práctica. Mencionar que hipótesis se realizan, rango de validez de las aproximaciones. **Máximo: 1 carilla**

Metodología - max 1.5 puntos

Describa el dispositivo experimental y la técnica de medida utilizada. Coloque un esquema y/o foto que permita visualizar el experimento. Detalle el procedimiento de medida. No se debe colocar un punteo de las actividades realizadas, sino redactar un párrafo con los aspectos importantes del procedimiento, analizando que debe tenerse en cuenta para cumplir las hipótesis realizadas. Por ejemplo, se debe explicar: ¿Por qué se deciden registrar un número grande de períodos en lugar de medir sólo un período?, ¿Por qué deben realizar el experimento en iguales condiciones?, ¿Por qué deben medir tantas veces lo mismo?. Establezcan al menos una diferencia entre medir el tiempo con cronómetro y con sensor.

Análisis de datos - max 4.5 puntos

- Presenten los datos de período en función del ángulo inicial añadiendo barras de error en los puntos experimentales. Discutan si era correcta o no la hipótesis de pequeñas oscilaciones para el ángulo elegido en las primeras medidas.
- Presenten el histograma de ocurrencias para el período medido con cronómetro y con sensor, con curva Gaussiana superpuesta (recuerden tener cuidado con el factor de normalización de la curva gaussiana). ¿Aplicaron criterio de eliminación de datos? Justificar brevemente.
- Comparen cualitativamente los histogramas y sus respectivos ajustes. ¿Esperan que los datos sigan una distribución normal? ¿Lo cumplen?

- Calculen los valores medios y desviaciones en cada caso, expresando el resultado final del período T con su incertidumbre.
- Calculen el resultado final para la aceleración gravitatoria, con su incertidumbre. Discutan si el valor es concordante* con el valor de referencia. Comentar sobre exactitud y precisión de la medida de g .
- Presenten la grafica del período medido con el fotosensor de forma consecutiva ¿Que pueden comentar respecto a la evolución del valor del periodo con el tiempo?

Que una medida sea concordante con otro valor implica que el valor a comparar se encuentra entre el rango de incertidumbre de la medida. Por ejemplo, $9,8 \pm 0,2$ concuerda con cualquier valor entre 9.6 y 10.0.

Recordar:

1. Las **ecuaciones** deben estar **numeradas y citadas** en el texto cuando se haga uso de ellas. Las **gráficas** deben estar numeradas y se debe colocar un **pie de figura** que describa cada una de ellas. Deben contener nombre de la magnitud y unidades en los ejes y leyenda en caso de ser necesario. Los puntos experimentales **NO** deben unirse y las curvas teóricas se deben realizar en trazo continuo y bien definido. Además, deberán acompañar las gráficas de una breve descripción y comentario de la información que extraen de ella. Todas las figuras deben ser citadas en el texto. Los **valores finales** deberán estar correctamente presentados, siguiendo el Criterio de cifras significativas. Se debe colocar la bibliografía utilizada, citándola en el texto cuando se haga referencia a información obtenida de la misma.
2. Utilice los conceptos de precisión, exactitud y confiabilidad entre otros para apoyar la discusión de sus resultados.

Conclusiones - max 1 punto

Redacten un párrafo (máximo 10 líneas) que incluya la discusión de las siguientes preguntas:

- ¿Los valores de aceleración obtenidos experimentalmente mediante el valor de T concuerdan con el valor de referencia? ¿Concuerdan entre sí?
 - ¿Observan alguna diferencia entre los dos instrumentos de medida empleados?
 - ¿El comportamiento estadístico seguido por los datos del período fue el esperado?
 - ¿Cómo depende el período de del ángulo inicial? ¿Los datos obtenidos fueron acordes a lo esperado?
-

3.3. Guía de trabajo de la Práctica 2 - Caracterización de dispositivos de medición de temperatura.



Física Experimental 1



Práctica 2 - Caracterización de dispositivos de medición de temperatura

1. Objetivos

Se busca comprender las principales características de funcionamiento de los dispositivos de medición de temperatura. Se estudiarán las características básicas de algunos termómetros usuales y realizar una calibración de alguno de ellos.

2. Fundamento teórico

Dispositivos de medición de temperatura

En muchos sistemas de física, química e ingeniería en general, la temperatura constituye una variable física importante que necesita ser monitoreada y controlada.

Los dispositivos de medición de temperatura se pueden clasificar (según la propiedad termométrica usada), como de funcionamiento mecánico (termómetro de mercurio, termómetro de alcohol y tira bimetalica) o de funcionamiento eléctrico (detector de temperatura de resistencia, termistor y termopar). En la práctica existen numerosos tipos de sensores de temperatura o termómetros que, según la aplicación específica, pueden ser los más adecuados. En la Tabla 1 se indican algunos tipos de termómetros y sensores de temperatura más comúnmente usados junto a algunas de sus características más notables (Gil y Rodríguez, 2001).

Tipo de termómetro	Rango nominal (°C)	Linealidad	Dif. Características notables
de mercurio	-10 a 300	buena	simple, lento, de lectura manual
termoresistencia (RTD)	-150 a 600	alta	mayor exactitud
termocupla	-150 a 1500	alta	requiere referencia de temperatura
termistor	-15 a 115	no lineal	muy sensible

Tabla 1: Algunos tipos de termómetros y sensores de temperatura usuales junto a algunas de sus características más notables.

En esta práctica, nos concentraremos en los sensores de temperatura de base eléctrica debido a su mayor precisión y facilidad para proporcionar mediciones para el procesamiento de señales. Estos sensores se basan en el principio de que la resistencia eléctrica del material del que están hechos presenta una dependencia con la temperatura.

En este punto es importante mencionar el proceso de calibración de un sensor. Los sensores a utilizar no brindan de manera directa el valor de la temperatura T a la que se encuentran, sino que lo hacen de manera indirecta, mediante el valor de la resistencia eléctrica R , siguiendo determinada relación entre dichas variables. Conocer completamente la relación $R(T)$ es lo que se denomina calibrar el sensor, de manera de que si conocemos R entonces podemos encontrar T . Como instrumento patrón para calibrar utilizaremos un termómetro de alcohol, dispositivo en el que vamos a confiar que la temperatura medida es correcta.

Sensores de temperatura resistivos, termo-resistencia - RTD

Los termómetros de resistencia o termómetros a resistencia son transductores de temperatura (“Resistance Temperature Detector”, RTD) que se basan en la dependencia de la resistencia eléctrica de un material con la temperatura. En ellos se aprovecha el efecto que tiene la temperatura en la conducción de los electrones para que, ante un aumento de temperatura, haya un aumento de la resistencia eléctrica que presentan.

La variación de la resistencia eléctrica R que presenta un sensor RTD con la temperatura T (para la mayoría de los materiales metálicos), puede ser expresada como se muestra en la Ec. (1):

$$R(T) = R_o(1 + \alpha_1(T - T_o) + \alpha_2(T - T_o)^2) \quad (1)$$

En la Ec. (1) T_o es la temperatura de referencia, R_o es la resistencia a la temperatura de referencia, y α_i son coeficientes constantes positivos. El número de términos incluidos en la expresión de $R(T)$ depende del material y de la exactitud. Típicamente, solo se incluye el coeficiente α_1 ya que se obtiene un comportamiento lineal para un gran intervalo de temperaturas. Por ejemplo, el Platino (Pt) presenta un comportamiento lineal de R vs. T para el rango de temperaturas que va desde -75 °C a 50 °C. En esta práctica, para el caso del sensor RTD, vamos a incluir solo el primer coeficiente, por lo que el modelo de resistencia en función de la temperatura que se estudiará será el que muestra la Ec. (2).

$$R(T) = R_o(1 + \alpha_1(T - T_o)) \quad (2)$$

La Figura 1 muestra un esquema de la estructura de un sensor RTD comercial, así como una imagen del mismo. Un material muy utilizado en la fabricación de estos dispositivos es el Pt, pues tiene una gran estabilidad, alto coeficiente de variación de la resistencia con la temperatura, alta reproductividad y puede usarse en un amplio rango de temperaturas (Gil y Rodríguez, 2001). Están protegidos externamente por una cápsula de alúmina. Cuando son de forma cilíndrica (se consiguen de distintos tamaños y formas), tienen un tamaño de unos 10 mm de largo y unos 2 mm de diámetro. Los más comunes se denominan Pt-100, lo que indica que está hecho con alambre de platino y presenta una resistencia de 100 Ω a 0 °C (Gil y Rodríguez, 2001).



Figura 1: Esquema de la estructura de un sensor de temperatura RTD e imagen digital de un dispositivo comercial.

Termistores

El termistor está formado por un material semiconductor y su funcionamiento se basa en la variación de resistividad que presenta el semiconductor con la temperatura. Cuando la resistencia

del semiconductor disminuye con un aumento de la temperatura decimos que tenemos un termistor NTC “Negative Temperature Coefficient”. Por otro lado, si la resistencia aumenta cuando aumenta la resistencia tendremos un termistor PTC “Positive Temperature Coefficient”. Una ecuación que se puede utilizar para vincular la temperatura con la resistencia en los termistores NTC es la Ec. (3):

$$R(T) = R_o \exp \left(B \left(\frac{1}{T} - \frac{1}{T_o} \right) \right) \quad (3)$$

En la Ec. (3) R_o es la resistencia del termistor a una determinada temperatura absoluta de referencia T_o , B es un parámetro del material (expresado en K), y T es la temperatura, también absoluta, a la cual se encuentra el termistor, para la que presentará una resistencia eléctrica R . Generalmente, el parámetro B es posible encontrarlo en las correspondientes hojas de datos que el fabricante del termistor provee al usuario. Un ejemplo de dicha hoja se ve en la Figura 2.

General technical data				
Climatic category	(IEC 60068-1)		55/125/21	
Max. power	(at 25 °C)	P_{25}	450	mW
Resistance tolerance		$\Delta R_0/R_n$	$\pm 5, \pm 10$	%
Rated temperature		T_R	25	°C
Dissipation factor	(in air)	δ_{th}	approx. 7.5	mW/K
Thermal cooling time constant	(in air)	τ_c	approx. 20	s
Heat capacity		C_{th}	approx. 150	mJ/K

Electrical specification and ordering codes			
R_{25} Ω	No. of R/T characteristic	$B_{25/100}$ K	Ordering code
15	1203	2900 ±3%	B57164K0150+000
22	1203	2900 ±3%	B57164K0220+000
33	1203	2900 ±3%	B57164K0330+000
47	1302	3000 ±3%	B57164K0470+000
68	1303	3050 ±3%	B57164K0680+000
100	1305	3200 ±3%	B57164K0101+000
150	1305	3200 ±3%	B57164K0151+000
220	1305	3200 ±3%	B57164K0221+000

Figura 2: Ejemplo de hoja de datos de un termistor donde se ve tabulado el parámetro B.

La Figura 3 muestra diferentes tipos de termistores. Si bien su tamaño es variable, en general los mismos son del orden de algunos milímetros.



Figura 3: Ejemplo de diferentes termistores.

Finalmente, en la figura 4 pueden observarse las curvas típicas de comportamiento de la resistencia con la temperatura para un sensor RTD de Pt (curva negra) y termistores de distinto valor de R_o (R_{25} en la figura). Para el caso del RTD puede verse que el corte en $T = 0$ coincide con el valor de R_o .

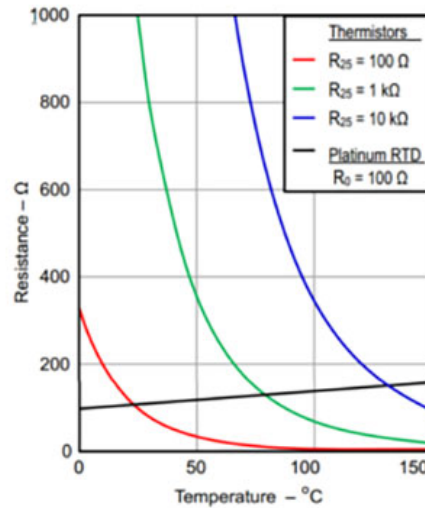


Figura 4: Curvas típicas de resistencia con la temperatura para un RTD de Pt y distintos termistores. Figura extraída de Texas Instruments

2.1. Ejercicio 1

1. De la Ec. (3), realizar un cambio de variables adecuado de manera que quede de la forma $y = Mx + N$.
2. De los coeficientes M y N , despejar los parámetros del termistor R_o y B , así como sus correspondientes incertidumbres.
3. Repetir los puntos 1 y 2, para el caso del RTD. Recordar que en esta práctica consideraremos para el RTD la siguiente relación: $R(T) = R_o(1 + \alpha_1(T - T_o))$

2.2. Ejercicio 2

1. De la Ecs. (2) y (3), despejar explícitamente la temperatura en función de la resistencia, y los correspondientes parámetros, para ambos dispositivos ($T = f(R)$).

3. Procedimiento experimental

Calibración del RTD y/o termistor

Para obtener los datos para la calibración del RTD y/o el termistor, se utilizará un baño de agua cuya temperatura podrá ser modificada. A su vez se contará con un termómetro de alcohol y un multímetro digital (elemento para medir la resistencia eléctrica).

1. Discutir la preparación del dispositivo experimental de manera que se pueda medir en paralelo la resistencia del RTD y/o termistor junto a la temperatura del baño. Analizar qué cuidados se deben tener. Considerar que para que la calibración sea razonablemente buena, se deben medir al menos 10 puntos en un intervalo de temperaturas que va desde la temperatura ambiente a 70 °C (aproximadamente).

2. Armar el dispositivo experimental sumergiendo los elementos y sin encender el baño térmico. Tomar la primer medición con la temperatura ambiente.
3. Variar la temperatura del agua y esperar que el sistema llegue a un nuevo equilibrio térmico. En dicho momento, registrar nuevamente la temperatura del baño y la resistencia del RTD/termistor.
4. Repetir el paso 3 de manera de tener al menos 10 puntos para la calibración.

4. Tratamiento de datos

Calibración del RTD y termistor

1. Graficar los datos experimentales obtenidos. ¿Siguen el comportamiento esperado? Si hay diferencias, explicar a qué pueden deberse.
2. Realizar los cambios de variable correspondiente de manera de linealizar los datos obtenidos.
3. Para el termistor, mediante el método de mínimos cuadrados, obtener los valores (junto a su incertidumbre) de los parámetros B y R_o . Verificar si existen valores de referencia para dichos parámetros y comparar en caso positivo.
4. En el caso de RTD, obtener los parámetros (junto a su incertidumbre) del modelo α_1 y R_o . Verificar si existen valores de referencia para dichos parámetros y comparar en caso positivo.
5. Graficar los datos linealizados y el ajuste realizado.
6. Graficar los datos experimentales junto al modelo original según sea el caso (termistor o RTD).

Referencias

- Gil, S. y Rodríguez, E., “Física Re-creativa”, Prentice Hall, 2001.

3.4. Pautas de corrección de informe: Caracterización de dispositivos de medición de temperatura.



Física Experimental 1



Guía Informe - Calibración de dispositivos de medición de temperatura

Este material es una guía para la realización del informe. Tiene información específica detallada para esta práctica. Su lectura debe complementarse con la Guía para escribir un informe donde encontrarán información general válida para todos los informes del curso.

Formato general - max 1.5 puntos

Incluye: organización, contenido general del informe, redacción y ortografía. Las graficas deben estar correctamente presentadas, con pie de figura y citadas en el texto.

Objetivos - max 0.5 puntos

Detallen los objetivos de la práctica en un máximo de 5 líneas

Fundamento teórico - max 1 punto

Explique el fundamento teórico de la práctica, sintetizando los conceptos que son importantes para el desarrollo de la misma. Deben explicarse únicamente los conceptos y ecuaciones más importantes que sean necesarias para el desarrollo de la práctica. Mencionar que hipótesis se realizan, rango de validez de las aproximaciones. **Máximo: 1 carilla**

Metodología - max 1.5 puntos

Describa el dispositivo experimental y la técnica de medida utilizada. Coloque un esquema y/o foto que permita visualizar la manera de calibrar el sensor de temperatura, junto a los demás materiales utilizados. Detalle el procedimiento de medida. No se debe colocar un punteo de las actividades realizadas, sino redactar un párrafo con los aspectos importantes del procedimiento, analizando que debe tenerse en cuenta para cumplir las hipótesis realizadas. Por ejemplo, se debe explicar: ¿Por qué se deciden registrar el número de mediciones realizadas?, ¿Cómo se llegó al mismo?, ¿Por qué deben realizar el experimento en iguales condiciones? Establezcan al menos una diferencia entre medir la temperatura con el termómetro de alcohol o si se hubiese utilizado un termómetro más preciso. ¿Hubiese sido necesario?

Análisis de datos - Caracterización del RTD/termistor - max 4.5 puntos

- Presenten de manera gráfica los datos medidos de resistencia eléctrica en función de la temperatura, añadiendo las barras de incertidumbre en los puntos experimentales. Discutir si los puntos son representativos de la temperatura de toda la masa de agua, si la misma logró ser uniforme al momento de la medida, etc.
- Realizar los cambios de variables correspondientes para lograr la linealización de las variables y poder aplicar el método de mínimos cuadrados, calculando los diferentes parámetros de interés. Reportar el valor del coeficiente de correlación lineal y discutir qué tan buenos fueron los ajustes y si fuese posible realizar algún cambio para mejorar los datos obtenidos.
- Obtener los parámetros de los modelos (B , R_0 , α_1) junto con sus incertidumbres y discutir sobre los mismos. Graficar los datos experimentales medidos junto al modelo obtenido por mínimos cuadrados.

- Buscar valores de referencia típicos de B , R_o , α_1 . Discutan si el valor es concordante* con el valor de referencia.

*Que una medida sea concordante con otro valor implica que el valor a comparar se encuentra entre el rango de incertidumbre de la medida. Por ejemplo, $9,8 \pm 0,2$ concuerda con cualquier valor entre 9.6 y 10.0.

Recordar:

1. Las **ecuaciones** deben estar **numeradas y citadas** en el texto cuando se haga uso de ellas. Las **gráficas** deben estar numeradas y se debe colocar un **pie de figura** que describa cada una de ellas. Deben contener nombre de la magnitud y unidades en los ejes y leyenda en caso de ser necesario. Los puntos experimentales **NO** deben unirse y las curvas teóricas se deben realizar en trazo continuo y bien definido. Además, deberán acompañar las gráficas de una breve descripción y comentario de la información que extraen de ella. Todas las figuras deben ser citadas en el texto. Los **valores finales** deberán estar correctamente presentados, siguiendo el Criterio de cifras significativas. Se debe colocar la bibliografía utilizada, citándola en el texto cuando se haga referencia a información obtenida de la misma.
2. Utilice los conceptos de precisión, exactitud y confiabilidad entre otros para apoyar la discusión de sus resultados.

Conclusiones - max 1 punto

Redacten un párrafo que incluya la discusión de las siguientes (entre otras que pueden responder) preguntas:

- ¿Los valores de B , R_o y α_1 obtenidos experimentalmente mediante mínimos cuadrados concuerdan con el valor de referencia?
 - En ambas partes de la práctica ¿obtienen el comportamiento esperado para el fenómeno bajo estudio? Justificar.
-

3.5. Guía de trabajo de la Práctica 3 - Oscilaciones amortiguadas.



Física Experimental 1



Práctica 3 - Oscilaciones Amortiguadas

Objetivos

En esta práctica, abordaremos el estudio del movimiento de un sistema masa-resorte con amortiguamiento viscoso, con los siguientes objetivos:

- Diseñar una experiencia que permita determinar la constante elástica del resorte utilizado en el sistema.
 - Evaluar el coeficiente de amortiguamiento viscoso en diferentes medios y condiciones.
 - Aplicar y reforzar los conceptos y herramientas de análisis de datos adquiridos en clases anteriores.
-

1. Fundamento Teórico

La elasticidad es la propiedad de un material de recuperar su tamaño y forma original después de ser comprimido o estirado por una fuerza externa. En muchos materiales (metales y minerales) la deformación es directamente proporcional a la fuerza. Esto se conoce como la ley de Hooke, publicada en 1678 por Robert Hooke, uno de los científicos experimentales más importantes de la historia. Sin embargo, si la fuerza supera determinado valor, el material queda deformado permanentemente y la ley de Hooke ya no es válida. La máxima fuerza que un material puede resistir en una sección determinada sin sufrir deformaciones permanentes se denomina límite de elasticidad.

Ahora supongamos un sistema masa resorte con amortiguamiento como el que se observa en la figura 1, donde sólo trabajaremos con la proyección según el versor $\hat{j}$ de la segunda ley de Newton.

En la dirección $\hat{j}$, la fuerza ejercida por el resorte (F_e) debido a su estiramiento o compresión es:

$$F_e = -k(y - l_0) \quad (1)$$

donde k es la constante elástica del resorte y l_0 su longitud natural, de modo que $(y - l_0)$ representa la elongación respecto a la longitud natural.

Por otro lado, la fuerza viscosa F_v ejercida por el medio (aire, agua, etc.) es mas compleja de modelar y depende de varios factores como la viscosidad del fluido, el tamaño del objeto sumergido y la velocidad relativa entre ellos. Para velocidades bajas y alta viscosidad, la fuerza se puede modelar como:

$$F_v = -b\dot{y} \quad (2)$$

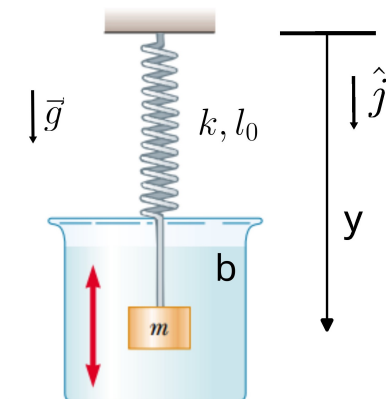


Figura 1: Sistema masa resorte con amortiguamiento

donde b es el coeficiente de amortiguamiento viscoso. A mayores velocidades el modelo se vuelve más complejo, existiendo una zona en donde se puede aproximar como dependiente de una constante por la velocidad al cuadrado. Para el problema acá planteado, esto último efecto puede ser una mejor aproximación, pero al ser no lineal, ya no se puede resolver las ecuaciones analíticamente. Por tal motivo, se trabajará con el modelo planteado en la ecuación 2.

Teniendo en cuenta las fuerzas que actúan sobre la masa descritas por las ecuaciones 1 y 2 además del peso, se llega a la siguiente ecuación de movimiento,

$$-k(y - l_0) - b\dot{y} + mg = m\ddot{y} \quad (3)$$

donde g es la aceleración gravitatoria terrestre y m la masa que oscila unida al resorte.

Si hallamos la posición de equilibrio:

$$\begin{cases} \dot{y} = 0 \\ \ddot{y} = 0 \end{cases} \rightarrow y_0 = l_0 + mg/k \quad (4)$$

Realizando el cambio de variable: $x = y - y_0$, se resuelve la ecuación 3 para obtener la ley $x(t)$ (ver ejercicio 1), que representa la posición de la masa respecto de la posición de equilibrio (elongación) en función del tiempo, involucrando los parámetros propios del sistema (masa, constante del resorte, constante de amortiguamiento).

Ejercicio 1:

Realizar el cambio de variable planteado en la parte anterior y obtener la ley horaria del movimiento.

Según la relación que vincule los parámetros del sistema existen tres tipos de soluciones para la ec.3:

- Si $b^2 > 4mk$ el movimiento de la masa es **sobreamortiguado** y el gráfico de la elongación en función del tiempo será como se muestra en la Figura 2.

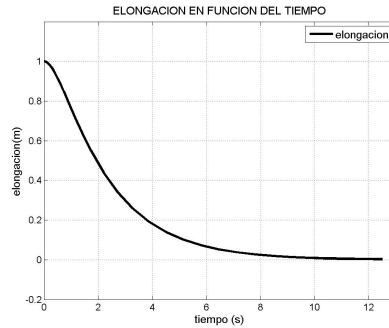


Figura 2: Gráfica de la elongación en función del tiempo para el caso sobreamortiguado.

- b) Si $b^2 < 4mk$, se dice que el movimiento es **subamortiguado** y el gráfico de la elongación en función del tiempo será como se muestra en la Figura 3. Este caso se va a trabajar en más detalle y la ley horaria $x(t)$ se analiza más adelante.

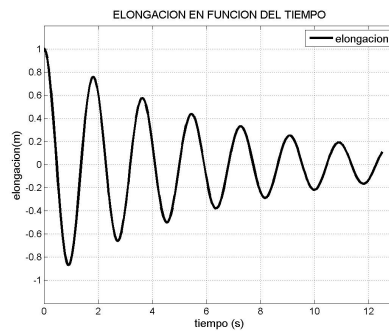


Figura 3: Gráfica de la elongación en función del tiempo para el caso subamortiguado.

- c) Si $b^2 = 4mk$ se dice que existe **amortiguamiento crítico** y el gráfico de la elongación en función del tiempo será como se muestra en la Figura 4.

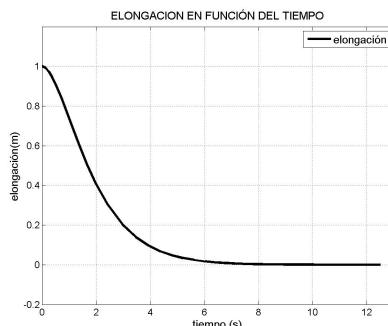


Figura 4: Gráfica de la elongación en función del tiempo para el caso de amortiguamiento crítico.

Ejercicio 2:

Para el caso subamortiguado:

- a) Verificar que la expresión $x(t)$ que se presenta a continuación es solución de la ecuación de movimiento obtenida en el Ejercicio 1, para el caso subamortiguado:

$$\begin{aligned}
 x(t) &= A(t) \sin(\omega t - \phi) \quad A(t) = A_0 e^{-t/\tau} \\
 \omega^2 &= \omega_0^2 - \left(\frac{1}{\tau}\right)^2 \\
 \tau &= \frac{2m}{b} \qquad \omega_0^2 = \frac{k}{m}
 \end{aligned} \tag{5}$$

Donde A_0 es una amplitud para la oscilación y la curva $A(t)$ es la envolvente de $x(t)$.

- b) Pensar como aplicaría el método de Mínimos Cuadrados para obtener el valor de b y de A_0 y sus respectivas incertidumbres. ¿Qué significado físico tiene el valor obtenido de A_0 ?
- c) Demostrar que si despreciamos el coeficiente de amortiguamiento viscoso (b) la constante del resorte esta dada por la expresión:

$$k = \frac{4\pi^2 m}{T^2} \tag{6}$$

2. Dispositivo experimental

Para realizar esta experiencia se cuenta con un sistema masa - resorte, que puede oscilar tanto en aire como en agua. La posición de la masa será determinada utilizando un sensor de movimiento que emite ondas de ultrasonido. También contarán con un dinamómetro, un instrumento capaz de medir fuerzas de empuje y tracción. El funcionamiento de ambos sensores se describe en el apéndice.

1. Recordando que la fuerza elástica se determina según la ecuación 1, discutan en grupos que metodología pueden utilizar para determinar el valor de la constante elástica k del resorte. Teniendo en cuenta que disponen del sensor de movimiento, un sensor de fuerzas y de varias masas diferentes. Consideren particularmente los siguientes casos:
 - 1.1. Medida dinámica: Si se mide al mismo tiempo con el sensor de movimiento y el sensor de fuerzas para determinar simultáneamente $y - l_o$ y $|\vec{F}_e|$ respectivamente.
 - 1.2. Medida estática: Determinar k usando diferentes masas y midiendo el estiramiento y la fuerza cuando las masas están en equilibrio. Comparar los dos valores de k obtenidos con los dos diferentes métodos.

En ambos casos piensen como se puede usar el método de mínimos cuadrados para determinar k con su incertidumbre.

2. Armar un sistema experimental que permita detectar las oscilaciones del sistema masa-resorte.
3. Antes de comenzar a medir se debe analizar qué cuidados experimentales deben tomarse para cumplir las hipótesis del modelo teórico que se quiere aplicar para el análisis.
4. Los datos obtenidos deberán ser exportados desde LoggerPro a un archivo de texto para poder analizarlos utilizando el programa SciDavis o similar. Visualizando la curva obtenida de las oscilaciones del sistema, infiera a cuál de los movimientos oscilatorios que estudiamos corresponde.
5. Encontrar b y su incertidumbre. Para obtener el valor de b , se deberá ajustar los valores de amplitudes máximas por una función adecuada, utilizando el método de mínimos cuadrados. Las medidas y el análisis de b debe ser realizado para un medio de amortiguamiento a elección: Aire o Agua.
6. Con los valores de b y de k obtenidos graficar la función elongación en función del tiempo que se desprende del modelo teórico (Ec. 5) y compararla con la gráfica obtenida experimentalmente.

A. APÉNDICE: Características de los sensores.

A.1. Detector de Movimiento por ultrasonido

El detector de movimiento emite pequeños paquetes de ondas de ultrasonido desde un transductor ubicado en su interior. El detector entonces escucha el eco de las ondas que fueron reflejadas por cierto objeto que se interpuso en su camino. El instrumento mide el tiempo que tardan las ondas en llegar al objeto y volver al detector, usando este tiempo y la velocidad del sonido en el aire, determina la distancia al objeto. La sensibilidad del circuito detector del eco se incrementa de a pasos por cada milisegundo que transcurre, con el fin de tener en cuenta la atenuación de las ondas si el objeto se encuentra lejos del detector.

Las ondas son emitidas en un cono con un ángulo de entre 15 y 20 grados, medidos desde el centro del eje de la fuente, como se muestra en la figura 5. De esta forma, el sensor sólo detecta objetos que se encuentran dentro del cono de emisión. Por otro lado, el sensor no es capaz de identificar la fuente del eco (sólo su distancia), por lo tanto, en caso de tener más de un objeto dentro del cono de emisión el sensor va a dar resultados erróneos. Las especificaciones del detector de movimiento proporcionadas por el fabricante se encuentran en la tabla 1.

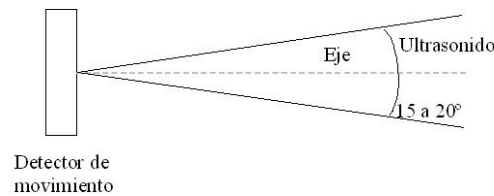


Figura 5: Esquema del detector de movimiento con el cono de ondas de ultrasonido

Tabla 1: Especificaciones del detector de movimiento

Frecuencia del ultrasonido	40KHz
Resolución	2mm
Exactitud típica	2mm
Rango mínimo	$0,4\text{m}$
Rango máximo	6m
Velocidad del ultrasonido	343m/s
Alimentación	$51\text{mA}@5\text{VDC}$

A.2. Sensor de Fuerzas (dinamómetro)

El sensor de fuerza de doble rango que se muestra en la figura 6 es un dispositivo que permite medir fuerzas de empuje y tracción. En este caso, el sensor mide la fuerza que se ejerce sobre su asa (ya sea de empuje o tracción). Se puede usar como reemplazo de una balanza de resorte manual, montarse en un soporte para medir fuerzas verticales o en un carro dinámico para estudiar colisiones. Puede entonces ser utilizado en una amplia variedad de experimentos, que incluyen, por ejemplo: colisiones, movimientos armónicos, ley de Hooke, etc.

En el caso del experimento de ley de Hooke con el sensor vertical, basta con colocar un extremo del resorte en el asa del sensor y el otro extremo en la masa. Se deja que la masa estire al resorte hasta que este alcance la longitud de equilibrio, de esta forma el sensor mide la fuerza ejercida por el resorte que es necesaria para que la masa esté en equilibrio.

Entre otras características el sensor permite medir fuerzas desde 0.01 Newton hasta 50 N. Tiene dos escalas diferentes (rangos): una de 10 N (con resolución de 0.01 N) y otra de 50 N (con resolución de 0.05 N).



Figura 6: Esquema del sensor de fuerza

3.6. Guías de trabajo de la Práctica 4 - Determinación de la velocidad del sonido, Determinación del calor específico de un metal y Movimiento en un plano inclinado.



Física Experimental 1



Práctica 4 - Determinación de la velocidad del sonido

1. Objetivos

Calcular la velocidad del sonido mediante dos métodos diferentes. Reforzar los conocimientos sobre tratamiento de datos adquiridos a lo largo del curso.

2. Fundamento teórico

2.1. Velocidad del sonido en el aire

Cuando escuchamos un sonido, lo que en realidad hacemos es detectar las variaciones de presión que se transportan en un medio (por ejemplo el aire). El sonido se puede representar como una onda mecánica tridimensional y longitudinal (esto último implica que la presión oscila en la misma dirección que el sentido de propagación de la onda). La ecuación 1 representa a la velocidad del sonido en el caso particular del aire comportándose como un gas ideal.

$$\|\vec{v}\| = \sqrt{\frac{\gamma RT}{M}} \quad (1)$$

Donde

- v representa la velocidad del sonido
- γ es la relación de calores específicos del gas a presión y a volumen constante, si nos mantenemos en la hipótesis de que el gas se comporta como ideal y adiabático, entonces $\gamma = 1,40$
- $R = 8,314 \text{ J/molK}$ es la constante universal de los gases ideales
- T es la temperatura absoluta
- M es la masa molar del aire que se puede tomar como $28,8 \times 10^{-3} \text{ kg/mol}$

Por lo tanto la velocidad del sonido termina prácticamente siendo función de la temperatura absoluta, por ejemplo si $T = 20^\circ\text{C}$

$$\|\vec{v}\| = \sqrt{\frac{\gamma RT}{M}} = \sqrt{\frac{(1,40)(8,314 \text{ J/mol} \cdot \text{K})(293 \text{ K})}{28,8 \times 10^{-3} \text{ kg/mol}}} = 344 \text{ m/s} \quad (2)$$

2.2. Ondas estacionarias en un tubo

Las ondas estacionarias resultan de la interferencia de dos ondas que se propagan en direcciones opuestas dentro del tubo, creando regiones de mínima y máxima amplitud conocidas como nodos y antinodos.

Las ondas viajeras siguen la ecuación $y_1(x, t) = A \sin(kx - \omega t)$, donde A es la amplitud, k es el número de onda $\left(k = \frac{2\pi}{\lambda}\right)$, con λ siendo la longitud de onda, y ω es la frecuencia angular ($\omega = 2\pi f$), con f siendo la frecuencia de oscilación de la presión, ver figura 1.

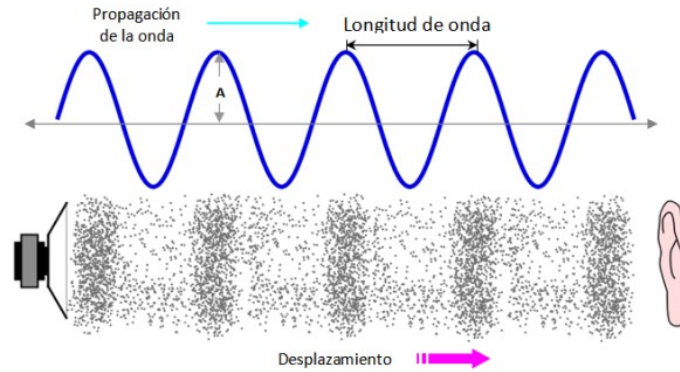


Figura 1: Propagación de una onda de sonido (abajo) y la representación de la oscilación la presión en cada punto (arriba). Fuente: Wikimedia Commons

2.2.1. Tubo abierto

Como se mencionó anteriormente, las ondas estacionarias se pueden tomar como la superposición de dos ondas, de las cuales la segunda es el resultado del reflejo de la primera en el borde de un tubo. La onda incidente que viaja hacia la derecha es:

$$y_1(x, t) = A \sin(kx - \omega t)$$

Si la relación entre el largo del tubo y la frecuencia es la adecuada, entonces la onda reflejada que se propaga hacia la izquierda es:

$$y_2(x, t) = A \sin(kx + \omega t)$$

Sumando ambas entonces tenemos:

$$y_e = A \sin(kx - \omega t) + A \sin(kx + \omega t) \quad (3)$$

Descomponiendo $\alpha = kx$ y $\beta = \omega t$, y utilizando la relación trigonométrica:

$$\sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \cos \alpha \sin \beta \quad (4)$$

Se obtiene que la ecuación de la onda estacionaria

$$y_e(x, t) = 2A \sin(kx) \cos(\omega t) \quad (5)$$

Ejercicio 1: Utilizar la relación en la ecuación (4) para deducir la ecuación (5).

El primer elemento que está multiplicando en la ecuación (5) ($\sin(kx)$) depende únicamente de la posición y generará en los casos extremos puntos de oscilación nula (nodos) y de oscilación máxima (antinodos), esto se puede ver en la figura 2, y funcionará como una cota superior para el segundo elemento ($\cos(\omega t)$) que depende del tiempo.

Si se continúa con el razonamiento anterior, una vez que se forman ondas estacionarias la longitud de onda correspondiente para que aparezcan nodos en los extremos del tubo sigue la relación

$$\lambda = \frac{2L}{n} \quad (6)$$

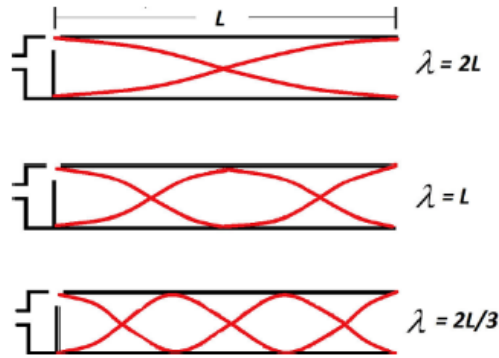


Figura 2: Representación de las ondas estacionarias en un tubo abierto por ambos extremos.

donde L es el largo del tubo y n ($n = 1, 2, 3, \dots$) es un natural al cual llamaremos número de armónico. Utilizando que

$$v = f\lambda \quad (7)$$

donde f es la frecuencia del tono y v la velocidad del sonido se llega a que las frecuencias en un tubo abierto en ambos extremos son

$$f = \left(\frac{v}{2L}\right)n \quad (8)$$

La ecuación 8 implica que las ondas estacionarias aparecen únicamente para determinadas frecuencias que terminan siendo función del número de armónico n (manteniendo la velocidad del sonido y el largo del tubo constantes), un factor fundamental a considerar para la práctica.

2.2.2. Tubo semi cerrado

Análogamente, si el tubo es semi cerrado como en la Figura 3, se puede demostrar que la longitud de onda sigue la relación

$$\lambda = \frac{4L}{n} \quad (9)$$

y por lo tanto la frecuencia es

$$f = \left(\frac{v}{4L}\right)n \quad (10)$$

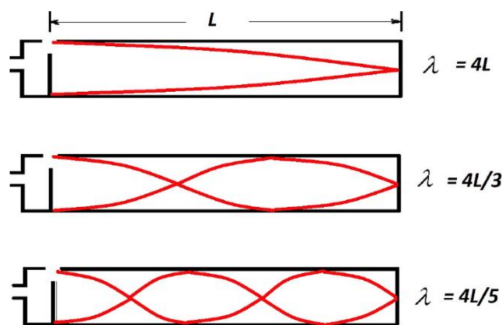


Figura 3: Representación de las ondas estacionarias en un tubo abierto por un solo extremo

3. Procedimiento experimental

3.1. Reflexión de pulso

El dispositivo consta de un tubo de cartón tapado en uno de sus extremos. En el extremo opuesto abierto se reproducirá un tono de corta duración¹, mientras que se recolectará la respuesta con un micrófono (capaz de registrar pequeñas variaciones en la presión del aire) en el mismo lado.

Los datos experimentales de esta parte de la práctica se tomarán con la ayuda del script 'Experimento 24: Velocidad del sonido' del programa Logger Pro que está configurado para empezar a medir inmediatamente luego de que el micrófono detecta una variación importante en la presión.

Se deberá diferenciar la señal emitida de la reflejada, analizando la variación de tiempo entre las cuales el micrófono detecta ambas señales, con esto y el largo del tubo se determinará la velocidad del sonido mediante la ecuación,

$$v = \frac{\Delta x}{\Delta t} \quad (11)$$

Se tomarán como mínimo 20 medidas y se realizará un tratamiento estadístico de las mismas para obtener un valor de Δt . Se deberá registrar la temperatura ambiente para luego realizar el cálculo de γ .

- Presentar la gráfica ilustrativa de la amplitud en función del tiempo, identificar y marcar en ella el tiempo t entre la señal original y su reflexión.
- Presentar el histograma de las medidas realizadas y obtener el valor de t con correspondiente incertidumbre obtenida a partir de tratamiento estadístico.
- Calcular v a partir del t de la parte anterior, con su incertidumbre correspondiente.
- Calcular de γ a partir del v anterior.

3.2. Ondas estacionarias

La idea de esta parte es utilizar la aplicación 'generador de frecuencia' o el archivo de audio² para ir variando las frecuencias y registrar las señales con el Logger Pro, en caso de usar la aplicación se sugiere variar las frecuencias desde 50Hz a 800Hz . Recordar que la frecuencia de resonancia va a detectarse al ver un máximo en la amplitud de la señal. Una vez que se identifica un máximo se deberá tomar el período local de la señal cerca del mismo y luego usar la ecuación,

$$f = \frac{1}{T} \quad (12)$$

Al final del experimento, se obtendrán una serie de valores de frecuencias y números de armónicos correspondientes. Luego, se procederá a graficar la primera variable en función de la segunda, y a partir de la pendiente de dicho gráfico se determinará la velocidad del sonido, como se muestra en la ecuación (8)

- Presentar la gráfica de f en función del número de armónico.
- Calcular v a partir de esa gráfica con su incertidumbre correspondiente.
- Calcular γ a partir del nuevo valor de estimación de v .

Discutan si los valores de la velocidad del sonido y de γ son concordantes con el valor de la referencia.

¹Descargar audio Reflexión de pulso

²Descargar audio Ondas Estacionarias

4. Bibliografía

- Física Universitaria, Sears, Francis W. and Zemansky, Mark W., Pearson Educación de México, 2009
- Experimentos de física con smartphones, Gómez González, Rubén and others, 2023

5. Apéndice

Simulador de ondas estacionarias <https://acortar.link/QKKzAy>



Física Experimental 1



Práctica 4 - Determinación del calor específico de un metal.

1. Objetivos

Esta actividad tiene como principal objetivo determinar el calor específico de un metal estudiando la transferencia de energía en forma de calor entre el metal y una masa de agua. Se espera que en esta actividad se puedan aplicar los conocimientos adquiridos hasta ahora en el curso en relación al análisis del modelo físico y se profundice en la discusión de los resultados obtenidos, y las fuentes de error en la experiencia.

2. Fundamento Teórico

Ley Cero de la Termodinámica y equilibrio térmico

Una de las variables más importantes (y utilizada de manera cotidiana) para la descripción de un sistema termodinámico es la temperatura. Definir la misma no es una tarea sencilla, y a menudo se puede llegar a perder rigurosidad cuando se incluyen en la discusión palabras como frío o calor.

Supongamos que se tienen dos cuerpos, A y B (que no necesariamente tienen que estar en contacto entre sí), y se desea estudiar si ambos se encuentran a la misma temperatura. Una manera de probar esto es poner en contacto térmico entre sí a los cuerpos A y B, para luego estudiar cómo varían a lo largo del tiempo diferentes propiedades termodinámicas que pueden caracterizarlos (volumen, presión, etc.). Si A y B estaban inicialmente a diferente temperatura, entonces las propiedades antes mencionadas variarán rápidamente al comienzo del contacto térmico, para luego ir estabilizándose a lo largo del tiempo. Cuando estas propiedades se mantengan constantes a lo largo del tiempo, se puede definir que los cuerpos están en *equilibrio térmico* entre sí, es decir, ambos tienen la misma temperatura.

Es posible realizar el mismo estudio a través de un tercer cuerpo (denominado C¹), sin necesidad de poner en contacto térmico los contactos A y B. Esto es resumido en la *ley cero de la termodinámica* que establece que *si los sistemas A y B están cada uno en equilibrio térmico con un tercer sistema C, entonces A y B están en equilibrio térmico entre sí*

¹Este tercer cuerpo puede ser definido como un termómetro.

Primera Ley de la Termodinámica

La primera ley de la termodinámica establece cómo varía la energía interna de un sistema (ΔU) debido al intercambio de energía en forma de calor (Q) y al intercambio de energía en forma de trabajo de frontera (W). Para un sistema cerrado (es decir, no entra ni sale masa del sistema en ningún momento), la Primera Ley se puede expresar como:

$$\Delta U = Q - W \quad (1)$$

A través del análisis microscópico de un sistema, es posible establecer cuánto vale la energía interna para líquidos, sólidos y gases. Para los primeros dos se puede escribir:

$$\Delta U = mc(T_f - T_i) \quad (2)$$

donde m representa la masa del sólido/líquido bajo estudio, c el calor específico y $T_f - T_i$ la diferencia de temperaturas entre el estado inicial (i) y el estado final (f).

Aplicación de la Primera Ley a un calorímetro

Un calorímetro (como el que va a ser utilizado en esta práctica) es un dispositivo que permite medir las cantidades de calor suministradas o recibidas por un cuerpo.

Es importante definir en este punto algunas hipótesis de trabajo. En primer lugar, vamos a suponer que el calorímetro no interviene en el proceso termodinámico, es decir, no absorbe ni libera calor (ni al ambiente ni a los cuerpos en su interior). En segundo lugar, vamos a suponer que el aire que quede dentro del calorímetro es nulo, de manera que tampoco interviene en el proceso termodinámico.

Supongamos que dentro del calorímetro se coloca en contacto térmico dos cuerpos A y B de masas m_A y m_B respectivamente. Cada uno se encuentra a una temperatura inicial $T_{A,i}$ y $T_{B,i}$. Después de cierto tiempo ambos cuerpos alcanzarán el equilibrio térmico a una temperatura de equilibrio T_{eq} .

Realizaremos un análisis aplicando la Primera Ley dentro del calorímetro, tomemos en cuenta la siguientes consideraciones:

- El sistema bajo estudio son los cuerpos A+B.
- El intercambio de calor con el ambiente es $Q = 0$, dado que el calorímetro ni absorbe ni libera calor al ambiente.
- El trabajo de frontera del sistema es $W = 0$, dado que el volumen dentro del calorímetro es constante.
- Aplicando la primera Ley: $\Delta U_{total} = \Delta U^A + \Delta U^B = Q - W = 0 \Rightarrow \Delta U^A + \Delta U^B = 0$

De las definiciones de energía interna:

$$\Delta U^A + \Delta U^B = m_A c_A (T_{eq} - T_{A,i}) + m_B c_B (T_{eq} - T_{B,i}) = 0 \quad (3)$$

Conociendo por lo tanto la temperatura final del sistema T_{eq} (por ejemplo, midiendo con un termómetro al momento del equilibrio térmico), es posible determinar el calor específico c_B del material bajo estudio (si se conocen los demás parámetros):

$$c_B = \frac{m_A c_A (T_{A,i} - T_{eq})}{m_B (T_{eq} - T_{B,i})} \quad (4)$$

3. Procedimiento experimental

En este experimento, a través del proceso de intercambio de calor entre cuerpos sólidos y agua, se determinará el calor específico de algunos metales (Aluminio y Bronce).

Dependiendo de qué magnitudes se elijan como variables y cuáles constantes, la ecuación 3 se puede reescribir de forma lineal o el modelo podrá tomar una expresión más compleja.

Por ejemplo, si se toman como variables la temperatura inicial del metal $T_{M,i}$ y la temperatura de equilibrio T_{eq} dicha ecuación se puede expresar como:

$$T_{eq} = b + aT_{M,i} \quad (5)$$

Pero si las variables son la temperatura inicial del agua $T_{ag,i}$, y la temperatura final de equilibrio T_{eq} , entonces la relación será

$$T_{eq} = d + cT_{ag,i} \quad (6)$$

Donde la dependencia de a , b , c y d con los parámetros fijos cambia según qué variables se elije.

Ejercicio 1 Determine los valores de las constantes a , b , c y d que aparecen en las ecuaciones 5 y 6

Se discutirá en el grupo acerca de distintas formas posibles de llevar a cabo la experiencia, y se elegirá una de ellas.

Cada grupo decidirá cómo realizará las medidas, qué magnitudes se tomarán como variables y cuáles se dejarán constantes. Deberá discutir cómo dejar dichas magnitudes constantes.

Se analizarán las limitaciones experimentales, qué factores se pueden controlar y cuáles no, y cómo estos pueden afectar en las medidas realizadas y en los resultados.

Algunas recomendaciones importantes para hacer las medidas:

- Predecir teóricamente los valores de T_{eq} esperados, según las condiciones iniciales a utilizar. Esto lo ayudará a discernir mejor cuando el sistema alcance el equilibrio.
- Usar del orden de 200 ml de agua en el calorímetro; o lo mínimo posible que deje el metal y el sensor de temperatura cubiertos.
- Tenga cuidado al manipular el agua y el metal caliente.
- Intente obtener como mínimo seis medidas para poder realizar un ajuste lineal.
- Tomar las precauciones necesarias para asegurar que las magnitudes que deban permanecer constantes no cambian a lo largo de la experiencia (y si esto no es posible analizar cómo esto influye en los resultados).
- Si los resultados son insatisfactorios, intente repetir el procedimiento o algunas medidas teniendo mayor cuidado en las condiciones de trabajo o mejorando alguna de las etapas del procedimiento.

4. Tratamiento de datos

A continuación se detallan algunos aspectos importantes para el análisis de datos y la discusión. Esta experiencia es la última del curso, hasta el momento se han introducido diferentes metodologías o conceptos necesarios para la realización de experimentos y comunicación de resultados como técnicas de medida, métodos de procesamiento. En particular el conjunto de ellos han sido reforzados en la práctica anterior. En este caso solo se utilizará el método de mínimos cuadrados para determinar el valor del calor específico del metal, y la propagación de incertidumbres para evaluar la incertidumbre del calor específico. Sin embargo, será de especial importancia el análisis y la discusión de los resultados, con énfasis en el proceso de medida, en cómo se pueden implementar mejoras, etc.

- Se utilizará el método de mínimos cuadrados para determinar el valor del calor específico del metal con su incertidumbre.
- Mostrar la gráfica de las medidas realizadas junto a la curva resultado del ajuste. Poner barras de errores en los datos experimentales.
- El sistema cumple el modelo físico? Discuta según el coeficiente de correlación.
- Calcular el calor específico del metal estudiado con su incertidumbre y comparar el mismo con su valor de referencia.
- Discutir cuales fueron las principales fuentes de incertidumbre en las de las medidas, además de como se podría mejorar el montaje experimental.
- Es importante que el análisis y la discusión sean profundos y exhaustivos.

5. Bibliografía

- Física, Resnick-Halliday-Krane.
- Física Re-Creativa, S. Gil y E. Rodríguez.

6. Apéndice

En la tabla 1 se muestran valores de referencia del calor específico para diferentes metales.

Material	Calor específico (J/kg K)	Calor específico (cal/g°C)
Aluminio	900	0.215
Bronce	385	0.092
Hierro	448	0.107
Aire	1006	0.241
Agua	4190	1.0
Hielo (agua estado sólido)	2095	0.5

Cuadro 1: Calor específico y calor latente de fusión para diferentes materiales.



Física Experimental 1



Práctica 4

Movimiento en un plano inclinado: determinación del coeficiente de fricción cinética.

1. Objetivos

Estudiar el movimiento de deslizamiento de un objeto por un plano inclinado, a partir del cual se determinará el coeficiente de fricción cinética entre las superficies del cuerpo y del plano sobre el cual desliza, aplicando diferentes herramientas de análisis estudiadas en el curso.

2. Fundamento teórico

2.1. Fricción estática y cinética

Desde un punto de vista dinámico, se pueden identificar interacciones que permiten describir el comportamiento de un cuerpo en determinadas condiciones. La interacción que da origen a la denominada fuerza de contacto es debida al contacto entre las superficies de dos cuerpos sólidos. Como se muestra en la figura 1, esta fuerza se suele descomponer en dos direcciones: en la dirección perpendicular (o normal) a la superficie de contacto llamada fuerza normal N y en la dirección paralela a la superficie de contacto llamada fuerza de rozamiento o fricción f . En la figura 1 también se representa una fuerza aplicada sobre la masa, F y la fuerza peso mg .

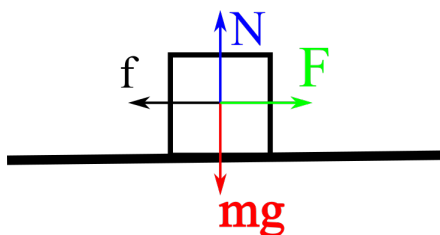


Figura 1: Diagrama de fuerzas sobre un cuerpo en reposo apoyado en un plano horizontal.

La fuerza de rozamiento entre superficies se debe a las irregularidades microscópicas de las superficies como se muestra en la figura 2 y el valor de los coeficientes de rozamiento asociados dependerán del material y de la rugosidad de las superficies.

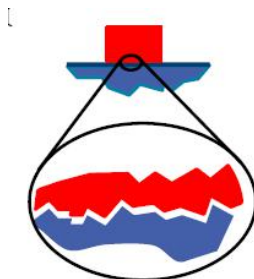


Figura 2: Representación de las irregularidades entre dos superficies en contacto.

Es posible distinguir dos tipos de fricción: estática f_e y cinética f_c . La fricción estática ocurre cuando las dos superficies en contacto se mantienen en reposo relativo entre sí y su valor es menor o igual a la normal por el coeficiente de rozamiento estático μ_e : $f_e \leq \mu_e N$. La fricción cinética ocurre cuando las dos superficies en contacto están en movimiento relativo entre sí y su valor es igual a la normal por el coeficiente de rozamiento cinético μ_c : $f_c = \mu_c N$.

Para ejemplificar lo anterior, considere un bloque de masa m , inicialmente en reposo y apoyado sobre una superficie horizontal, al cual se le aplica una fuerza F paralela a la superficie de apoyo, cuyo módulo aumenta lentamente desde 0 N . En la figura 3 se esquematiza la dependencia entre el módulo de la fuerza de rozamiento f y la fuerza variable F que se le aplica al bloque. A partir del momento en el cual la fuerza F supera a la fuerza de rozamiento estático máximo $f_{e(\text{máx})} = \mu_e N$, el cuerpo empieza a moverse y, por lo tanto, el cuerpo comienza a experimentar la fricción cinética f_c .

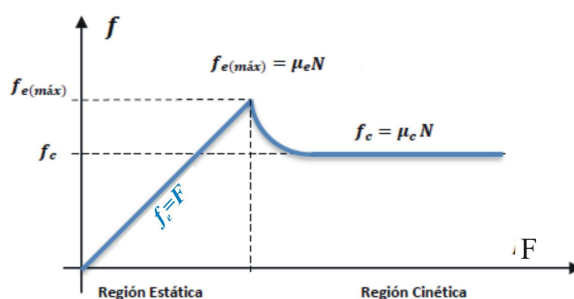


Figura 3: Fuerza de fricción en función de una fuerza F de módulo variable aplicada sobre el cuerpo.

2.2. Deslizamiento por un plano inclinado

Considere un objeto apoyado sobre un plano inclinado, formando un ángulo α con la horizontal. El objeto se mantendrá en reposo sobre el plano hasta que se alcance cierto ángulo α_1 , a partir del cual el cuerpo comienza a deslizarse. En la figura 4 se muestra el digrama de fuerzas sobre el cuerpo de masa m al variar el ángulo de inclinación hasta llegar a α_1 .

Una vez superado el ángulo crítico α_1 , la fuerza que actúa es la fricción cinética y el cuerpo estará deslizando con una cierta aceleración. Para obtener el coeficiente de rozamiento cinético entre las superficies en contacto, podríamos determinar la aceleración a que experimenta el cuerpo para un cierto ángulo α y a partir de ella calcular μ_c .

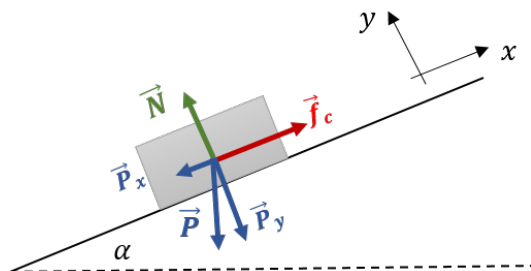


Figura 4: Diagramas de fuerzas para un objeto sobre un plano inclinado un ángulo α

- **Ejercicio 1:** Deducir la expresión del coeficiente de rozamiento cinético aplicando la 2ª ley de Newton.
- **Ejercicio 2:** Escribir la expresión para el desplazamiento $x(t)$ del cuerpo sobre el plano inclinado.

3. Procedimiento experimental

Diseñar un dispositivo que permita estudiar el movimiento de un cuerpo por un plano inclinado (ejemplo ilustrativo en la figura 5). Se determinará el coeficiente de fricción cinético a partir de la dos herramientas de análisis desarrolladas en el curso (mínimos cuadrados y estadística).

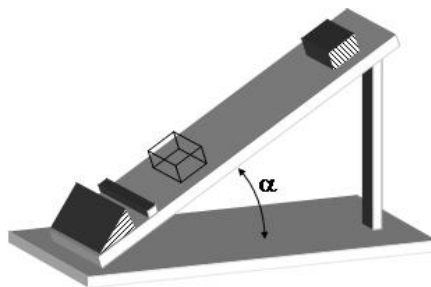


Figura 5: Esquema del dispositivo a utilizar.

1. Medida del coeficiente de rozamiento cinético μ_c .

- a) **Mínimos cuadrados:** A partir de la medida de $x(t)$, aplicar el método de mínimos cuadrados para validar el modelo cuadrático correspondiente a una aceleración constante. Luego de validar el modelo, obtener la aceleración del objeto y de ella el coeficiente de rozamiento cinético.
- b) Para determinar $x(t)$ se utilizará un sensor de posición Vernier, ya implementado en la práctica 3.
- c) **Tratamiento estadístico:** Discutir como debe diseñarse la experiencia para cumplir las hipótesis necesarias para realizar un tratamiento estadístico. ¿Cuál es la variable aleatoria que pueden medir?, y ¿cómo de dicha variable se puede obtener el coeficiente de rozamiento cinético?.
- d) Puede determinar la variable utilizando el mismo sensor que para el caso anterior.

4. Tratamiento de datos

1. Determine el valor del coeficiente de fricción cinético por el método estadístico y mínimos cuadrados.
2. Para el caso de medidas estadísticas: Verificar si las medidas (de la variable elegida, tiempo o aceleración) realizadas con el tratamiento estadístico siguen una distribución Gaussiana. Recordar que para que ambos gráficos puedan ser comparados deben encerrar la misma área.
3. Para el caso de ajuste por método de mínimos cuadrados: ¿Su sistema cumple el modelo físico? Discuta según el coeficiente de correlación.
4. Comparar el valor del coeficiente de fricción cinética determinado con los dos métodos. Discutir cual les parece el método más adecuado para medir μ_c .
5. Realizar una búsqueda bibliográfica de coeficientes de rozamiento obtenidos para materiales similares a los que usaron en la práctica. Comparar, si posible, los valores de μ obtenidos con los reportados para superficies similares.
6. Discutir cuáles serían las principales fuentes de incertidumbre en cada una de las medidas, además de como se podría mejorar el montaje experimental.

Anexo 4

Programa del curso Física Experimental 1



Programa de FÍSICA EXPERIMENTAL 1

1. FÍSICA EXPERIMENTAL 1

2. CRÉDITOS

5 créditos

3. OBJETIVOS DE LA UNIDAD CURRICULAR

Objetivo general de enseñanza:

Acercar a los estudiantes a la actividad experimental, a partir de propuestas de aula que incluyen: el diseño de experiencias prácticas de laboratorio; la utilización de diferentes herramientas de análisis de datos aplicadas a problemas experimentales; el cálculo de incertidumbres y la comunicación efectiva de resultados experimentales de forma escrita y oral.

Objetivos de aprendizaje:

Para las diferentes actividades experimentales propuestas durante el curso, se espera que los estudiantes logren:

- Reconocer las magnitudes físicas a medir y seleccionar los parámetros asociados al sistema físico de interés.
- Implementar el montaje experimental siguiendo los lineamientos propuestos.
- Aplicar una metodología para medir las magnitudes físicas y parámetros de la experiencia.
- Analizar los puntos débiles en el montaje experimental y/o el procedimiento y definir mejoras para los mismos.



- Identificar y analizar las diversas fuentes de incertidumbre en la realización de medidas de magnitudes físicas, y evaluar su impacto en el resultado.
- Utilizar herramientas de análisis de datos experimentales, en particular las estadísticas y de ajustes por mínimos cuadrados e interpretar los resultados obtenidos al aplicarlas, e interpretar los resultados obtenidos al aplicarlas.
- Reconocer los distintos instrumentos de medición, recordar sus características de funcionamiento y utilizarlos en la medida de cantidades físicas en el procedimiento experimental.
- Realizar y analizar gráficos de medidas experimentales y de modelos teóricos.
- Analizar los resultados experimentales e interpretarlos en comparación con el modelo físico propuesto, y obtener conclusiones sobre el diseño de la experiencia y sus resultados.
- Presentar resultados de forma escrita en un formato preestablecido, con varias instancias a lo largo del curso para mejorar y reforzar los aprendizajes.
- Elaborar una presentación y exponer de forma oral los resultados obtenidos en una de las experiencias.

4. METODOLOGÍA DE ENSEÑANZA

El curso se desarrolla en modalidad de laboratorio, donde se realizan actividades experimentales.. Los estudiantes, distribuidos en grupos de 12 - 15 estudiantes (aproximadamente) por docente, concurren, con una frecuencia quincenal, a una clase presencial de 3 horas de duración. Cada unidad quincenal se compone de un conjunto de recursos escritos y/o audiovisuales para ser estudiados previo a la instancia presencial. Durante cada quincena se estima una dedicación horaria fuera de clase de 7 hs. Durante las clases presenciales, los estudiantes trabajan preferentemente en parejas o tríos, y llevan adelante la actividad experimental pautada, con el acompañamiento y orientación docente. Se trata de una propuesta de trabajo activo, que busca la participación e involucramiento de los estudiantes. Se estimula el intercambio grupal para el análisis de las implicancias del modelo teórico en el diseño experimental y la discusión acerca de las diferentes posibilidades de llevar adelante la experiencia. Se pueden incluir en las semanas sin clase presencial, otras instancias de trabajo sincrónico para abordar algunos temas específicos, como por ejemplo medidas e incertidumbre, producción escrita, presentación de resultados, construcción de gráficos, entre otros de los objetivos del curso.



Otras actividades a realizar durante el curso y que forman parte de la evaluación corresponden a: cuestionarios de evaluación y autoevaluación, entrega de informes escritos la semana posterior a cada actividad experimental, presentación oral de un tema asignado en la última clase del curso. Se pueden incluir otras instancias de evaluación continua durante las clases en forma escrita u oral.

Se utiliza la plataforma EVA para disponer las consignas de trabajo y los materiales. Se dispone en EVA un foro de consultas atendido por el equipo docente, donde además se promueve el intercambio entre estudiantes. Además, se ofrecen clases de consulta presenciales y/o virtuales sincrónicas durante todo el semestre.

5. TEMARIO

El curso aborda cuatro temas principales con relación al tratamiento de datos experimentales, que se desarrollan durante la primera mitad del curso.

1. Mediciones e incertidumbres
2. Modelado de problemas físicos.
3. Tratamiento estadístico de datos
4. Método de Mínimos cuadrados

Además, cada año se seleccionan 4 o 5 prácticas experimentales que corresponden a contenidos incluidos en los programas de Física 1 y 2. Como ejemplo, las actividades experimentales pueden incluir temas relativos a: leyes de Newton, movimientos oscilatorios, leyes de los gases, transferencia de calor entre otros.

6. BIBLIOGRAFÍA

Para todos los temas del curso se proporcionan materiales en formato escrito y/o audiovisual, realizados por el equipo docente.



Tema	Básica	Complementaria
Mediciones e incertidumbres	(1,3,4)	
Modelado de problemas físicos.	(1)	
Tratamiento estadístico de datos	(1)	
Método de Mínimos cuadrados	(1)	
Temas correspondientes al contenido de las actividades experimentales	(1,2)	

6.1 Básica

1. Notas del curso y videos elaborados por el equipo docente.
2. R. Resnick, D. Halliday and K. Krane, Física 1 (CECSA, 4ta. edición)
3. J. R. Taylor, Introducción al análisis de errores: El estudio de las incertidumbres en las mediciones físicas, Reverté, 4ta. Ed, 2018.
4. Joint Committee for Guides in Metrology (JCGM) - BIPM, Evaluation of measurement data — Guide to the expression of uncertainty in measurement, 2008.
5. G. Perera, Probabilidad y estadística, Ed. Fin de siglo, 2011.

6.2 Complementaria

Se puede sugerir bibliografía complementaria de acuerdo a las actividades experimentales que se realicen en cada edición, haciéndola disponible en la página del EVA del curso.

7. CONOCIMIENTOS PREVIOS EXIGIDOS Y RECOMENDADOS

7.1 Conocimientos Previos Exigidos: Conocimientos de los contenidos

correspondientes al examen de la UC Física 1 y del curso de la UC Física 2.

7.2 Conocimientos Previos Recomendados: Conocimientos de Cálculo diferencial e integral a nivel básico.

ANEXO A

Para todas las Carreras

Esta primera parte del anexo incluye aspectos complementarios que son generales de la unidad curricular.

A1) INSTITUTO

Instituto de Física

A2) CRONOGRAMA TENTATIVO

Las actividades experimentales pueden variar en diferentes ediciones del curso, manteniéndose dentro de los temas previstos en el programa.
Todas las clases son de 3 horas.

Semana 1	Clase inicial- Presentación del curso, armado de grupos. Tema: proceso de medida e incertidumbres
Semana 2	Trabajo fuera de clase - Entrega de ejercicios
Semana 3	Tema 2- Tratamiento estadístico - Práctica experimental 1
Semana 4	Trabajo fuera de clase - Realización del informe
Semana 5	Tema 3 - Método de mínimos cuadrados Práctica experimental 2
Semana 6	Trabajo fuera de clase - Realización del informe
Semana 7	Clase sobre discusión de informes y presentaciones
Semana 8	Trabajo fuera de clase - Entrega de una tarea



Semana 9	Práctica experimental 3
Semana 10	Trabajo fuera de clase - Realización del informe
Semana 11	Práctica experimental 4
Semana 12	Trabajo fuera de clase - Realización del informe
Semana 13	Presentaciones orales
Semana 14	Presentaciones orales 2

A3) MODALIDAD DEL CURSO Y PROCEDIMIENTO DE EVALUACIÓN

Esta unidad curricular tiene asistencia obligatoria a clase. Sólo admite la aprobación del curso, sin opción de examen posterior.

Todas las actividades con evaluación se corrigen en base 10, con 5 como puntaje mínimo de aprobación.

Los informes grupales se corrigen utilizando pautas o rúbricas de evaluación, con criterios establecidos y publicados en el EVA del curso al momento de realizar cada experiencia.

Las actividades individuales escritas pueden ser de múltiple opción o de desarrollo, teniendo como objetivo ayudar a los estudiantes en el seguimiento del curso y a la preparación previa de cada clase.

El peso relativo de cada ítem evaluado está dado de la siguiente forma:

Actividades grupales: 75 % (incluye informes, otras entregas grupales, manejo del dispositivo experimental, presentación final)

Actividades individuales: 25 % (cuestionarios, instancias de evaluación orales o escritas, involucramiento en la actividad experimental, actitud proactiva a lo largo del curso)

Al comienzo de cada curso se comunica el detalle de las actividades a desarrollar.

Cada grupo de actividades de evaluación tiene un máximo de actividades reprobadas, de acuerdo con la cantidad de actividades totales que tenga cada edición del curso.



A4) CALIDAD DE LIBRE

No corresponde calidad de libre.

A5) CUPOS DE LA UNIDAD CURRICULAR

No tiene cupo.

ANEXO B para la(s) carrera(s) XXX

(Un anexo distinto para cada carrera que tome la unidad curricular. En caso de que a dos o más carreras les corresponda información idéntica en este anexo, se utilizará el mismo anexo, explicitando cuáles son todas esas carreras.)

Esta(s) parte(s) del anexo incluye(n) los aspectos que son particulares de cada carrera que tome la unidad curricular.

B1) ÁREA DE FORMACIÓN

El área de formación (materia, según la anterior nomenclatura) identifica las grandes áreas temáticas ligadas a un sector de la ciencia o de la técnica. Cada comisión de carrera evaluará a qué área de formación corresponde la unidad curricular.

B2) UNIDADES CURRICULARES PREVIAS

Curso:

Examen: