



*Universitat
Abat Oliba CEU*

Sesgo cognitivo en la evaluación de contenido en Educación Secundaria

TRABAJO DE FINAL DE MÁSTER

Autor: Aniol Castany Feixas

Tutora: Dra. Franciele Corti

Máster universitario en Formación del Profesorado de Educación Secundaria Obligatoria
y Bachillerato, Formación Profesional y Enseñanza de Idiomas

Año: 2025

DECLARACIÓN

Declaro que el material de este documento, que ahora presento, es fruto de mi propio trabajo. Cualquier ayuda recibida de otras personas ha sido citada y reconocida dentro de este documento. Hago esta declaración sabiendo que incumplir las normas relativas a la presentación de trabajos puede acarrear consecuencias graves. Soy consciente de que el documento no se aceptará salvo que se entregue con esta declaración.

Firma:

Aniol CASTANY FEIXAS

*“Si vols construir un vaixell, no comencis per buscar fusta, tallar taules o distribuir el treball;
primer has d'evocar en els homes l'anhel de mar lliure i ample”*

ANTOINE DE SAINT-EXUPÉRY

Resumen

La evaluación de contenido es una parte fundamental en el proceso de enseñanza-aprendizaje. Puede verse influenciado por la aparición de sesgo cognitivos, que son una distorsión sistemática e inconsciente que afecta a los juicios, la toma de decisiones y el procesamiento de la información. Existen varios tipos de sesgos cognitivos, como el sesgo de anclaje, el efecto halo/cuerno, el sesgo de comparación, el sesgo de confirmación y el sesgo de afinidad, entre otros. Para determinar si este fenómeno está presente en la corrección de contenido en Educación Secundaria, se diseña una investigación que compara las calificaciones obtenidas por los estudiantes antes y después de aplicar técnicas de reducción de sesgo cognitivo.

Los resultados confirman la presencia de sesgo cognitivo, incrementando la nota en un 6,5% de media. Además, se favorece al 40% del alumnado, aunque no se demuestra que dependa del rendimiento académico previo. Entre los sesgos más presentes destacan el sesgo de anclaje, el efecto cuerno y el sesgo de confirmación (de expectativas negativas). Al contrario de lo esperado, la comparación entre alumnos reduce el sesgo cognitivo. Por último, los alumnos con más afinidad con el profesorado son los más perjudicados.

Con todo, se pone de manifiesto la necesidad de implementar estrategias de reducción del sesgo cognitivo entre el profesorado para que el alumnado pueda confiar en un proceso de evaluación justo e imparcial que refleje verdaderamente su rendimiento académico.

Resum

L'avaluació de contingut és una part fonamental en el procés d'ensenyament-aprenentatge. Pot veure's influenciat per l'aparició de biaixos cognitius, que són una distorsió sistemàtica i inconscient que afecta als judicis, la presa de decisions i el processament de la informació. Existeixen diversos tipus de biaixos cognitius, com el biaix d'ancoratge, l'efecte halo o banya, el biaix de comparació, el biaix de confirmació i el biaix d'afinitat, entre altres. Per a determinar si aquest fenomen està present a la correcció de contingut a l'Educació Secundària, es dissenya una investigació que compara les qualificacions obtingudes pels estudiants abans i després d'aplicar tècniques de reducció del biaix cognitiu.

Els resultats confirmen la presència del biaix cognitiu, incrementant la nota en un 6,5% de mitjana. A més, s'afavoreix al 40% de l'alumnat, encara que no es demostra que depengui del rendiment acadèmic previ. Entre els biaixos més presents destaquen el biaix d'ancoratge, l'efecte banya i el biaix de confirmació (d'expectatives negatives). Al contrari del que s'esperava, la comparació entre alumnes redueix el biaix cognitiu. Per últim, els alumnes amb més afinitat amb el professorat són els més perjudicats.

Amb tot, es posa de manifest la necessitat d'implementar estratègies de reducció del biaix cognitiu entre el professorat per tal que l'alumnat pugui confiar en un procés d'avaluació just i imparcial que reflecteixi veritablement el seu rendiment acadèmic.

Abstract

Content assessment is a fundamental part of the teaching-learning process. It can be influenced by the emergence of cognitive biases, which are systematic and unconscious distortions that affect judgment, decision-making, and information processing. There are several types of cognitive biases, such as anchoring bias, the halo/horn effect, comparison bias, confirmation bias, and affinity bias, among others. To determine whether this phenomenon is present in content correction in high school, a study is designed to compare the grades obtained by students before and after applying cognitive bias reduction techniques.

Results confirm the existence of cognitive bias, increasing a 6,5% the grade's average. Moreover, 40% of the students are favored, although there is no evidence that this phenomenon is affected by previous academic achievements. Among the most prevalent biases, there are anchoring bias, the horn effect and confirmation bias (of negative expectations). Contrary to what was expected, comparison between students reduces cognitive bias. Finally, the higher affinity students have with the teachers, the more disadvantaged they are.

However, it is evidenced that strategies to reduce cognitive bias among teachers need to be implemented, so that students can trust a fair and impartial assessment process that truly reflects their academic performance.

Palabras clave / Paraules clau / Keywords

Sesgo cognitivo – evaluación – alumnado – profesorado

Biaix cognitiu – avaluació – alumnat – professorat

Cognitive bias – assessment – students – teachers

Índice

1	Introducción	10
2	Marco teórico	12
2.1	El sesgo cognitivo	12
2.1.1	Definición de sesgo cognitivo	12
2.1.2	Teorías sobre el sesgo cognitivo	13
2.1.3	Tipos de sesgos cognitivos	17
2.1.4	El fenómeno de estudio	21
2.2	Estado de la cuestión	21
2.3	Vulnerabilidad al sesgo cognitivo en la evaluación de contenido	23
2.4	Mitigación de los sesgos cognitivos	29
2.4.1	Causas que propician los sesgos cognitivos en contextos educativos	29
2.4.2	Estrategias transversales para la mitigación de los sesgos cognitivos	30
2.4.3	Estrategias específicas para la mitigación de los sesgos cognitivos	31
3	Metodología	35
3.1	Formulación de la pregunta de investigación e hipótesis	35
3.2	Objetivos	36
3.3	Diseño de la investigación	36
3.4	Técnica de muestreo y muestra del estudio	37
3.5	Instrumentos y técnicas de recogida de datos	38
4	Resultados	39
4.1	Perfil de la muestra	39
4.2	Análisis de resultados sobre el primer objetivo	39
4.3	Interpretación de los resultados	45
4.4	Análisis de resultados sobre el segundo objetivo	48
5	Conclusiones	54
5.1	Revisión de los objetivos y discusión	54
5.2	Limitaciones de la investigación	55
5.3	Líneas prospectivas	56
6	Bibliografía	57
	ANEXOS	60

Anexo 1. Tabla de dimensiones del cuestionario.....	61
Anexo 2. Resultados de la doble corrección.....	63
Anexo 3. Respuestas del cuestionario	66

Índice de figuras

Figura 1 Frecuencia absoluta de estudiantes en función del % de variación de los resultados (Total).....	41
Figura 2 Frecuencia absoluta de estudiantes en función del % de variación de los resultados (Grupo A).....	44
Figura 3 Frecuencia absoluta de estudiantes en función del % de variación de los resultados (Grupo B).....	44
Figura 4 Frecuencia absoluta de estudiantes en función del % de variación de los resultados (Grupo C).....	45
Figura 5 Resultados relativos al efecto halo/cuerno.....	50
Figura 6 Resultados relativos al sesgo de confirmación.	51
Figura 7 Resultados relativos al sesgo de afinidad.	52
Figura 8 Resultados relativos al sesgo de anclaje.	53
Figura 9 Resultados relativos al sesgo de comparación.	53

Índice de tablas

Tabla 1 Clasificación de los tipos de evaluación.	29
Tabla 2 Estrategias específicas de mitigación de los sesgos cognitivos más presentes en la evaluación de contenido.	33
Tabla 3 Muestra de profesores y alumnos.	38
Tabla 4 Muestra de los alumnos según rendimiento académico.	39
Tabla 5 Puntuaciones en la primera y segunda corrección y diferencia entre ellas, en promedio.....	40
Tabla 6 Frecuencias absolutas de los rangos de diferencia entre la primera y la segunda corrección.....	41
Tabla 7 Puntuaciones en las dos correcciones según el tipo de estudiante.	42
Tabla 8 Frecuencia absoluta de estudiantes en función del % de variación de los resultados.	43
Tabla 9 Resultados estadísticos de la prueba t-student para la muestra total.	46
Tabla 10 Resultados estadísticos de la prueba t-student para el grupo A.	47
Tabla 11 Resultados estadísticos de la prueba t-student para el grupo B.	47
Tabla 12 Resultados estadísticos de la prueba t-student para el grupo C.	47

Tabla 13	Resumen de las respuestas E2, E4-E10 y E12-E13 en frecuencias absolutas. ...	49
Tabla 14	Resumen de las respuestas E2, E4-E10 y E12-E13 en porcentajes.....	49
Tabla 15	Resumen de las respuestas E3 y E11. Media aritmética.	50
Tabla 16	Tabla de dimensiones del cuestionario.....	61
Tabla 17	Resultados de la doble corrección.....	63
Tabla 18	Respuestas del cuestionario.	66

1 Introducción

El proceso de evaluación es una parte fundamental en cualquier sistema educativo, ya que determina el rendimiento académico de los estudiantes, orienta las decisiones pedagógicas y tiene un impacto directo en el aprendizaje. Sin embargo, los métodos de evaluación no siempre garantizan un juicio imparcial o justo, dado que los evaluadores —en este caso, los profesores— son susceptibles a la influencia de diversos sesgos cognitivos. Estos sesgos son distorsiones sistemáticas en el juicio y la toma de decisiones que afectan a la objetividad, y que pueden influir en cómo los docentes interpretan y califican el rendimiento de sus alumnos.

Aunque el sesgo cognitivo se ha estudiado ampliamente en otros campos, como la psicología, la economía y la toma de decisiones, su impacto en el contexto educativo, y en particular, en la educación secundaria, sigue siendo un área de investigación menos explorada. La elección de este tema es una respuesta a la experiencia vital del autor y a la gama de experiencias que ha acumulado entorno a la temática de la evaluación de contenido, así como la correcta aplicación de la justicia. En concreto, surge de la percepción de que los estudiantes con mejores calificaciones tienden a ser evaluados de manera más favorable. Asimismo, se observa que los estudiantes con bajo desempeño académico suelen ser calificados con mayor severidad. Por ello, esta investigación se centra en investigar cómo los sesgos afectan la evaluación del contenido en la educación secundaria, abordando la influencia en los resultados académicos de los estudiantes.

Investigar este fenómeno, por tanto, puede ofrecer implicaciones relevantes para la comunidad educativa. En primer lugar, se trata de una cuestión de justicia que afecta a todos los estudiantes; los que más sobresalen, los que, a menudo, pasan desapercibidos, y lo que les queda mucho camino por recorrer para llegar a los logros mínimos. En segundo lugar, es beneficioso para docentes, puesto que, al ser más conscientes de este fenómeno pueden reducir su impacto, ofreciendo una mayor objetividad y rigurosidad durante la tarea evaluativa. Finalmente, el estudio de los sesgos cognitivos en la evaluación también tiene implicaciones en la sociedad y en el modelo educativo, especialmente en la formación de los futuros profesores y en el diseño de sistemas de evaluación más transparentes y objetivos.

Con todo, se pretende dar respuesta a la pregunta: ¿existen sesgos cognitivos en la evaluación de contenido en Educación Secundaria? Además, se pretende cuantificar su impacto para esclarecer si las calificaciones reflejan exclusivamente el desempeño académico o si están condicionadas por factores ajenos a los contenidos evaluados.

El trabajo se estructura en tres partes principales. Primero, se desarrolla el marco teórico que aborda los conceptos clave sobre sesgos cognitivos y su relevancia en la evaluación

educativa. Luego se presenta la metodología, donde se definen los objetivos específicos y las hipótesis de partida, así como los instrumentos utilizados para llevar a cabo la investigación. Por último, se presentan, analizan y discuten los resultados obtenidos, generando así las conclusiones que permiten responder a la pregunta de investigación original.

2 Marco teórico

El estudio de cualquier disciplina exige el conocimiento de la definición de la misma. Por eso, se empieza el marco teórico abordando el concepto de sesgo cognitivo, así como las principales teorías desarrolladas en las últimas décadas, las cuales, a su vez, moldean el propio concepto de sesgo cognitivo. Después se presentan los tipos de sesgo cognitivo más relevantes y se describe con exactitud el fenómeno de estudio. A continuación, se expone el estado de la cuestión, presentando las últimas aportaciones en relación al fenómeno de estudio, lo cual permite ver la línea de investigación a trazar. Finalmente, se relaciona la evaluación de contenido con la vulnerabilidad al sesgo cognitivo y se presentan estrategias para la reducción de este fenómeno.

2.1 El sesgo cognitivo

2.1.1 Definición de sesgo cognitivo

El término sesgo cognitivo fue utilizado por los psicólogos Amos Tversky y Daniel Kahneman en 1974 y lo definieron como un fenómeno psicológico que se refiere a las distorsiones sistemáticas que afectan el juicio, la toma de decisiones y el procesamiento de la información. Se puede afirmar que los sesgos son producto de atajos mentales, conocidos como heurísticas, que el cerebro utiliza para simplificar el procesamiento de información en situaciones complejas o ambiguas. Si bien las heurísticas son necesarias para manejar la sobrecarga cognitiva, pueden dar lugar a errores sistemáticos en el razonamiento, llevando a decisiones que no son objetivas ni racionales (Tversky y Kahneman, 1974).

A esta definición se puede añadir que los sesgos cognitivos son consecuencia de la racionalidad limitada del ser humano, quien se ve obligado a tomar decisiones dentro de limitaciones de tiempo, información y recursos cognitivos (Simon, 1947). En este sentido, también se expresan otros autores (Nisbett y Ross, 1980) cuando definen el sesgo cognitivo como el error sistemático que surge de la interacción entre los procesos cognitivos limitados y las influencias sociales, culturales y contextuales.

Es interesante detectar que no siempre se define el sesgo cognitivo como un error, como en el caso de Tversky y Kahneman (1974) o Nisbett y Ross (1980), pero de igual forma lo entiende Slovic (1987) cuando afirma que se producen heurísticas y sesgos en la evaluación de la percepción del riesgo. Es decir, en contextos de incertidumbre, afirma que se producen atajos mentales que distorsionan la realidad. Así lo entienden también Todd y Gigerenzer (2000), cuando señalan que los sesgos cognitivos no son necesariamente errores, sino subproductos de heurísticas que han evolucionado para ser útiles en la mayoría de las circunstancias, aunque pueden fallar en algunos contextos específicos.

En resumen, se puede definir un sesgo cognitivo como una (i) distorsión (ii) sistemática (iii) de la información disponible, del juicio objetivo o de una decisión racional (iv) generado por las limitaciones cognitivas.

- (i) El sesgo como una distorsión. La realidad es percibida o entendida de forma distinta a la que es. Esta afirmación presupone que la realidad es verdadera y que se puede conocer.
- (ii) El sesgo es sistemático. Esto indica que bajo unas mismas condiciones y un mismo contexto, siempre se produce el efecto de la distorsión. Consecuentemente, el sesgo no es intencional, sino inconsciente.
- (iii) Se puede distorsionar la información misma, el juicio que se hace sobre esta información o la decisión que se ha tomado en base al juicio hecho.
- (iv) El sesgo cognitivo está arraigado en el mismo proceso cognitivo, es decir, en las formas en que las persona perciben, procesan, interpretan y recuerdan la información. En última instancia, es inherente a la misma naturaleza humana y en la forma en cómo la mente opera.

Si bien es correcto afirmar que estas características pueden constituir una definición consistente de los sesgos cognitivos, esta puede ser matizada bajo distintas perspectivas, según las teorías que lo explican.

2.1.2 Teorías sobre el sesgo cognitivo

De las múltiples teorías que existen para explicar el fenómeno del sesgo cognitivo, seguramente la más influyente es la de Tversky y Kahneman (1974), que abordó por primera vez este concepto. En la década de 1970 publicaron su primer estudio, tratando de explicar la naturaleza de los sesgos cognitivos, cómo se originan y cómo afectan a la toma de decisiones. Posteriormente, varios autores propusieron su visión u ofrecieron críticas a estos primeros autores. A continuación, se presentan algunas de las teorías más relevantes.

1. Teoría de las Heurísticas de Tversky y Kahneman

Tal y como se ha mencionado, Tversky y Kahneman fueron pioneros en el estudio del sesgo cognitivo. La primera teoría que formularon fue la de las heurísticas. Esto significa que las personas utilizan atajos mentales para procesar la información de manera más rápida y eficiente para poder tomar decisiones, especialmente en aquellas situaciones donde hay incertidumbre o falta algún tipo de información que se debe completar. No obstante, en el proceso de completar esta información, se pueden cometer errores sistemáticos. Estos errores sistemáticos son precisamente lo que llamaron sesgos cognitivos.

En concreto, Tversky y Kahneman (1974), en su primer estudio *Judgment under Uncertainty: Heuristics and Biases*, identificaron dos heurísticas comunes en el campo cognitivo, que son

la heurística de representatividad y la de disponibilidad. La primera consiste en juzgar la probabilidad de un suceso basándose en lo similar que es a un prototipo conocido. Por ejemplo, en el artículo se realiza un experimento, describiendo a una persona cerrada, que no le gusta interaccionar, con gran aprecio al orden, apasionado por el detalle... Luego se pregunta a la gente a qué profesión pertenece, si a un granjero o a un bibliotecario. Es sorprendente como la mayoría responden que la descripción pertenece a un bibliotecario, aunque haya muchos más granjeros que bibliotecarios en Estados Unidos. Es decir, estadísticamente, la probabilidad de acertar escogiendo granjero es más elevada que eligiendo bibliotecario, sin embargo, la mayoría escoge bibliotecario.

La segunda heurística, la de disponibilidad, consiste en evaluar la probabilidad de un suceso por la facilidad con la que vienen a la mente ejemplos semejantes. Un ejemplo que ponen los mismos autores es que se puede evaluar (de forma sesgada) el riesgo de ataque al corazón en personas de mediana edad recordando los acontecimientos similares entre los conocidos que uno tiene. Es evidente que hay muchos factores que pueden influir en tal evaluación, como por ejemplo, que el sujeto recuerde más vivamente un ataque al corazón de un familiar (que si lo hubiera leído en una estadística), o la facilidad con la que puede imaginar tal suceso, que es mucho más elevada que, por ejemplo, imaginar otro tipo de enfermedad mucho menos común.

Por todo ello, se puede afirmar que las heurísticas de representatividad y de disponibilidad, frecuentes en el ámbito cognitivo, pueden originar sesgos de esta misma índole. Ambas teorías son comúnmente aceptadas, por explicar adecuadamente el origen de los sesgos cognitivos. Además, la teoría de las heurísticas es la más extendida de entre todas las teorías que intentan explicar el fenómeno de los sesgos cognitivos, hasta tal punto que, hoy en día, se puede hallar en diversos campos de la medicina o la salud. Por ejemplo, el diccionario médico de la Clínica Universidad de Navarra (2023) define el sesgo cognitivo de este modo:

El sesgo cognitivo es un fenómeno estudiado tanto en psicología como en medicina, que hace referencia a la tendencia sistemática de pensar de manera distorsionada debido a la interpretación sesgada de la información disponible. Se trata de una desviación de la toma de decisiones racionales o del juicio objetivo, lo que puede influir significativamente en las percepciones y actitudes de los individuos hacia diversos aspectos de la vida, incluyendo la salud y el bienestar.

Aunque los sesgos cognitivos pueden parecer errores o defectos en el pensamiento, en realidad son producto de una serie de atajos mentales o heurísticas que el cerebro utiliza para procesar información de manera más eficiente. Dado que el cerebro humano tiene una capacidad limitada para procesar información, estas heurísticas pueden ser útiles para tomar decisiones rápidas en situaciones complejas o inciertas. Sin embargo, también pueden llevar a errores sistemáticos o sesgos en el pensamiento. (párr. 1-2)

2. Teoría del Anclaje y Ajuste

Otra teoría clave en el estudio de los sesgos cognitivos es la teoría del anclaje y ajuste, que fue desarrollada por los mismos autores que la teoría de las heurísticas. Tversky y Kahneman (1974) afirman que “en muchas situaciones, las personas hacen estimaciones a partir de un valor inicial que es ajustado para obtener la respuesta final” (p. 1128). Este valor inicial es *el ancla*. El problema de este método radica en que el ajuste posterior normalmente es insuficiente (Epley y Gilovich, 2006) debido, precisamente, al sesgo cognitivo subyacente.

Este sesgo se manifiesta en muchos contextos, como en las negociaciones o en la toma de decisiones económicas. En el campo de la educación, por ejemplo, cuando un profesor recibe una evaluación de un estudiante como sobresaliente en un área puede, en consecuencia, sobrevalorar sus calificaciones en otros aspectos de su desempeño, debido a la influencia del primer juicio.

3. Teoría de la racionalidad limitada

Antes de que Tversky y Kahneman desarrollaran sus teorías, Simon (1947) expuso la teoría de la racionalidad limitada. Aunque pueda parecer una teoría sencilla, que parte del sentido común, su originalidad radica en que fue el primero que la aplicó en el campo de las ciencias sociales y, en concreto, en el ámbito de la economía.

Simon (1947) sostiene que la racionalidad de las personas está limitada por tres dimensiones fundamentalmente. La primera, es la información disponible. La segunda, es la limitación cognitiva de la persona que quiere conocer. Y la tercera es el tiempo del que se dispone para tomar la decisión. Todo ello le sirvió para poner en entredicho que las personas maximizan la utilidad de sus decisiones de manera exhaustiva. Por el contrario, sostiene que a veces es difícil evaluar todas las opciones disponibles y sus consecuencias, por lo que es preferible hacer una decisión más rápida y menos acurada, que gastar el tiempo en evaluar todas las opciones posibles para conseguir un resultado perfecto. En consecuencia, cuando se toman estas decisiones más rápidas (y menos acuradas), se pueden producir sesgos cognitivos. El autor puntualiza que estos fenómenos no se pueden considerar simplemente errores, sino que son adaptaciones que el ser humano tiene que hacer para adaptarse al ambiente que le rodea, que a menudo es complejo, y al pensamiento humano, que es siempre limitado.

4. Teoría de la Cognición Social y los Sesgos Motivacionales

Fiske y Taylor (1991) desarrollaron una teoría que va más allá de la teoría de Simon (1947). Propuso que los sesgos cognitivos no solo se originan por las limitaciones en el procesamiento de la información, sino que también se pueden estar influenciados por factores distintos al cognitivo. En particular, planteó que los factores sociales y emocionales tienen un peso relevante en estos procesos.

Los autores desarrollan esta teoría afirmando que las personas a menudo hacen juicios, no en base a una racionalidad (más o menos afectada por las limitaciones cognitivas), sino que tienen una preferencia por aquellas decisiones, juicios o valoraciones que refuerzan sus propios pensamientos o actitudes. Esto significa que están más predispuestos a aceptar teorías menos sólidas, pero que coinciden con sus valores a otras teorías, que pudieran ser más fundamentadas, pero que no comparten. Precisamente esto es lo que teorizó, Klayman (1995), bajo el nombre de sesgo de confirmación.

En el campo de la educación, este sesgo puede ser muy relevante, puesto que los docentes pueden favorecer a estudiantes que compartan sus creencias o valores, o que sean más afines a ellos, mientras que pueden discriminar a otros que estén más alejados de sus posturas.

5. Teoría del Procesamiento Dual

Esta teoría fue planteada por dos autores de forma paralela. El primero, Kahneman (2011), en su continua investigación sobre el procesamiento de la información y la toma de decisiones, motivo por el cual ganó el premio Nobel en 2002, y el segundo, Evans (2008). Ambos estudiaron los procesos de cognición y descubrieron que el cerebro humano procesa la información de dos modos básicos, es decir, existen dos sistemas de pensamiento. Por un lado, identificaron el Sistema 1, que es rápido, automático y basado en intuiciones (o heurísticas) y es responsable de la mayoría de las actividades que requieren de una respuesta en un corto intervalo de tiempo. Por otra, identificaron el Sistema 2, que es opuesto al Sistema 1, que es más lento, deliberado y consciente, por lo que es el sistema idóneo para las actividades que precisan de mayor análisis o que son más complejas.

Si bien es verdad que los sistemas de pensamiento no son excluyentes, normalmente en las decisiones predomina un sistema sobre el otro. Por tanto, cuando predomina el Sistema 1 sobre el Sistema 2, existe mayor probabilidad de que se produzcan sesgos cognitivos, dadas las características inherentes a este (rapidez, automatismo e intuición).

6. Teoría de la racionalidad ecológica

Todd y Gigerenzer (2000) definen la teoría de la racionalidad ecológica como un “comportamiento adaptativo resultante del ajuste entre los mecanismos de la mente y la estructura del entorno en el que opera” (p. 728). Los autores establecen diálogo con Simon en el contexto de la teoría de la racionalidad limitada y con Tversky y Kahneman (1974), en cuanto a las heurísticas. Precisamente en este término es donde divergen Todd y Gigerenzer (2000) y Tversky y Kahneman (1974) ya que los primeros critican la visión que tienen los últimos del concepto de heurística. Dentro de la teoría de la racionalidad ecológica, Todd y Gigerenzer (2000) no consideran las heurísticas como simples atajos mentales que generan sesgos cognitivos, sino que sostienen que son efectivas en el contexto adecuado. En concreto, los autores distinguen entre heurísticas rápidas y frugales que, aunque son

mecanismos simples, son altamente útiles para realizar decisiones con un esfuerzo cognitivo bajo.

2.1.3 Tipos de sesgos cognitivos

Con todo lo anterior, se observa que cada teoría está enfocada en explicar un tipo de sesgo cognitivo. O, dicho de otro modo, las teorías cognitivas muestran una faceta del fenómeno del sesgo cognitivo, pues tratan de explicar sus causas y las distintas manifestaciones posibles. Aunque los sesgos cognitivos son numerosos y a veces parecidos, a continuación se presentan algunos de los más relevantes.

1. Sesgo de representatividad

Este tipo de sesgo explicado anteriormente surge de la teoría de las heurísticas de representatividad (Tversky y Kahnema, 1974). Tal y como se ha explicado, el sesgo de representatividad consiste en juzgar la probabilidad de un suceso basándose en la similitud a un prototipo conocido. Es decir, en lugar de basarse en la racionalidad o en la estadística, se sigue a un prototipo o una categoría preexistente.

Este sesgo puede dar lugar a distorsiones en cualquier tipo de contexto, pero especialmente en educación, puede ser perjudicial. Por ejemplo, si al corregir un examen un profesor juzga a un alumno como excelente, por la respuesta de las primeras preguntas, puede llevarle a pensar que el estudiante cumple con el perfil o prototipo de estudiante brillante, por lo que puede influir positivamente en la corrección de los ejercicios posteriores. Lo mismo pasa con el polo negativo.

2. Sesgo de disponibilidad

Tal y como se ha comentado previamente, al igual que el sesgo de representatividad, se basa en la teoría que lleva el mismo nombre. Tversky y Kahneman (1974) la conceptualizaron como la tendencia a juzgar la probabilidad o frecuencia de un evento basándose en la facilidad con que ejemplos relevantes vienen a la mente. Es un sesgo que afecta principalmente el juicio que se hace en torno a un suceso específico. En el ámbito educativo, este sesgo se puede presentar en momentos, como cuando un docente reprende a un estudiante que ha mostrado una actitud particularmente molesta. El educador puede que recuerde más intensamente este episodio; por lo tanto, en la evaluación de la actitud, podría juzgar que el estudiante ha estado molestando en múltiples ocasiones.

3. Sesgo de anclaje

De nuevo, se explica a través de una teoría de Tversky y Kahneman (1974), en concreto la del anclaje y ajuste. Furnham y Boo (2011) resaltan la definición de los primeros autores afirmando que “el efecto anclaje es la influencia desproporcionada sobre las personas que toman decisiones para que emitan juicios que son sesgados hacia un valor presentado

inicialmente”¹ (párr. 4). Es decir, consiste en basarse demasiado en la primera información que se recibe, el "ancla", y tomar decisiones posteriores influenciadas por esta, incluso si esta es irrelevante o incorrecta. Esto significa que la persona procesa la información de modo que da una importancia excesiva a la primera información, solamente por el hecho de ser la primera.

Se asegura en el mismo estudio que esta teoría es muy robusta, dado que ha sido probada durante más de 40 años de investigación, por lo que ha sido verificada en múltiples contextos. En este sentido, se afirma que “lo más impresionante es la cantidad de estudios que han demostrado la solidez de los efectos de anclaje con juicios muy diferentes, por ejemplo, la estimación del tiempo”² (Furnham y Boo, 2011, párr. 6).

En el contexto educativo, esto puede implicar que un docente evalúe el desempeño de un estudiante en función de su primer trabajo, sin tener en cuenta sus progresos posteriores, lo que puede dar lugar a evaluaciones injustas.

4. Efecto halo/cuerno:

El efecto halo (o, en su versión negativa, el efecto cuerno), es un fenómeno que se circunscribe dentro de la teoría de la racionalidad limitada. El nombre fue acuñado por Edward Thorndike (1920), quién demostró, a través de una investigación en el ámbito militar, que la impresión general de una persona puede influir en la percepción de características específicas, aun cuando estas últimas no guardan relación directa con la primera. En concreto, el autor hizo puntuar cinco cualidades de forma independiente, tales como la apariencia física, la inteligencia, el liderazgo, las cualidades personales y el valor al servicio. No obstante, los resultados de la investigación mostraron que éstas guardaban relación estrecha con la impresión general, atenuando deficiencias o potenciando cualidades.

Solomon Asch (1946) contribuyó en la ampliación de este concepto y demostró que un atributo como la calidez o la frialdad de una persona influye en la forma en que las personas interpretan otros rasgos, incluso aquellos que pasan desapercibidos y no se pueden observar. Esta tendencia resalta cómo los primeros juicios pueden generar una percepción global que domina la evaluación posterior.

En esta dirección, Nauts et al. (2014), comentando las conclusiones del trabajo de Asch (1946), afirman que el autor:

¹ Nota. Traducción del autor. A lo largo de todo el trabajo, las traducciones son realizadas por el mismo autor. *The anchoring effect is the disproportionate influence on decision makers to make judgments that are biased toward an initially presented value.*

² *What is most impressive is the number of studies that have demonstrated the robustness of the anchoring effects with very different judgements, for instance, time estimation.*

concluyó que los perceptores se forman impresiones coherentes y unitarias de los demás. Para tales impresiones unitarias, los perceptores atribuyen diferentes significados y pesos a los rasgos, asignando roles centrales a algunos rasgos (estos determinan el significado/función de los otros rasgos) y roles periféricos a otros (su significado/función es determinado por los rasgos centrales). Los rasgos interactúan dinámicamente para dar forma a la interpretación de cada uno³. (p. 154)

Por ejemplo, en el caso de Thorndike (1920), los rasgos centrales serían la impresión general, y los rasgos periféricos serían la apariencia física, la inteligencia, el liderazgo, las cualidades personales y el valor al servicio. Otros autores como Nisbett y Wilson (1977) también relacionaron experimentalmente cómo las percepciones globales de algunas cualidades afectaban a la percepción de otras habilidades específicas, como por ejemplo la amabilidad.

Una aplicación práctica en el contexto de la evaluación de contenido podría ser que un estudiante que destaca por su popularidad fuera evaluado más positivamente en su desempeño académico que otro con un perfil más discreto.

5. Sesgo de comparación

En el sesgo de comparación -o sesgo de contraste- el agente que toma la decisión “percibe la información en contraste con lo que la precedió”⁴ (Bhargava y Fisman, 2014, p. 444). Es decir, en el campo de la evaluación de contenido, es la tendencia a evaluar un desempeño en relación con otro previo, en lugar de basarse en criterios objetivos. Son muy frecuentes en correcciones secuenciales. Por ejemplo, un docente podría puntuar un examen más severamente si el trabajo anterior revisado fue excepcionalmente bueno.

Una vez definidos los primeros cinco tipos de sesgos cognitivos, se procede a definir otros tres sesgos que contienen un componente emocional o social en su génesis.

6. Sesgo de confirmación

Se puede caracterizar el sesgo de confirmación, también denominado sesgo a mi favor (*myside bias*), como “la tendencia a considerar y ponderar de manera positiva aquellos argumentos consistentes con las propias opiniones, al mismo tiempo que se ignoran y se evalúan de manera negativa aquellos argumentos que contradicen estas opiniones y creencias previas” (Perkins, 2019, citado por Massolo y Traversi, 2021, p. 62).

En el contexto educativo, un docente puede, por ejemplo, juzgar más favorablemente a un estudiante que el curso pasado tuvo un rendimiento académico elevado. Por el contrario, si

³ [Asch] concluded that perceivers form coherent, unitary impressions of others. For such unitary impressions, perceivers attribute different meanings and weights to traits, assigning central roles to some traits (these determine the meaning/function of other traits) and peripheral roles to others (their meaning/function is determined by central traits). Traits interact dynamically in shaping each other's interpretation.

⁴ [a decision maker] perceives information in contrast to what preceded it.

un estudiante ha demostrado un bajo desempeño, el docente puede ser más propenso a notar los errores y subestimar sus logros.

7. Sesgo de expectativa

Keoma (2021) define el sesgo de expectativa del siguiente modo:

Es la tendencia de las personas a percibir en función de sus expectativas, filtrando la información que supone más importante, circunstancia que conduce a una interpretación distorsionada o inexacta de una situación específica. Este sesgo está relacionado directamente con las emociones, los deseos, las creencias o ganas de que algo suceda, razón por la cual, la percepción juega un papel fundamental. (p. 17)

Y, dado que el trabajo se desarrolla en el ámbito educativo, es de especial relevancia y merecen una atención particular las aportaciones de Rosenthal y Jacobson (1968), que vincularon el sesgo de expectativa con el efecto Pigmalión. Este último fenómeno se refiere a cómo las creencias o expectativas de las personas afectan al desempeño de otras, con lo que, aplicado en el campo de la educación, se traduce en cómo las creencias de los profesores pueden generar sesgos positivos (o negativos) en los estudiantes. Otro ejemplo podría ser que un profesor que espera que un estudiante tenga éxito puede, sin darse cuenta, ofrecerle más apoyo (moldea su comportamiento), lo que aumenta la probabilidad de que el estudiante cumpla esa expectativa.

8. Sesgo de afinidad

Este sesgo se produce cuando un evaluador favorece a las personas con las que comparte características similares con él, como antecedentes culturales o intereses. En educación, puede manifestarse cuando los profesores tienden a favorecer a estudiantes que se parecen a ellos o comparten sus valores.

Entre todos estos fenómenos vistos, se va a agrupar en un solo tipo aquellos sesgos que presenten características muy similares. Por un lado, se puede cuestionar si el sesgo de anclaje y el sesgo de comparación se podrían analizar conjuntamente, pues comparten una característica general común, que es el hecho de comparar el objeto de evaluación con otro ejemplo de características similares. Además, todos estos sesgos afectan al juicio del evaluador, es decir, afectan al mismo elemento de cognición. Sin embargo, se considera que el sesgo de anclaje tiene suficientemente peso en la bibliografía como para considerarlo como un tipo de sesgo independiente.

La simplificación que se propone es la de los sesgos de confirmación y de expectativa. Ambos sesgos comparten un mecanismo esencial, que consiste en proyectar sobre una persona unas creencias o expectativas y buscar su confirmación en ellas. Por tal razón, se

decide simplificar los dos tipos de sesgos en uno, que de ahora en adelante se tratará como sesgo de confirmación.

2.1.4 *El fenómeno de estudio*

Una vez presentado el concepto de sesgo cognitivo, su definición, el recorrido histórico de las principales teorías y los distintos tipos que existen, se procede a conceptualizar el fenómeno que se quiere estudiar. En concreto, se pretende estudiar el fenómeno del sesgo cognitivo en relación con la evaluación de contenido en Educación Secundaria, para determinar si es más favorable en estudiantes con mejor rendimiento académico previo. Se analizan los distintos tipos de sesgo cognitivo y la potencial relación que existe con el sesgo cognitivo en estudiantes con alto rendimiento académico.

De entre los 7 tipos definidos en el apartado “2.1.3 Tipos de sesgos cognitivos”, hay algunos que pueden presentar más incidencia y que, a la vez, pueden ser detectados con mayor facilidad, por lo que van a ser objeto de estudio:

- *Sesgo de anclaje*. Puede ser que el profesor recuerde en exceso la primera información académica recibida (ya sea muy buena o muy mala).
- *Efecto halo/cuerno*. El profesor puede verse influenciado por una característica sobresaliente del estudiante, en el caso de estudio, sería la brillantez académica (o la falta de esta).
- *Sesgo de comparación*. Es común que los estudiantes con mejores notas sean objeto de comparaciones con otros estudiantes con peores calificaciones. En este sentido, se puede sobrevalorar su rendimiento.
- *Sesgo de confirmación*. Se puede dar el caso que los profesores proyecten las expectativas de unos resultados académicos excelentes durante los primeros exámenes en pruebas posteriores. Por el contrario, se puede proyectar una expectativa negativa o desesperante en estudiantes con malas calificaciones académicas.
- *Sesgo de afinidad*. Se puede dar el caso que haya estudiantes con buenos resultados académicos que tengan más afinidad (aunque sea intelectual) con los profesores.

2.2 Estado de la cuestión

Tal y como se ha introducido en el primer capítulo, el fenómeno del sesgo cognitivo es una disciplina ampliamente estudiada en campos como la psicología, la sociología o la economía, pero la investigación llevada a cabo en el ámbito educativo es reducida (Sola et al., 2020).

Hay distintos factores que hacen compleja la búsqueda de información sobre la cuestión planteada. Un primer escollo es que la búsqueda debe estar acotada en el ámbito cognitivo, dado que existen otros tipos de sesgo como, por ejemplo, los sesgos culturales; los sesgos que se dan en el material académico usado por los docentes, como los libros de texto; los sesgos raciales (Clark y Zygmunt, 2014); los sesgos de género... Otra cuestión a considerar es que la investigación debe estar centrada en la evaluación de contenido y no solamente en la relación de enseñanza-aprendizaje entre el profesor y el alumno. Por ejemplo, Allen y Coole (2012), investigan acerca del sesgo de confirmación en la enseñanza y se centran en el modo de evaluar (mediante un pre-test y un post-test), por lo que sí que inciden en la evaluación del contenido. No obstante, se centran en alumnos de posgrado, que es precisamente la tercera dificultad de la investigación; debe estar dirigida a alumnos de secundaria. Finalmente, hay un último grupo de investigaciones que se centran fundamentalmente en ofrecer estrategias para mitigar el sesgo cognitivo. Sola et al. (2020) sostienen que en los últimos años surge “una nueva línea de estudio que se focaliza en la educación anti-sesgos” (p. 500). El mismo autor, citando los trabajos de Gaias (2015) y Sparks y Edwards (2012), destaca al respecto que:

ofrecen un marco teórico y práctico para la comprensión de cómo incorporar el estudio de los sesgos cognitivos, su detección y reducción en la práctica y la política de clase, así como la forma de enseñar a los estudiantes para contrarrestar activamente los prejuicios y la discriminación. (Sola et al., 2020, p. 500)

En esta línea, el autor ofrece dos ejemplos de artículos científicos que tratan el tema de la mitigación del sesgo cognitivo mediante la gamificación, una línea que también está en auge en los últimos años:

Por ejemplo, el juego “*Mitigating Analyst Cognitive Bias by Eliminating Task Heuristics (Macbeth)*” que es usado para entrenar a los analistas de la CIA (Dunbar y cols., 2013; 2014) o “*Missing: The pursuit of Terry Hughes (missing)*”, el cual, es un entrenamiento para adultos para identificar y mitigar los sesgos cognitivos (Symborski y cols., 2014). Esta clase de juegos trata de recrear las situaciones que llevan al uso de sesgos cognitivos. Tienen como objetivo hacer consciente al participante de sus sesgos y les enseña a reducirlos. (Sola et al., 2020, p. 501)

Con todo, se constatan dos fenómenos. El primero, que existen algunas investigaciones sobre sesgos cognitivos en educación y, de modo más o menos tangencial, abordan la temática de la evaluación de contenido y la reducción del sesgo cognitivo, aportando conocimiento y valor en la disciplina de estudio. Y el segundo, que no se ha podido comprobar la existencia de ningún estudio científico que trate directamente sobre el sesgo cognitivo en la evaluación de contenido en Educación Secundaria y, mucho menos, en el ámbito de la educación en Cataluña.

En consecuencia, se usará la aportación de distintos autores para acabar de confeccionar cuerpo del trabajo e ir relacionando la evaluación de contenido con la potencial vulnerabilidad al sesgo cognitivo, así como las distintas estrategias para la mitigación del mismo. Esta tarea se desarrolla, respectivamente, en los apartados “2.3 Vulnerabilidad al sesgo cognitivo en la evaluación de contenido” y “2.4 Mitigación de los sesgos cognitivos”.

2.3 Vulnerabilidad al sesgo cognitivo en la evaluación de contenido

Todos los sesgos descritos anteriormente inciden en la figura del evaluador como sujeto que padece el sesgo cognitivo, afectado por los errores de juicio, la distorsión en la percepción, la falta de memoria equilibrada, la traición de los sentimientos... No obstante, cabe la posibilidad que el fenómeno de estudio no solo se produzca *en o a causa del* sujeto que realiza la evaluación, sino que el contenido mismo a ser evaluado y el tipo de evaluación escogida pueden influenciar en dicho proceso.

Es importante hacer notar que estos dos elementos presentan una característica que los diferencia. Por un lado, el contenido a ser evaluado es extrínseco al proceso mismo de evaluación, es decir, viene dado desde fuera, por lo que no puede ser modificado ampliamente. Es decir, el sujeto que evalúa no puede hacer incidencia directa en este factor para reducir el sesgo cognitivo. Por otra parte, en cambio, sí que puede influir más en el método de evaluación escogido, de forma que, en sí mismo, presente una mínima vulnerabilidad al sesgo. Por ello, a continuación se analiza la influencia del tipo de evaluación en el sesgo cognitivo.

Para comprender los tipos de evaluación, primero es preciso recordar brevemente qué es evaluar. Para Stufflebeam y Shinkfiel (1995), autores de las principales teorías evaluativas, “la evaluación es el enjuiciamiento sistemático de la valía o mérito de un objeto” (p. 19). Los mismos autores, de forma más pormenorizada, detallan que:

la evaluación es el proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y el mérito de las metas, la planificación, la realización y el impacto de un objeto determinado con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados. (p. 92)

Es decir, la tarea evaluativa implica un proceso fundamental de enjuiciamiento de un valor, aunque su actividad no se circunscribe solamente a esta función. Por ejemplo, Navaridas (2013) sostiene que la evaluación no solo se centra en calificar el desempeño del alumnado, sino también en diagnosticar necesidades, retroalimentar procesos y garantizar la calidad del sistema educativo.

Dado que la evaluación persigue múltiples fines, es de esperar que se alcancen de múltiples formas. Por ello, existen diversos tipos de evaluación y diversas clasificaciones de los mismos. Entre ellas, Díaz (2002) destaca que:

desde hace ya varias décadas se ha incorporado al discurso docente [...] la clasificación que distingue a los tipos de evaluación por el momento en que se introducen en un determinado episodio, proceso o ciclo educativo. Estas tres clases de evaluación son llamadas: diagnóstica, formativa y sumativa. (p. 396)

Esta clasificación a menudo recibe el nombre de clasificación según la finalidad, aunque es cierto que la evaluación diagnóstica suele estar al principio del proceso educativo, la evaluación formativa suele acompañar dicho proceso y la sumativa tiende a aplicarse al final. Es por este motivo que Díaz (2002) lo distingue según el momento en que se introducen. Ambas clasificaciones pueden ser válidas a la vez, dado que no son excluyentes. Así lo entiende Rosales (1990) cuando explica las diferencias entre la evaluación diagnóstica, formativa y sumativa, resaltando el momento de aplicación y la finalidad que persiguen.

Aunque no son tan predominantes, existen otros autores que clasifican los tipos de evaluación según otros criterios. Por ejemplo, Del Valle (2013) enfatiza el enfoque metodológico; Heredia (2009) incide en el estándar de comparación; Ravela (2013) vuelve a clasificar la evaluación según la finalidad, pero presenta otra distinción según el agente evaluador; Mora-Vicarioli (2019) afirma que existen “múltiples clasificaciones según la forma en que se lleve a cabo, el momento en que ocurra, el sujeto que la lleve a cabo, entre otros” (p.71).

Por todo ello y dado que la clasificación de los tipos de evaluación puede ser muy dispar según los autores consultados, se elabora un resumen de los criterios de clasificación más relevantes. En él, se incluyen: el momento en que se realiza la evaluación, la finalidad que tiene, el enfoque metodológico adoptado, el estándar de comparación utilizado, el agente evaluador que la lleva a cabo, la modalidad escogida y la extensión de la misma. A continuación, se explica tal clasificación que se resume finalmente en la **Tabla 1**. Además, se relaciona cada tipo de evaluación con la vulnerabilidad frente al sesgo cognitivo y se califica de modo gradual en caso de que exista. De esta forma, se pretende dar respuesta a si los tipos de evaluación juegan un papel destacado en la aparición del sesgo cognitivo y, en caso afirmativo, se identifica cuáles tienen una mayor relevancia.

Empezando el análisis, aparece la primera clasificación según el momento en que se produce la evaluación. Dentro de este criterio existen cuatro tipos: la evaluación inicial, la continua, la final y la diferida. La evaluación inicial, tal y como el nombre indica, se aplica al comienzo del proceso de enseñanza-aprendizaje con el propósito de diagnosticar los conocimientos previos, habilidades y necesidades del estudiante. La evaluación continua se desarrolla a lo largo de todo el proceso educativo y, tal como indica Mora-Vicarioli (2019), “busca que el estudiante vaya recibiendo realimentación con las diferentes actividades que realiza” (p. 73). La evaluación final se lleva a cabo al término de un curso, unidad o ciclo educativo con el fin de medir los aprendizajes logrados en función de los objetivos previamente establecidos. Suelen utilizarse exámenes, trabajos o proyectos para analizar el

rendimiento y en Educación Secundaria, suele expresarse en forma de calificaciones numéricas. Por último, la evaluación diferida se realiza después de un período de tiempo considerable tras la finalización del proceso de enseñanza. Su objetivo es evaluar la retención, la profundidad y solidez del conocimiento adquirido.

De estas descripciones, se infiere que la evaluación final es el tipo de evaluación más vulnerable al sesgo cognitivo por diversas razones. En primer lugar, por el efecto de disponibilidad, dado que los evaluadores pueden recordar con más peso el desempeño reciente del estudiante, sin considerar su progreso completo. En segundo lugar, por el sesgo de expectativa, ya que si un docente espera que un estudiante tenga un rendimiento alto o bajo, puede interpretar su desempeño final de manera sesgada. Finalmente, por el efecto halo, es decir, que una buena impresión previa del estudiante pueda influir en la calificación final sin considerar su rendimiento real.

En contraste, los otros tipos de evaluación -considerando el momento en que se hace la evaluación- son menos vulnerables al sesgo cognitivo. La evaluación continua o procesual, al ser progresiva, permite corregir sesgos al integrar múltiples mediciones y retroalimentaciones. La evaluación inicial tiene menos sesgos acumulados, aunque puede verse afectada por el sesgo de confirmación o de expectativa del evaluador. Y la evaluación diferida, al realizarse con distancia temporal, puede reducir el impacto de juicios emocionales inmediatos, aunque sigue siendo vulnerable al sesgo de disponibilidad.

En segundo lugar, existe la evaluación según la finalidad, que distingue tres tipos: la diagnóstica, la formativa y la sumativa. Rosales (1990) analiza la finalidad de la evaluación diagnóstica afirmando que:

tiene como misión específica determinar las características de la situación inicial para la puesta en marcha de un determinado proceso didáctico, y servir de base por lo tanto, para decisiones sobre la programación o diseño del mismo. También se atribuye a la evaluación diagnóstica la profundización en el conocimiento de las causas de determinados problemas a lo largo de la enseñanza, en cuyo caso serviría de base para decisiones relativas a su recuperación. (p. 34)

Es iluminador ver como la evaluación diagnóstica no solo ofrece información sobre la situación inicial, sino que su finalidad última es de carácter pragmático pues pretende servir de herramienta para orientar las decisiones sobre las estrategias que se van a aplicar a futuro, a la vez que permite detectar las carencias en los procesos de enseñanza. A la vez, sirve como retroalimentación inicial para estudiantes y padres, es decir, proporciona una oportunidad para que ambos pueden incidir en las áreas que necesitan refuerzo y mantener aquellas donde se exhibe dominio.

La evaluación formativa, al igual que la diagnóstica, no pretende centrarse en los resultados, sino que en cuanto al proceso didáctico se refiere, "sirve para determinar la naturaleza de desarrollo del mismo y constituye el punto de partida para decisiones de perfeccionamiento"

(Rosales, 1990, p. 34). Es decir, de nuevo, tiene carácter informativo y de retroalimentación, que sirve como herramienta de discernimiento en la mejora continua y permite a los estudiantes tomar conciencia de sus fortalezas y debilidades a medida que van aprendiendo. La evaluación sumativa es la más conocida y practicada “tiene lugar al final de un determinado proceso didáctico, constata los resultados del mismo y sirve de base para adoptar decisiones de certificación, de promoción o repetición, de selección” (Rosales, 1990, p. 34). Como en los otros tres casos, sirve como herramienta para tomar decisiones, aunque esta vez, bajo una perspectiva claramente calificadora y numérica.

Tal y como se ha hecho notar, los tres tipos de evaluación (diagnóstica, formativa y sumativa), de forma general, están asociados de forma paralela a los momentos iniciales, medios y finales del proceso evaluativo, por lo que se deduce que adolecen de los mismos tipos de sesgos cognitivos. Esto significa que la evaluación sumativa es más vulnerable a los sesgos cognitivos que la evaluación diagnóstica o la formativa. El hecho de que la evaluación sumativa esté enfocada en medir los resultados de forma numérica, le hace especialmente vulnerable.

En tercer lugar, si la distinción se realiza según el enfoque metodológico, se pueden distinguir dos grandes tipos: la evaluación cuantitativa y la cualitativa, y la mezcla de las dos, que se define como evaluación cuali-cuantitativa. Del Valle (2013) resume con gran precisión los dos primeros enfoques y la necesidad de la mutua interconexión entre ellos. Puntualiza que existen:

dos tendencias: una cuantitativa y la otra cualitativa. La primera centra su atención exclusivamente en la cantidad de productos y resultados efectuados; la segunda describe e interpreta los procesos que tienen lugar en todo el entorno educativo. Creemos que la evaluación educativa comprende ambos aspectos: evaluar es medir, y evaluar es predecir, aunque más bien a corto y mediano plazo. Por lo que optamos por centrar la evaluación no únicamente en la medición de los resultados finales, sino en la recogida de información sobre las variables que intervienen en el proceso de enseñanza y aprendizaje, como subraya la tendencia cualitativa, pero cuantificadas en una medida, como reclama la tendencia cuantitativa. Todo ello encaminado a una toma de decisiones que regirá el investigador según su paradigma. (p. 109)

Es decir, la primera se centra más en medir los resultados a través de pruebas estandarizadas y técnicas numéricas, por lo que permite ser más objetiva. En contraste, la segunda se detiene en el proceso mismo de aprendizaje para recoger información útil, por lo que, al ser más descriptiva, puede resultar más subjetiva. Dada esta estrecha relación entre evaluación cuantitativa y cualitativa y el grado de objetividad y subjetividad, se deduce que esta última es más vulnerable a los sesgos cognitivos.

En cuarto lugar, si se analiza la evaluación de contenido según el estándar de comparación, existe la evaluación normativa, la criterial y la personalizada. Según Popham (1983, citado por Heredia, 2009), la evaluación normativa

tiene como objetivo ver la posición relativa de cada alumno respecto a los demás y cumple su función cuando con ella se sabe quién es mejor (o peor) que quién. Así, trata de identificar el talento del alumno, de manera que se puedan establecer comparaciones y luego hacer una selección entre los alumnos. (p. 42)

Por tanto, la evaluación normativa tiene una función discriminatoria -no en sentido peyorativo- que permite ordenar la adquisición de conocimiento de los alumnos. A este planteamiento se contraponen la evaluación criterial, cuya característica fundamental es la de “apreciar el logro de los objetivos por parte de cada alumno sin compararlo con el de sus compañeros” (Gómez Arbo, 1990, citado por Heredia, 2009, p. 43). Es, frecuentemente, el tipo de evaluación más utilizada en el aula. Se podría considerar un tercer tipo dentro de este criterio, que es la evaluación personalizada. Se adapta a las características individuales de cada estudiante, considerando su ritmo de aprendizaje, necesidades y capacidades. Es frecuente en educación inclusiva y atención a la diversidad.

Dada la naturaleza de la educación normativa, no se puede afirmar que presente el sesgo de comparación, pues la comparación con otros estudiantes se realiza de forma sistemática e intencional, por lo quedaría excluida propiamente de la definición de sesgo cognitivo. No obstante, puede resultar vulnerable al sesgo de anclaje o el efecto halo, al comparar el resto de las calificaciones con el primer estudiante elegido o con un estudiante que sobresalga por sus características personales o académicas. Al tener que ajustar la calificación según el desempeño del resto de estudiantes, se puede provocar un sesgo de afinidad, según el estudiante en cuestión comparta más o menos características con el evaluador. Así lo defiende Sadler (2009):

La calificación basada en normas suele considerarse injusta porque los estudiantes no tienen control sobre la pertenencia al grupo de referencia, pero sus calificaciones dependen de la clasificación en relación con otros. Al externalizar el marco de referencia, la calificación basada en criterios es también un contrapunto para evaluar los trabajos de los estudiantes en contraposición a otros marcos de referencia como los gustos o preferencias no explicitadas del docente, o el nivel de desempeño previo de cada estudiante.⁵ (p. 175)

Por tanto, se puede afirmar que el uso de rúbricas, asociado al modelo criterial, disminuye el impacto de los sesgos cognitivos. Castejón et al. (1998), citados por Heredia (2009), aseguran que las pruebas criteriosales “pueden asegurar en mayor medida una evaluación objetiva del conocimiento de cada alumno en distintas fases, etapas o momentos del ciclo

⁵ *Norm-referenced grading is often considered unfair because students do not have control over the membership of the reference group, yet their grades depend on ranking relative to others. By externalising the reference framework, criteria-based grading is also a counter to evaluating student works against such other frameworks as the teacher's unarticulated tastes or preferences, or each student's previous level of performance.*

formativo” (p. 41). Por último, la evaluación personalizada puede caer fácilmente en sesgos de confirmación, de expectativa y de afinidad y el efecto halo/cuerno aunque, en principio, debería evitar el sesgo de comparación.

En quinto lugar, si se distingue el tipo de evaluación por el agente evaluador, existe la evaluación interna y la externa. Según Ravela (2013) “una evaluación es externa cuando es realizada por un evaluador distinto del sujeto o institución evaluada y es interna cuando es realizada por el propio sujeto” (p. 3). Por otro lado, existe la autoevaluación, que es llevada a cabo por los mismos estudiantes, la heteroevaluación, realizada por docentes u otros agentes educativos y la coevaluación, donde los mismos compañeros efectúan la evaluación.

En cuanto a la evaluación externa, puede ser mucho menos vulnerable que la evaluación interna en cuanto a los sesgos cognitivos, ya que reduce todos aquellos que están vinculados con el conocimiento de los estudiantes, como el efecto halo y los sesgos de confirmación, de expectativa y de afinidad.

De los otros tres tipos de evaluación, la coevaluación puede estar sujeta a sesgos como el sesgo de afinidad, el sesgo de confirmación, el sesgo de comparación y el efecto halo. La autoevaluación también puede estar sesgada, dado que la imagen que uno tiene de sí mismo no siempre se corresponde con la realidad. En cuanto a la heteroevaluación, que es el modo de evaluación más usado, puede presentar habitualmente todos los sesgos mencionados en el trabajo.

En sexto lugar, se distingue la evaluación según la modalidad, que diferencia entre el modo presencial, que se lleva a cabo en un entorno físico donde los estudiantes y docentes interactúan directamente, y la evaluación a distancia, realizada mediante plataformas digitales o medios remotos. Aunque la evaluación a distancia a menudo puede resultar más objetiva, no es por la naturaleza misma de la evaluación a distancia, sino por las limitaciones tecnológicas en las que se ven envueltos los evaluadores, los cuales tienden a utilizar una evaluación más cuantitativa o de respuesta única. Por tanto, a priori, no se puede afirmar que una u otra (presencial o a distancia) tenga mayor o menor vulnerabilidad al sesgo cognitivo, por su misma naturaleza.

Por último, se analiza el último criterio de tipo de evaluación, que es según la extensión. En ella se distinguen la evaluación parcial y la global, donde se evalúa solo una parte o la totalidad del contenido. En sí, esta distinción no afecta en la facilidad en que se puedan producir los sesgos cognitivos, aunque es coherente afirmar que, a mayor contenido a corregir, más aumenta la probabilidad, por estadística, de que se produzcan los sesgos cognitivos por la fatiga y el cansancio del evaluador.

En la **Tabla 1** se presenta un resumen de los criterios de clasificación más relevantes, analizados previamente, así como su vulnerabilidad frente al sesgo cognitivo y el modo de afectación en forma de escala gradual.

Tabla 1

Clasificación de los tipos de evaluación.

Criterio	Tipos	Vulnerabilidad al sesgo
Según el momento	Inicial	Baja-Media
	Procesual (continua)	Baja
	Final	Alta
	Diferida	Baja
Según la finalidad	Diagnóstica	Baja-Media
	Formativa	Baja
	Sumativa	Alta
Según el enfoque metodológico	Cuantitativa	Baja
	Cualitativa	Alta
	Cuali-cuantitativa	Media
Según el estándar de comparación	Normativa	Alta
	Criterial	Baja
	Personalizada	Media
Según el agente evaluador	Interna	Alta
	Externa	Baja
	Autoevaluación	Media
	Heteroevaluación	Alta
	Coevaluación	Media
Según la modalidad	Presencial	Indiferente
	A distancia	Indiferente
	Mixta	Indiferente
Según la extensión	Parcial	Baja
	Global	Media

Nota. Tabla de elaboración propia.

2.4 Mitigación de los sesgos cognitivos

2.4.1 Causas que propician los sesgos cognitivos en contextos educativos

Como se deduce del análisis presentado en el inicio del marco teórico, las causas o factores que propician los sesgos cognitivos son variados. Se pueden agrupar en distintos factores:

- La propia naturaleza heurística del pensamiento humano.
- Las emociones en el juicio profesional.
- El contexto cultural y social.

Además, hay dos factores que también pueden ser el origen de los sesgos cognitivos, y que merece la pena destacar.

Por un lado, las mismas instituciones pueden propiciar sesgos cognitivos cuando no ofrecen formación suficiente para detectarlos y erradicarlos, en la medida de lo posible, de las prácticas del profesorado. Si existe falta de retroalimentación de la actividad docente o ausencia de trabajo colaborativo, la percepción del propio trabajo queda reducida a la visión que cada docente pueda tener de la misma, cosa que puede incrementar este fenómeno. Por último, si existe una exigencia por cumplir unos determinados estándares, puede llevar a los docentes a adoptar mecanismos de corrección más rápidos y subjetivos, en lugar de favorecer la objetivación en la evaluación de contenido.

Por otro, hay que considerar el agotamiento profesional, conocido como *burnout*. Es “una respuesta al estrés crónico en el trabajo (a largo plazo y acumulativo), con consecuencias negativas a nivel individual y organizacional” (Martínez Pérez, 2010, p. 3). De modo habitual se asocia con un estado de agotamiento físico, mental y emocional propiciado por el estrés laboral. Además, este fenómeno “tiene gran relevancia en el ámbito educativo, especialmente en el desempeño docente” (Briceño, 2024, p. 4). El agotamiento crónico, en consecuencia, puede modificar el estado de ánimo, alterando las funciones cognitivas como la percepción, el juicio o la memoria, de tal forma que se favorezca la aparición de los sesgos cognitivos.

2.4.2 Estrategias transversales para la mitigación de los sesgos cognitivos

Para mitigar la influencia de los sesgos cognitivos, es necesario implementar distintas estrategias que minimicen sus causas. Se entiende por estrategias transversales todas aquellas acciones que se pueden aplicar en la reducción de cualquier tipo de sesgo cognitivo. De algún modo, van más allá del tipo de sesgo en concreto, pues son como el *background* general. Entre los más comunes, se pueden destacar la sensibilización, formación y reflexión, el uso de la tecnología y la mejora de las condiciones laborales.

Un primer paso esencial es la sensibilización y reflexión sobre la existencia y los efectos de los sesgos cognitivos. Puesto que en la naturaleza misma de los sesgos cognitivos está el hecho que son inconscientes, tomar consciencia de ellos ayuda en gran medida a la reducción de los mismos. Para ello, es deseable la actuación en dos niveles. El primero, a nivel colectivo, donde es necesaria la formación en este ámbito. Y el segundo, a nivel individual, ofreciendo ayuda a cada docente en particular para que identifique y explicita los momentos en los que puede haber padecido el influjo de este fenómeno.

En segundo lugar, el uso de la tecnología también ofrece alternativas para mitigar los sesgos cognitivos. Permiten a los docentes preparar material con múltiples fuentes de información, de tal forma que amplían la visión del profesor en el modo de evaluar el contenido. También son útiles en esta tarea puesto que permiten la posibilidad de correcciones instantáneas, basadas en la objetividad que la tecnología requiere. Además, ofrecen retroalimentación

personalizada a los estudiantes, por lo que puede reducir sesgos tales como el sesgo de afinidad o el efecto halo.

Finalmente, es importante implementar condiciones laborales que reduzcan la aparición de sesgos cognitivos. Ramos (1999), citado por Martínez Pérez (2010), considera importante la intervención en el ámbito individual, social y organizacional, “en referencia a la eliminación o disminución de estresores en el trabajo” (p. 70). Solo así los docentes pueden tomar decisiones equilibradas y justas.

En conclusión, mitigar los sesgos cognitivos en el ámbito educativo requiere un enfoque multidimensional que combine sensibilización, reflexión y formación profesional, la integración de tecnología y la mejora de las condiciones laborales. Estas estrategias no solo mejoran la práctica docente a nivel personal y social, sino que logran un sistema de evaluación más justo y equilibrado.

2.4.3 Estrategias específicas para la mitigación de los sesgos cognitivos

Aunque las estrategias transversales para la mitigación del sesgo cognitivo sean fundamentales para la consecución de este objetivo, es preciso implementar estrategias específicas, es decir, acciones que se concentren en un tipo de sesgo en particular. Bajo esta premisa, se analizan los distintos tipos de sesgo cognitivo, ofreciendo estrategias de mitigación para cada uno en concreto.

- *Sesgo de representatividad.* Ocurre cuando los evaluadores juzgan de manera errónea la probabilidad de un suceso. Se puede mitigar mediante el uso de una rúbrica. Brookhart (2013) la define como “un conjunto coherente de criterios para el trabajo de los estudiantes que incluye descripciones de niveles de desempeño en el criterio”⁶ (p. 4). El mismo autor establece que “el propósito principal de una rúbrica es evaluar el desempeño”⁷ (p. 4).
- *Sesgo de disponibilidad.* Ocurre cuando los evaluadores toman decisiones basadas en la información más fácil de recordar o más reciente. Afecta principalmente al juicio y a la memoria del evaluador, por lo que la principal estrategia para erradicar este sesgo es la evaluación ciega, es decir, a supresión de cualquier característica (nombre, sexo...) que pueda relacionar el estudiante evaluado con el contenido mismo evaluado. De esta forma, se obliga al evaluador a emplear un registro sistemático mediante rúbricas y notas objetivas y a no depender de la memoria. Para limitar el uso de la memoria también puede ser útil espaciar la corrección. Por otra

⁶ a coherent set of criteria for students' work that includes descriptions of levels of performance quality on the criteria.

⁷ the main purpose of a rubrics is to assess performances.

parte, para disminuir el sesgo en el juicio, es conveniente la revisión de decisiones, es decir, reflexionar sobre las calificaciones y compararlas con criterios establecidos.

- *Sesgo de anclaje*. Consiste en basarse demasiado en la primera información que se recibe (el ancla). Afecta principalmente al modo en como el evaluador procesa la información y el orden en el cual este la recibe. De esto se deduce que las estrategias más plausibles para mitigar este sesgo son la adopción de un método aleatorio de corrección para no depender del primer trabajo corregido, el uso de rúbricas para objetivar dicha evaluación y la revisión de todos los trabajos para no estar vinculado de forma desproporcionada con el primer o los primeros trabajos corregidos. En este sentido, también puede resultar útil espaciar las correcciones o emplear una doble corrección para no dar demasiada importancia al primer trabajo corregido.
- *Efecto halo*. Se basa en impresiones generales que distorsionan los juicios posteriores. Se puede mitigar separando las evaluaciones de conducta y de desempeño académico, utilizando rúbricas que definan claramente los criterios específicos para cada aspecto del aprendizaje. Otra forma para no asociar la impresión general de la persona que se corrige con los aspectos puramente académicas es el uso de la evaluación ciega o de la doble corrección.
- *Sesgo de comparación*. Se produce en comparación con otras evaluaciones. Afecta principalmente a la percepción y al juicio del evaluador y se puede mitigar mediante el uso de rúbricas que definan claramente los niveles de desempeño esperado. Otro modo de reducción del sesgo sería espaciar temporalmente las correcciones, dividiéndolas en grupos más pequeños, para no comparar el desempeño del estudiante con el del grupo. Además, el factor de la aleatoriedad puede ayudar a reducir el sesgo, puesto que se evitan comparaciones, por ejemplo, con el anterior de la lista, que puede ser que sea siempre un estudiante brillante.
- *Sesgo de confirmación*. Tiende a beneficiar a los estudiantes que el evaluador considera mejores y a perjudicar a los peores. Afecta principalmente al modo en como el evaluador busca, interpreta o recuerda información de una persona en concreto. También tiene incidencia en la emisión del juicio posterior. Para mitigar este sesgo es preciso desvincular el contenido del sujeto evaluado, por lo que la evaluación ciega o la doble corrección son las principales herramientas para ello. Otro modo de mitigar este sesgo es mediante, una vez más, el uso de rúbricas que definan claramente los niveles de desempeños deseados.
- *Sesgo de expectativa*. Al igual que el sesgo de confirmación, puede beneficiar a los mejores estudiantes y perjudicar a los peores, pero en este caso no según una información previamente disponible, sino según las expectativas del evaluador. Las

técnicas más útiles pueden ser la evaluación ciega, el uso de la doble corrección y el uso de rúbricas objetivas.

- *Sesgo de afinidad*. Depende del vínculo emocional o social del evaluador con el evaluado. Se podría mitigar con la doble corrección, es decir, que otra persona corrigiera el mismo contenido para mitigar las conexiones que pudiera haber entre los primeros. De modo similar, implementar la coevaluación o la evaluación por pares podría ayudar a objetivar la evaluación. Otro método efectivo es la supresión del nombre, es decir, la evaluación ciega, para prevenir este tipo de preferencias. Finalmente, el uso de rúbricas objetivas también es una herramienta útil para separar lo personal de lo académico.

Por tanto, analizando en conjunto todos los sesgos descritos, se observa que existen distintas técnicas que pueden usarse con la finalidad de mitigar más de un sesgo. En la **Tabla 2** se pueden identificar las cinco estrategias más comunes, por frecuencia de aparición.

Tabla 2

Estrategias específicas de mitigación de los sesgos cognitivos más presentes en la evaluación de contenido.

Tipo de sesgo	Rúbricas	Evaluación ciega	Doble corrección	Espaciar la corrección	Aleatoriedad
Sesgo de representatividad	x				
Sesgo de disponibilidad	x	x		x	
Sesgo de anclaje	x		x	x	x
Efecto halo/cuerno	x	x	x		
Sesgo de comparación	x			x	x
Sesgo de confirmación	x	x	x		
Sesgo de expectativa	x	x	x		
Sesgo de afinidad	x	x	x		

Nota. Tabla de elaboración propia.

Por tanto, las cinco estrategias específicas más comunes son el uso de rúbricas, la evaluación ciega, el espaciado de la corrección, el uso de la aleatoriedad y la doble corrección. Todas ellas han sido estudiadas por algún autor, que defiende su utilidad en la reducción del sesgo cognitivo.

Por ejemplo, Sadler (2009) plantea que “dividir los juicios holísticos en partes más manejables se considera una forma de aumentar la apertura de los estudiantes y lograr una mayor objetividad en la calificación”⁸ (p. 159). Además, sostiene que “comparar la calidad del trabajo de un estudiante con un criterio fijo y unos ‘estándares’ es educacionalmente más

⁸ *Breaking down holistic judgments into more manageable parts is seen as a way to increase openness for students and achieve more objectivity in grading.*

defendible que hacer comparaciones de cómo otros estudiantes del curso rinden en las mismas tareas o equivalentes”⁹ (p. 175). Queda claro, pues, que el uso de rúbricas permite la reducción del sesgo de comparación, además de conllevar otros beneficios, como el aumento la transparencia.

En cuanto a la adopción de métodos de evaluación ciega o anónima, Gómez Díaz y Martínez López (1997) sostienen que “la subjetividad de las calificaciones puede disminuir, por ejemplo, haciendo que los ejercicios sean anónimos” (p. 141).

Por otro lado, Goldbach et al. (2021), concluyen que hay evidencias que demuestran la existencia de sesgos cuando se corrigen exámenes de forma secuencial. Es decir, a los primeros estudiantes se les corrige de modo más favorable que a los últimos. De ahí la necesidad de la aleatoriedad en la corrección de contenido, ya que reduce el sesgo de anclaje y el sesgo de comparación, entre otros. También Joy (2025) reconoce que “simplemente aleatorizando los alumnos y haciendo paradas frecuentes durante la corrección puede prevenir a los alumnos recibir la peor parte”¹⁰ (p. 20). Por último, es ampliamente reconocido que la corrección por múltiples evaluadores reduce significativamente distintos tipos de sesgo, como el de anclaje, el de confirmación y el efecto halo.

⁹ *Comparing the quality of a student's work with fixed criteria and 'standards' is educationally more defensible than making comparisons with how other students in the course perform on the same or equivalent tasks.*

¹⁰ *Simply randomizing students on Canvas and taking frequent breaks during grading could prevent students from receiving the sort end of the stick.*

3 Metodología

La investigación se concibe principalmente bajo el paradigma positivista y es aplicada con un método deductivo, es decir, parte de hipótesis con el fin de llegar a conclusiones. También hay una parte inductiva, donde se formulan unas preguntas que posteriormente se evalúan para determinar el grado de correlación con los resultados, esto es, para llegar a la teoría. En el presente estudio se pretende explicar y comparar, pero a la vez comprender e interpretar un fenómeno en base a una muestra. En cuanto al método, se prioriza una metodología cuantitativa, tanto en la comparación documental como en la elaboración del cuestionario.

3.1 Formulación de la pregunta de investigación e hipótesis

En cuanto a la pregunta de investigación, se podría formular de la siguiente manera: ¿los alumnos de Educación Secundaria con mejor nota previa son evaluados de forma más favorable debido a los sesgos cognitivos?

Para dar respuesta a la pregunta, es preciso definir claramente las variables de estudio:

- Variable independiente (VI): los resultados académicos generales. Es una variable numérica que, en el caso de estudio, se determina por la nota final del primer trimestre de la asignatura del contenido avaluado.
- Variable dependiente (VD): el sesgo cognitivo. Es una variable numérica que puede ser escurridiza, pero que en el presente estudio se va a determinar mediante análisis estadístico de las notas de una prueba de contenido, comparando las dos calificaciones realizadas por el mismo profesor antes y después de aplicar técnicas para limitar el sesgo cognitivo. Es decir, no se mide el sesgo cognitivo como tal, sino el efecto en los resultados académicos de la limitación del sesgo cognitivo.

Relacionada con la pregunta de investigación, surge la hipótesis principal del trabajo, que guiará en todo momento la investigación. Sostiene la siguiente proposición:

- H₁: Cuanto mayor sea la nota previa del alumno, más le beneficia el sesgo cognitivo del profesor en su evaluación.
Por ello, se investiga si los resultados académicos, definidos como variable independiente (VI), afectan a la aparición de los sesgos cognitivos, definidos como variable dependiente (VD).

Además, cabe preguntarse si los sesgos cognitivos son de una misma naturaleza o si, por el contrario, existen varios tipos de sesgo cognitivo presentes en la evaluación de contenido en Educación Secundaria. Por ello, después de determinar si existe o no el sesgo cognitivo, se estudiará si existe alguna vinculación con alguno de los sesgos cognitivos presentados,

tales como el sesgo de anclaje, el efecto halo/cuerno, el sesgo de comparación, el sesgo de confirmación y el sesgo de afinidad.

3.2 Objetivos

De la pregunta de investigación y la hipótesis principal plantada, se define el objetivo central del estudio, que es determinar si los alumnos con mejores resultados académicos previos en Educación Secundaria son evaluados de forma más favorable debido a los sesgos cognitivos del evaluador (Obj₁). En caso afirmativo, se pretende identificar qué tipo o tipos de sesgos son los predominantes en la evaluación del alumnado en Educación Secundaria (Obj₂), lo cual sería el segundo objetivo del trabajo.

3.3 Diseño de la investigación

Para investigar el fenómeno del sesgo cognitivo, el evaluador o la evaluadora, que es el profesor o profesora habitual de la asignatura, corrige y puntúa un contenido académico, que debe cumplir con los siguientes requisitos:

- Debe ser una pregunta del examen trimestral del segundo trimestre.
- Debe ser una pregunta abierta, de comparar, relacionar, expresar una opinión, justificar... De este modo, se crea la necesidad de establecer una rúbrica en la segunda evaluación de contenido.
- Todo el alumnado debe contestar la misma pregunta.

Además, la primera corrección debe seguir los siguientes criterios:

- *Método*. La corrección del contenido se debe llevar a cabo con las herramientas habituales de corrección. Si de normal se utilizan rúbricas de evaluación, se deben utilizar, pero si, por el contrario, no se utilizan, se deben obviar.
- *Identificación*. La corrección se debe llevar a cabo con un dato identificador del alumno para que el profesor pueda relacionar el contenido corregido con el alumno que lo ha realizado. Por ejemplo, mediante el nombre y apellidos.
- *Evaluador*. El profesor habitual de la asignatura del contenido avaluado.
- *Tiempo*. La corrección debe ser efectuada en el período posterior a la realización de la prueba, de la forma habitual. Las preguntas pueden ser corregidas de forma secuencial, por preguntas, con espaciado de corrección...
- *Orden*. Corregir los exámenes con el orden habitual de corrección, de forma deliberada o aleatoria.

A continuación, se implementan las estrategias específicas de mitigación de los sesgos cognitivos más presentes en la evaluación de contenido, especificadas en la **Tabla 2**. En concreto, se incide sobre:

- *Método.* La corrección del contenido se debe llevar a cabo mediante el uso de una rúbrica elaborada por el propio profesor o profesora de la asignatura. En ella se determinan los criterios de puntuación de la pregunta.
- *Identificación.* La corrección se debe llevar a cabo mediante la evaluación ciega, es decir, sin ningún dato que permita asociar el contenido corregido con la persona que lo ha realizado. Se utiliza un pseudónimo para tal efecto.
- *Evaluador.* Para minimizar el sesgo cognitivo se implementaría la doble corrección llevada a cabo por un agente externo a la institución educativa. No obstante, no se puede llevar a cabo por la complejidad que envuelve, por tanto, será el mismo profesor o profesora titular de la asignatura quien corrija el contenido. De esta forma, además, no se añaden variables extrañas no controladas.
- *Tiempo.* La corrección debe ser efectuada en un periodo de tranquilidad. Además, se debe fragmentar la corrección del contenido en, como mínimo, tres veces.
- *Orden.* Se debe corregir los exámenes de forma totalmente aleatoria.

Se evalúa de nuevo el contenido aplicando las estrategias anteriores. Todo ello se lleva a cabo para determinar si un mismo contenido académico presenta distintas cualificaciones antes y después de aplicar las estrategias específicas para la mitigación del sesgo cognitivo. A continuación, se relacionan los datos obtenidos de las dos correcciones con cada alumno y su correspondiente nota del primer examen trimestral. Luego se agrupan los alumnos en función de este último criterio. Finalmente se comparan los resultados globales y por grupos. Con los resultados globales, se pretende determinar si existe sesgo cognitivo en general y mediante el análisis por grupos se pretende identificar si este es más beneficioso en estudiantes con mejores notas previas y viceversa.

Paralelamente, se pasa un cuestionario (ver apartado “3.5 Instrumentos y técnicas de recogida de datos”) a los profesores sobre cada alumno del que ha corregido contenido. En él, se hacen unas preguntas que tienen que ver con los tipos de sesgo, con el fin de determinar qué influencia tiene cada uno en la aparición del sesgo cognitivo.

3.4 Técnica de muestreo y muestra del estudio

Teniendo en cuenta que la población de los alumnos de Educación Secundaria en Cataluña es de unos 340.000 estudiantes, con un nivel de confianza del 95%, la muestra necesaria para poder extrapolar los resultados muestrales a toda la población es de 384 estudiantes.

No obstante, el objetivo principal no es extrapolar los resultados a toda la población. Además, la técnica empleada es una técnica no probabilística por accesibilidad en el centro donde se realizan las prácticas, por lo que los recursos son limitados. Al final, se obtiene una muestra total de 100 alumnos, que se clasifica de forma estratificada en tres grupos según el rendimiento académico:

- A. Estudiantes con nota final del primer trimestre igual o superior a 8.
- B. Estudiantes con nota final del primer trimestre comprendida entre 5 y 7,9.
- C. Estudiantes con nota final del primer trimestre inferior a 5.

De este modo, ningún estudiante queda potencialmente fuera del estudio y se mantienen las proporciones naturales de los resultados obtenidos en la calificación final del primer trimestre. La muestra final corregida por cada profesor se muestra en la **Tabla 3**.

Tabla 3

Muestra de profesores y alumnos.

Profesores	Alumnos
Profesor 1	8
Profesor 2	23
Profesor 3	26
Profesor 4	21
Profesor 5	22
Total	100

Nota. Tabla de elaboración propia.

3.5 Instrumentos y técnicas de recogida de datos

Para lograr el objetivo principal del trabajo, la técnica de recogida de datos utilizada es el análisis de documentos, en concreto, la comparación documental. Tal y como se describe en el apartado “3.3 Diseño de la investigación”, se comparan las preguntas de los exámenes trimestrales, tanto antes como después de la aplicación de las técnicas para reducir el sesgo cognitivo. Dado que se pretende extraer un resultado numérico como diferencia entre las dos cualificaciones, se considera una técnica cuantitativa.

Además, para determinar la influencia de los tipos de sesgo, se elabora un instrumento cuantitativo en forma de cuestionario sobre cada alumno. En él, se formulan una serie de preguntas relacionadas con el sesgo de anclaje, el efecto halo, el sesgo de comparación, el sesgo de confirmación y el sesgo de afinidad. En la **Tabla 16** (Anexo I) se muestra la tabla de dimensiones utilizada para generar el cuestionario.

4 Resultados

Una vez el profesorado ha realizado todas las correcciones de contenido y las encuestas de cada alumno, se realiza el análisis de los datos. A continuación, se presenta el análisis de los resultados relativos al primer objetivo, tanto a nivel global de toda la muestra como por tipos de alumno, según el rendimiento académico previo. Posteriormente se interpretan los resultados obtenidos. Más tarde, se analizan los resultados relativos al segundo objetivo, vinculando la aparición del sesgo cognitivo con los distintos tipos presentados.

4.1 Perfil de la muestra

La muestra ha sido de 100 alumnos, estratificados en tres grupos. Tal y como se puede observar en la **Tabla 4**, la muestra de alumnos obtenida no es homogénea respecto a la cantidad de respuestas corregidas por cada profesor, puesto que el Profesor 1 imparte clases en Bachillerato, donde hay menos alumnos. Respecto a los Profesores 2, 3, 4 y 5, sí que se observa una muestra más homogénea en cuanto al número de respuestas corregidas y en cuanto a la distribución de los alumnos dentro de las tres categorías, a excepción del Profesor 4, que no dispone de ningún alumno con calificación final del primer trimestre entre 8 y 10.

Tabla 4

Muestra de los alumnos según rendimiento académico.

Profesores	Alumnos			Total
	Grupo A [0,5)	Grupo B [5,8)	Grupo C [8-10]	
Profesor 1	3	2	3	8
Profesor 2	6	15	2	23
Profesor 3	5	13	8	26
Profesor 4	5	16	0	21
Profesor 5	7	11	4	22
Total	26	57	17	100

Nota. Tabla de elaboración propia.

Por tanto, podemos considerar que la muestra obtenida es válida para el estudio en cuestión, aunque no es suficientemente representativa para sacar conclusiones a nivel población, tal y como se ha indicado en el apartado “3.4 Técnica de muestreo y muestra del estudio”.

4.2 Análisis de resultados sobre el primer objetivo

Las calificaciones de la primera y la segunda corrección se detallan en la **Tabla 17** del Anexo 2. Como cada profesor corrige las preguntas sobre una puntuación distinta y no es lo mismo restar 0,1 punto a una pregunta de 0,5, que restar 0,1 punto a una pregunta que vale 1 punto,

se decide estandarizar los resultados de las correcciones. Así, todos los valores se transforman en una escala de valor sobre 100 puntos. De este modo, la calificación obtenida se puede interpretar como el porcentaje de la pregunta que es correcta (ver **Tabla 17**). Por ejemplo, el tercer estudiante, que ha obtenido 0,5 puntos sobre 0,75 en la primera corrección, obtiene un porcentaje del 66,6%, mientras que, en la segunda corrección, calificado con un 0,4 sobre 0,75, obtiene un porcentaje del 53,3%.

Para ver la diferencia que hay entre ambas correcciones, se decide restar el valor de la calificación de la primera corrección con el de la segunda, dado que esta última se considera como el valor real o, como mínimo, el más fiable. Con esta operación, se obtiene directamente el porcentaje de aumento o disminución del valor de la nota que presenta la primera corrección sobre la segunda. Este valor se puede observar en la última columna de la **Tabla 17**. En el caso de ejemplo, la variación es de $66,6 - 53,3 = 13,3\%$.

Repitiendo la operación para todos los casos, se obtienen los resultados de la **Tabla 5**:

Tabla 5

Puntuaciones en la primera y segunda corrección y diferencia entre ellas, en promedio.

	Correcciones		Diferencia (%)	Desv. est.
	1ª (%)	2ª (%)		
Total	73,4	66,9	6,5	22,1

Nota. Tabla de elaboración propia.

Es decir, se favorece en un 6,5% a los estudiantes debido al efecto del sesgo cognitivo en los profesores. Es importante mencionar que la desviación estándar es de 22,1 puntos e indica la dispersión de la muestra. Para un análisis más detallado de la distribución y la dispersión de los resultados, se elabora una tabla de frecuencias absolutas (ver **Tabla 6**), donde se detalla la diferencia entre la primera y la segunda corrección, con rangos de diferencia de un 10%, a excepción del valor 0 que, al tener significado por sí mismo, se decide clasificar en una categoría propia. De modo semejante, se pueden observar los mismos resultados de forma gráfica en la **Figura 1**.

Tabla 6

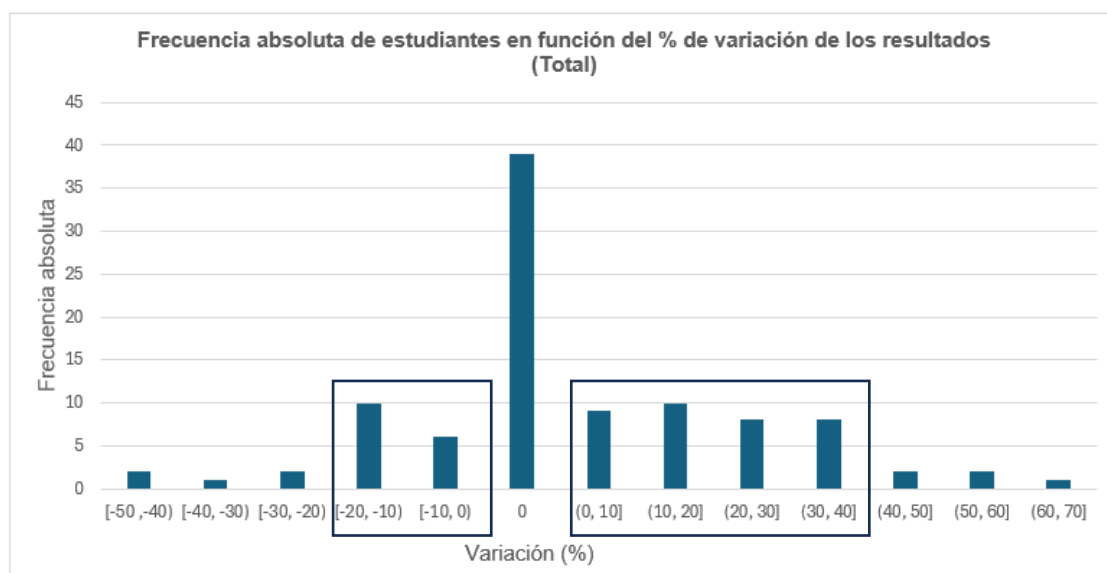
Frecuencias absolutas de los rangos de diferencia entre la primera y la segunda corrección.

	Rango de diferencia %	Frecuencia absoluta	Total de alumnos
Alumnos con la 1ª corrección más baja que la 2ª	[-50, -40)	2	21
	[-40, -30)	1	
	[-30, -20)	2	
	[-20, -10)	10	
	[-10, 0)	6	
Igual	0	39	39
Alumnos con la 1ª corrección más alta que la 2ª	(0, 10]	9	40
	(10, 20]	10	
	(20, 30]	8	
	(30, 40]	8	
	(40, 50]	2	
	(50, 60]	2	
	(60, 70]	1	

Nota. Tabla de elaboración propia.

Figura 1

Frecuencia absoluta de estudiantes en función del % de variación de los resultados (Total)



Nota. Tabla de elaboración propia.

Estos resultados muestran que existe una diferencia entre la primera y la segunda corrección en más de la mitad de los alumnos. Solo el 39% de los alumnos¹¹ es calificado con la misma nota en ambas correcciones. De la muestra restante, un 21% obtiene una diferencia negativa, es decir, presenta una calificación más baja en la primera corrección que en la segunda, mientras que el otro 40% obtiene una diferencia positiva, esto es, presenta una calificación más alta en la primera corrección que en la segunda.

Por un lado, los alumnos que presentan una diferencia negativa se encuentran principalmente en los rangos [-10, 0) y [-20, -10). En estos 2 rangos hay 18 de los 21 alumnos. Por otro lado, los alumnos que presentan una diferencia positiva se encuentran principalmente en los rangos (0, 10], (10, 20], (20, 30], (30, 40]. En estos cuatro rangos hay 35 de los 40 alumnos. En total, en estos 6 rangos se puede observar que hay una presencia más o menos constante de alumnos (entre 6 y 10 en cada uno de ellos). Por el contrario, en los otros rangos, inferiores al -20% y superiores al 40%, hay una presencia residual de alumnos (como máximo 2 por rango), que representan solamente el 10% del total de la muestra, por lo que se les puede considerar como valores extremos.

Con todo, se observa en la **Figura 1** que hay una presencia más marcada de alumnos en los rangos positivos, hecho que hace desviar la media hacia los valores positivos, en concreto, del 6,5%.

Una vez analizada la diferencia entre ambas correcciones y la distribución de los resultados, es preciso determinar en qué tipo de estudiante -clasificado según el rendimiento académico previo- se produce esta diferencia. Para ello, se analiza el promedio de las notas de los tres grupos (A, B y C), comparando las dos correcciones entre ellas. Los resultados obtenidos se muestran en la **Tabla 7**.

Tabla 7

Puntuaciones en las dos correcciones según el tipo de estudiante.

Grupo	Cantidad	Rango de notas	Correcciones		Diferencia (%)	Desv. est.
			1ª (%)	2ª (%)		
A	17	[8, 10]	85,3	79,2	6,1	18,3
B	57	[5, 8)	73,3	66,6	6,7	22,1
C	26	[0, 5)	66,0	59,6	6,5	25,2
Promedio			73,4	66,9	6,5	22,1

Nota. Tabla de elaboración propia.

Es decir, los estudiantes con mejor rendimiento académico previo (grupo A), muestran una diferencia del 6,1% en las correcciones; los estudiantes con un peor rendimiento académico

¹¹ Aunque la **Tabla 6** muestre frecuencias absolutas, dado que la muestra es de 100 alumnos, se puede interpretar como un porcentaje.

(grupo C), un 6,5% -coincidiendo con la media- y el resto (grupo B), un 6,7%. Los tres valores están próximos a la media y, a priori, no parece que exista una diferencia significativa en función del rendimiento previo. En cuanto a la dispersión, hay una tendencia en función del rendimiento académico; cuanto menor es el rendimiento académico, más dispersión presentan las correcciones.

Se modo paralelo, en la **Tabla 8** se analizan las frecuencias absolutas en función del rendimiento académico previo para los grupos A, B y C; viéndose gráficamente el resultado en la **Figura 2**, la **Figura 3** y la **Figura 4**, respectivamente. Como se puede observar, la **Figura 3** es la más semejante a la **Figura 1**, donde está representada el total de la muestra. Seguramente es porque más de la mitad de la muestra se encuentra en el grupo B. Por otra parte, la distribución de alumnos dentro del grupo A también sigue de modo parecido la distribución de alumnos totales, con una marcada presencia de alumnos con variación nula i el resto estando en unos rangos comprendidos entre el -30% y el 40%. Por el contrario, las calificaciones de los alumnos del grupo C presentan un patrón menos parecido a la tendencia general; no hay una presencia tan marcada de alumnos con variación nula en los resultados, aunque parece cierto que la gran mayoría está también entro los valores -20% y 30%. Con todo, los valores extremos no son más que 3 alumnos.

Tabla 8

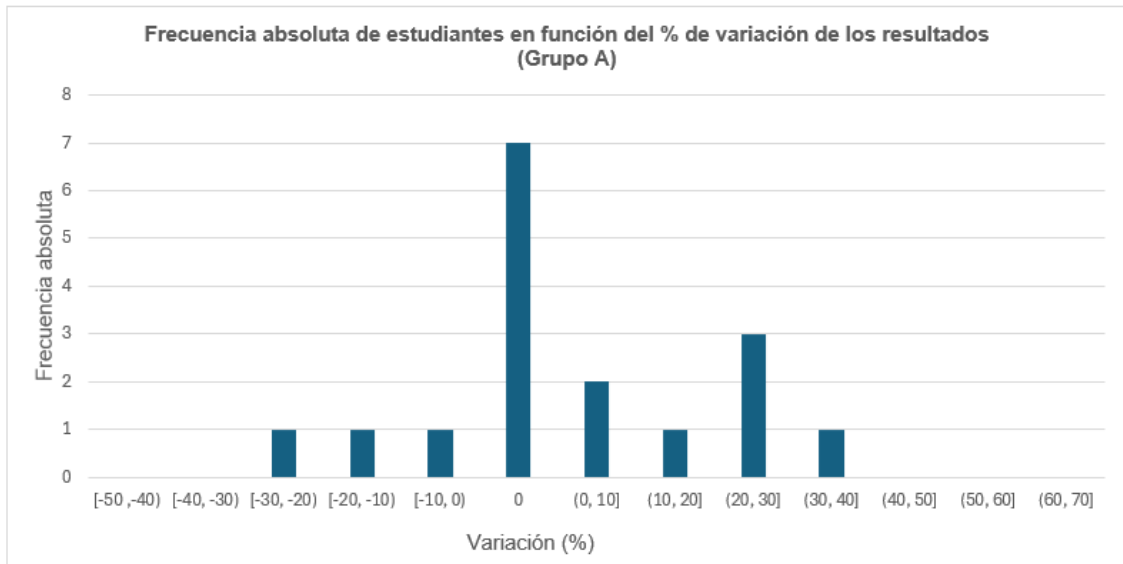
Frecuencia absoluta de estudiantes en función del % de variación de los resultados.

	Rango de diferencia %	Frecuencia absoluta Total	Frecuencia absoluta Grupo A	Frecuencia absoluta Grupo B	Frecuencia absoluta Grupo C	Total de alumnos
Alumnos con la 1ª corrección más baja que la 2ª	[-50, -40)	2	0	1	1	21
	[-40, -30)	1	0	1	0	
	[-30, -20)	2	1	1	0	
	[-20, -10)	10	1	4	5	
	[-10, 0)	6	1	4	1	
Igual	0	39	7	26	6	39
Alumnos con la 1ª corrección más alta que la 2ª	(0, 10]	9	2	4	3	40
	(10, 20]	10	1	4	5	
	(20, 30]	8	3	2	3	
	(30, 40]	8	1	7	0	
	(40, 50]	2	0	2	0	
	(50, 60]	2	0	1	1	
	(60, 70]	1	0	0	1	
Total		100	17	57	26	100

Nota. Tabla de elaboración propia.

Figura 2

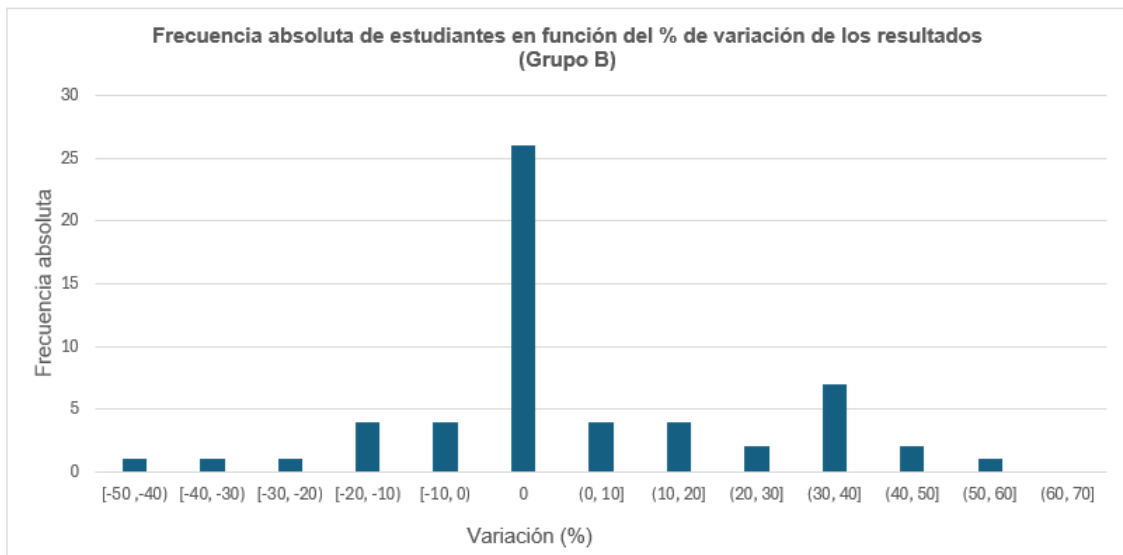
Frecuencia absoluta de estudiantes en función del % de variación de los resultados (Grupo A)



Nota: Figura de elaboración propia.

Figura 3

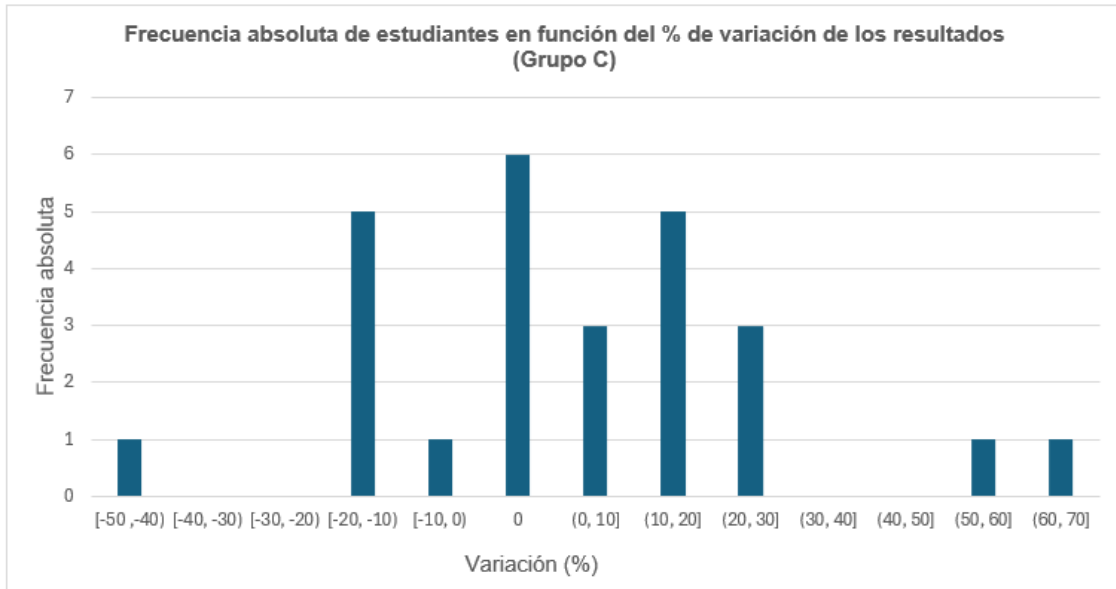
Frecuencia absoluta de estudiantes en función del % de variación de los resultados (Grupo B)



Nota: Figura de elaboración propia.

Figura 4

Frecuencia absoluta de estudiantes en función del % de variación de los resultados (Grupo C)



Nota. Figura de elaboración propia.

4.3 Interpretación de los resultados

La diferencia de medias entre las dos muestras pre y post test señala que ha habido una diferencia del 6,5% entre las dos correcciones. Por tanto, es indiscutible que, a nivel muestral, existe una diferencia entre la primera y la segunda corrección, puntuando de forma más favorable la primera. No obstante, a nivel poblacional, ¿es realmente significativa esta diferencia? Aunque el tamaño de la muestra no es suficientemente grande para inferir resultados, puede resultar interesante ver si, a nivel poblacional, los resultados son estadísticamente relevantes.

La prueba estadística que mejor se adapta para el cálculo de una diferencia entre medias es el test t-student. Admite 3 variantes:

- Prueba t para medias de dos muestras emparejadas.
- Prueba t para dos muestras suponiendo varianzas iguales
- Prueba t para dos muestras suponiendo varianzas desiguales.

En el caso de estudio, como se hace un análisis pre-test y un análisis post-test en el mismo grupo de alumnos, se debe escoger la prueba para medias de dos muestras emparejadas. En esta prueba, se presentan dos hipótesis, la H_0 , conocida como hipótesis nula y la H_1 , conocida como hipótesis alternativa.

$$\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 \neq 0 \end{cases}$$

La hipótesis nula sostiene que la diferencia entre las medias es cero; por el contrario, la hipótesis alternativa sostiene que la diferencia entre las medias es diferente de cero. Por

tanto, la confirmación de la hipótesis nula significa que la diferencia entre las medias de un par de grupos no es estadísticamente significativa, mientras que la confirmación de la hipótesis alternativa significa lo contrario. Para ello, se calcula el valor p. Si este es menor que el nivel de significación deseado (α), entonces se rechaza la hipótesis nula y se concluye que existe una diferencia significativa entre las medias. Si es mayor, se asume lo contrario. Además, como en la hipótesis alternativa hay una desigualdad, la prueba es de dos colas, por lo que el nivel de significación divide entre dos, pasando a ser $\left(\frac{\alpha}{2}\right)$. Se asume en este caso un nivel de significación de 0,05, que es un valor frecuente para este parámetro. Entonces, el nivel de significación de dos colas es $\frac{\alpha}{2} = 0,025$.

Realizando el test t-student para los conjuntos de datos de la primera y la segunda corrección, se obtienen los valores de la **Tabla 9**.

Tabla 9

Resultados estadísticos de la prueba t-student para la muestra total.

	Variable 1	Variable 2
Media	73,43	66,92
Varianza	678,38	794,75
Observaciones	100	100
Coefficiente de correlación de Pearson	0,67	
Diferencia hipotética de las medias	0	
Grados de libertad	99	
Estadístico t	2,94	
P(T<=t) dos colas	0,0040	
Valor crítico de t (dos colas)	1,98	

Nota. Tabla de elaboración propia mediante herramienta de análisis de datos de Excel

Es decir, el valor t-student es 2,94 y el p-valor a dos colas es de 0,0040. Por tanto, como $p < \frac{\alpha}{2}$, se puede rechazar la hipótesis nula y aceptar la hipótesis alternativa. Es decir, hay evidencias estadísticas para afirmar que las medias de la primera y la segunda corrección son significativamente diferentes con un nivel de confianza del 95%, a nivel poblacional.

Si se analizan los resultados según el rendimiento académico previo, se observa que la diferencia entre medias es muy parecida, por lo que cabría esperar resultados semejantes. Sin embargo, el tamaño de la muestra se reduce, por lo que es necesario hacer de nuevo el test de significancia para comprobar si, en estos casos, hay una diferencia estadística suficientemente relevante como para afirmar la presencia del sesgo cognitivo a nivel poblacional. Los datos obtenidos para el grupo A, B y C se muestran en la **Tabla 10**, la **Tabla 11** y la **Tabla 12**, respectivamente.

Tabla 10*Resultados estadísticos de la prueba t-student para el grupo A.*

	Variable 1	Variable 2
Media	85,29	79,22
Varianza	251,47	417,75
Observaciones	17	17
Coefficiente de correlación de Pearson	0,51	
Diferencia hipotética de las medias	0	
Grados de libertad	16	
Estadístico t	1,37	
P(T<=t) dos colas	0,1901	
Valor crítico de t (dos colas)	2,12	

Nota. Tabla de elaboración propia mediante herramienta de análisis de datos de Excel**Tabla 11***Resultados estadísticos de la prueba t-student para el grupo B.*

	Variable 1	Variable 2
Media	73,27	66,61
Varianza	680,65	745,93
Observaciones	57	57
Coefficiente de correlación de Pearson	0,66	
Diferencia hipotética de las medias	0	
Grados de libertad	56	
Estadístico t	2,28	
P(T<=t) dos colas	0,0264	
Valor crítico de t (dos colas)	2,00	

Nota. Tabla de elaboración propia mediante herramienta de análisis de datos de Excel**Tabla 12***Resultados estadísticos de la prueba t-student para el grupo C.*

	Variable 1	Variable 2
Media	66,03	59,55
Varianza	848,02	1049,46
Observaciones	26	26
Coefficiente de correlación de Pearson	0,67	
Diferencia hipotética de las medias	0	
Grados de libertad	25	
Estadístico t	1,31	
P(T<=t) dos colas	0,2018	
Valor crítico de t (dos colas)	2,06	

Nota. Tabla de elaboración propia mediante herramienta de análisis de datos de Excel

A la luz de los resultados, el p-valor de los tres grupos es superior a $\frac{\alpha}{2} = 0,025$. Por tanto, no se puede rechazar la hipótesis nula, es decir, no hay evidencias estadísticas, en ninguno de los tres niveles de rendimiento, para afirmar que las medias de la primera y la segunda corrección son significativamente diferentes con un nivel de confianza del 95%, a nivel poblacional. No obstante, el grupo B presenta un p-valor muy cercano, de 0,0264. Los valores del grupo A (0,1901) y del grupo C (0,2018) distan más del nivel de significación máximo permitido de 0,025, necesario para rechazar la hipótesis nula.

Aunque pueda parecer contradictorio, los resultados son opuestos debido al tamaño de la muestra. En ningún caso la muestra es suficiente como para extrapolar los resultados a nivel poblacional, pero se puede comprobar como en el grupo B, donde la muestra es mayor, casi se obtiene la misma conclusión que con la muestra total.

4.4 Análisis de resultados sobre el segundo objetivo

Tal y como se ha constatado, existe una diferencia del 6,5% entre la primera y la segunda corrección, por lo que, a nivel muestral, se puede afirmar que existe sesgo cognitivo en la evaluación de contenido. Pero ¿qué tipo o tipos de sesgo son más relevantes? Para responder a esta pregunta, se usan los datos provenientes del cuestionario. Las preguntas relativas a este apartado son de la E2 a la E13:

- E2: ¿Recuerdas cuál fue la calificación del primer trimestre del alumno?
- E3: Escríbela
- E4: ¿Recuerdas alguna calificación destacada (buena o mala) en el 1º o 2º trimestre?
- E5: ¿Reconoces alguna característica general del alumno que destaque positivamente?
- E6: ¿Reconoces alguna característica general del alumno que destaque negativamente?
- E7: ¿Has puesto como ejemplo al estudiante en algún trabajo/explicación...?
- E8: ¿Has utilizado algún examen/prueba del alumno para compararlos con otros alumnos?
- E9: ¿Has pensado alguna vez "va a sacar muy buena nota" antes de corregir el contenido?
- E10: ¿Has pensado alguna vez "va a sacar muy mala nota" antes de corregir el contenido?
- E11: ¿Cómo calificarías tu relación con el alumno?
- E12: ¿Tienes afinidad con el alumno en algún tema particular?
- E13: ¿Tienes discrepancias con el alumno en algún tema particular?

Las respuestas a estas preguntas se pueden observar en la **Tabla 18** (ver Anexo 3)¹². A modo de resumen, se elabora la **Tabla 13**, en frecuencias absolutas, y la **Tabla 14**, en porcentajes, para poder comparar mejor los resultados. Se elaboran las tablas respetando los tres grupos mediante los cuales se clasifican las respuestas:

- Diferencia en las correcciones superior a cero ($D > 0$): alumnos favorecidos.
- Diferencias en las correcciones igual a cero ($D = 0$): alumnos ni favorecidos ni perjudicados.
- Diferencia en las correcciones inferior a cero ($D < 0$): alumnos perjudicados.

Por ejemplo, de la pregunta 2 (E2) del cuestionario, para alumnos con diferencia positiva, 12 respuestas obtienen un “Sí” y 28 respuestas un “No”. Dado que el total es 40, las respuestas afirmativas, en porcentaje, se pueden calcular como $\frac{12}{40} = 30\%$ y las negativas como el complementario, es decir, $\frac{28}{40} = 70\%$. De este modo, se pueden comparar los porcentajes de respuestas afirmativas y negativas, en función de la diferencia entre las correcciones. En el caso de las preguntas E3 y E11, al ser respuestas numéricas, se calcula la media aritmética (ver **Tabla 15**) en función de las tres categorías mencionadas.

Tabla 13

Resumen de las respuestas E2, E4-E10 y E12-E13 en frecuencias absolutas.

Respuesta	Diferencia	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13
Sí	Diferencia > 0	12		24	34	20	8	3	16	12		14	3
	Diferencia = 0	4		12	17	11	5	5	12	0		10	3
	Diferencia < 0	5		12	17	13	2	2	10	6		11	1
No	Diferencia > 0	28		16	6	20	32	37	24	28		26	37
	Diferencia = 0	18		10	5	11	17	17	10	22		12	19
	Diferencia < 0	16		9	4	8	19	19	11	15		10	20

Nota: Tabla de elaboración propia.

Tabla 14

Resumen de las respuestas E2, E4-E10 y E12-E13 en porcentajes.

Respuesta	Diferencia	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13
Sí	Diferencia > 0	30		60	85	50	20	8	40	30		35	8
	Diferencia = 0	18		55	77	50	23	23	55	0		45	14
	Diferencia < 0	24		57	81	62	10	10	48	29		52	5
No	Diferencia > 0	70		40	15	50	80	93	60	70		65	93
	Diferencia = 0	82		45	23	50	77	77	45	100		55	86
	Diferencia < 0	76		43	19	38	90	90	52	71		48	95

Nota: Tabla de elaboración propia.

¹² *Nota.* No se ha contestado el cuestionario para todos los alumnos. Se ha obtenido una muestra total de 83 alumnos, descartando solamente 17 alumnos, los cuales no habían presentado diferencia entre las dos correcciones.

Tabla 15

Resumen de las respuestas E3 y E11. Media aritmética.

Operación	Diferencia	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13
Media	Diferencia > 0		4,9								3,7		
	Diferencia = 0		7,8								3,5		
	Diferencia < 0		5,2								3,6		

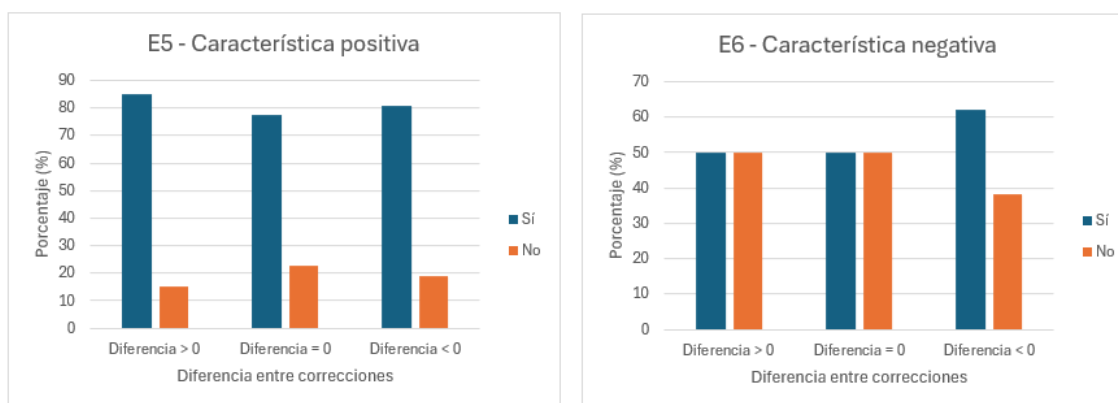
Nota: Tabla de elaboración propia.

A continuación, se detalla el resultado obtenido. Tal y como se puede observar en la tabla de dimensiones, hay sesgos que se describen con un aspecto positivo y otro negativo, como por ejemplo el efecto halo/cuerno, donde se pregunta por una característica positiva (E5) y otra negativa (E6); el sesgo de confirmación, donde se pregunta por una expectativa buena (E9) o una mala (E10), y el sesgo de afinidad, en el que se pregunta si se tiene una buena (E12) o mala (E13) relación con el alumno. En este tipo de sesgos se contrastará principalmente los alumnos con diferencia positiva entre la primera y la segunda calificación y los alumnos con diferencia negativa.

A tal efecto, se elabora la **Figura 5**, la **Figura 6** y la **Figura 7**. En la **Figura 5**, el porcentaje de alumnos en los que se reconoce una característica positiva no varía en función de la diferencia obtenida en las dos calificaciones. No obstante, los alumnos perjudicados reciben un porcentaje (62%) más alto de respuestas donde se reconoce una característica negativa de ellos, en comparación con los alumnos favorecidos (50%). Es decir, se deduce que se produce el efecto cuerno, pero no el efecto halo.

Figura 5

Resultados relativos al efecto halo/cuerno.



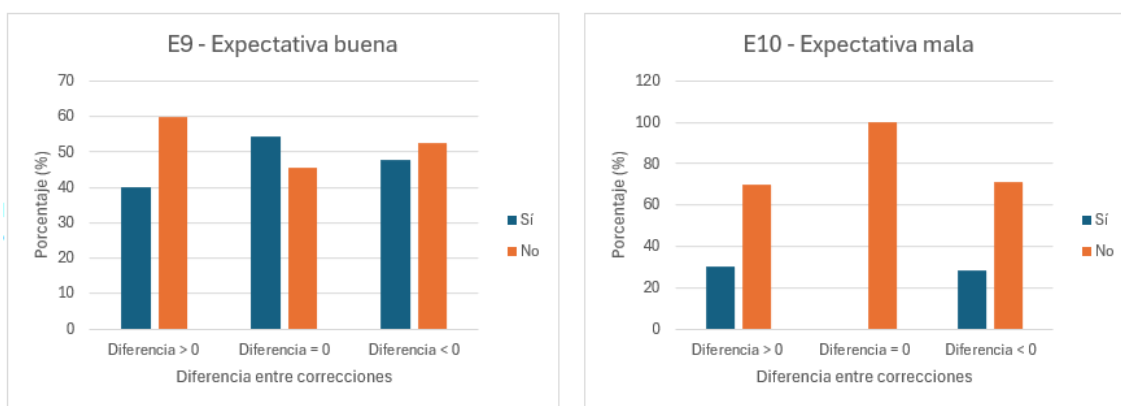
Nota: Figura de elaboración propia.

En cuanto al sesgo de expectativa, tal y como se puede comprobar en la **Figura 6**, el porcentaje de alumnos con expectativa buena es del 40% para los alumnos con diferencia positiva, mientras que es del 48% en los alumnos con diferencia negativa. Es decir, los profesores no tienen mejor expectativa en los alumnos beneficiados por el sesgo cognitivo.

En cuanto a las expectativas negativas, los resultados son prácticamente idénticos en alumnos con diferencia positiva (un 30%) respecto a alumnos con diferencia negativa (un 29%). Además, no existe ningún profesor que responda que tiene una expectativa mala y que haya cometido sesgo en las correcciones. De esto se deduce que la proyección de expectativas malas no influye en que el sesgo cognitivo sea favorable o desfavorable, aunque sí influye en que aparezca o no aparezca. Por tanto, para reducir el sesgo cognitivo es importante no proyectar expectativas malas sobre los alumnos.

Figura 6

Resultados relativos al sesgo de confirmación.



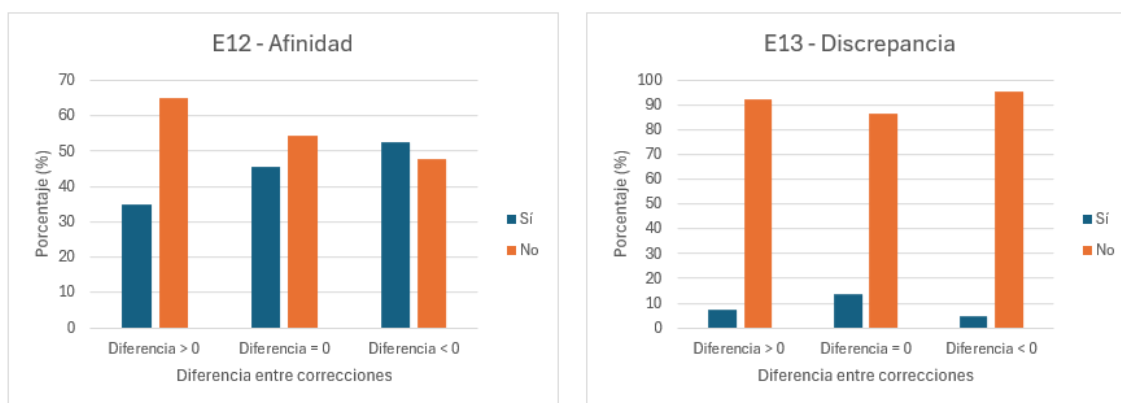
Nota: Figura de elaboración propia.

Si se analiza el sesgo de afinidad, se puede observar en la **Figura 7** que la afinidad en alumnos con diferencia positiva es del 35%, en alumnos sin diferencia es del 45% y en alumnos con diferencia negativa es del 52%. Es decir, al revés de lo que se podría pensar, los alumnos perjudicados por el sesgo cognitivo son los que más afinidad tienen con el profesorado y, al contrario, los alumnos beneficiados son los que menos sesgo cognitivo sufren por parte de sus correctores.

En cuanto al análisis de las discrepancias, se puede observar en la misma **Figura 7** que muy pocos profesores han afirmado tener discrepancias con algún alumno, por lo que se considera que el profesorado puede no haber sido del todo sincero, al ser una pregunta en cierto modo comprometedor.

Figura 7

Resultados relativos al sesgo de afinidad.



Nota: Figura de elaboración propia.

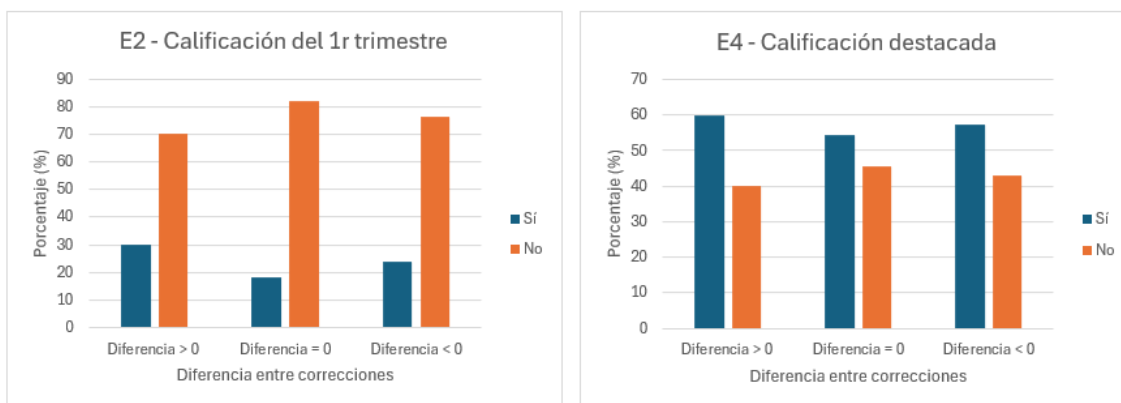
Existe un segundo grupo de sesgos los cuales no son definidos mediante una pregunta con aspecto positivo y otro negativo, sino con dos preguntas acerca del sesgo en cuestión. Estos son el sesgo de anclaje y el sesgo de comparación. En el primero, solo se pretende saber si se recuerda la calificación del primer trimestre (E2) o si se recuerda una calificación destacada (E4), pero se no distingue entre una calificación buena o mala. En el sesgo de comparación ocurre algo semejante; solo se pregunta si se ha puesto como ejemplo en público (E7) o si se ha utilizado para compararlo en privado (E8), pero en ningún caso pregunta si el ejemplo o la comparación han sido destacando aspectos positivos o negativos. Entonces, para el análisis de estas dos cuestiones se va a comparar principalmente los alumnos con diferencia positiva o negativa entre las dos calificaciones (es decir, distinta de cero) con los alumnos con diferencia nula.

Analizando el sesgo de anclaje mediante la **Figura 8**, se puede observar que el porcentaje de alumnos cuyos profesores recuerdan la nota del primer trimestre¹³ es más elevado en alumnos que han sufrido sesgo cognitivo de modo favorable (30%) y desfavorable (24%), respecto aquellos alumnos que no han sufrido variación (18%). Aunque es cierto que no existe una diferencia muy notoria, se puede afirmar que recordar la nota del trimestre anterior afecta a la aparición del sesgo cognitivo. El mismo patrón se repite en la pregunta E4, si bien es cierto que las diferencias son muy pequeñas.

¹³ Se toma como criterio para la validez de las respuestas E2 la información contenida en la pregunta E3, que pide que se escriba la nota en caso de que se recuerde. Si la nota varía en más de 1 punto con la nota real, se rechaza como válida y se califica como que no se recuerda. No obstante, no sucede en ninguno de los resultados, dado que la mayor diferencia es de 1 punto.

Figura 8

Resultados relativos al sesgo de anclaje.



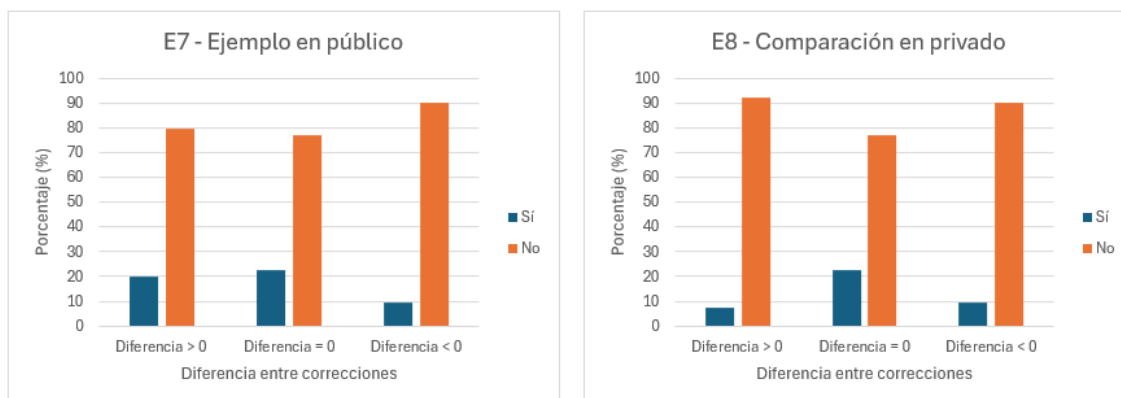
Nota: Figura de elaboración propia.

Finalmente, si se analiza el sesgo de comparación se observa en la **Figura 9** que el porcentaje de alumnos que son puestos como ejemplo en público es parecido en alumnos favorecidos (20%) respecto aquellos en los que no se produce diferencia (23%). La diferencia es más notoria cuando se compara este último grupo con los alumnos que reciben una diferencia negativa en sus correcciones; hay un 13% menos de alumnos que son puestos como ejemplo.

En cuanto a la comparación de los resultados de los alumnos en privado, los alumnos que no sufren sesgo son comparados en privado un 23% de las veces, mientras que los alumnos que sufren sesgo solo son comparados un 8% (diferencia positiva) y un 10% (diferencia negativa). Tal y como se puede comprobar, no el hecho de comparar una calificación, sea en público o en privado, favorezca la aparición del sesgo cognitivo en el profesorado. Es más, en vistas a los resultados, se puede afirmar que cuando se comparan los resultados hay menos sesgo cognitivo.

Figura 9

Resultados relativos al sesgo de comparación.



Nota: Figura de elaboración propia.

5 Conclusiones

5.1 Revisión de los objetivos y discusión

Después de investigar acerca del fenómeno del sesgo cognitivo en la evaluación de contenido en Educación Secundaria, en el centro donde se han desarrollado las prácticas de formación del profesorado, se llega a varias conclusiones.

En relación con el primer objetivo, se comprueba que existe sesgo cognitivo en la corrección de contenido en la muestra seleccionada de 100 alumnos, con una diferencia entre las calificaciones de un 6,5%. Además, considerando la muestra total, hay evidencias estadísticas para sostener que las medias de la primera y la segunda corrección son significativamente diferentes con un nivel de confianza del 95%.

Por el contrario, analizando los resultados en función del rendimiento académico previo y, por tanto, reduciendo la muestra, no se puede sostener que los porcentajes de diferencia obtenidos en las dos correcciones sean estadísticamente significativos en ninguno de los tres casos. El grupo A, con 26 alumnos con un rendimiento académico entre 8 y 10 puntos obtiene, de media, una diferencia entre las calificaciones de un 6,1%; el grupo B, con 57 alumnos con un rendimiento entre 5 y 7,9, obtiene un 6,7% y el grupo C, con 17 alumnos con calificaciones previas inferiores a 5, obtiene un 6,5%. No obstante, sí que se observa una tendencia: cuanto menor es el rendimiento académico previo, más dispersión presentan los resultados. Este fenómeno puede ser debido a varios factores como, por ejemplo, una mayor dificultad en la comprensión de la letra en el contenido a corregir, dado que los estudiantes de este grupo suelen mostrar una peor caligrafía.

En relación con el segundo objetivo, también se obtienen conclusiones relevantes. Se observa que el sesgo de anclaje influye en la aparición de sesgo cognitivo. Concretamente, los profesores que recuerdan la calificación obtenida por el alumnado en el trimestre anterior tienden a favorecer o perjudicar su evaluación posterior.

En cuanto al efecto halo, los datos no respaldan su presencia en la corrección de contenido de la muestra. Por el contrario, el efecto cuerno sí que se evidencia en los resultados, lo que sugiere que una impresión negativa previa puede influir negativamente en la evaluación de contenido.

Al contrario de lo esperado, el hecho de comparar resultados entre evaluaciones contribuye a disminuir el sesgo cognitivo en la muestra seleccionada, reduciendo así las diferencias entre ambas correcciones.

En cuanto al sesgo de confirmación, se determina que la existencia de expectativas positivas por parte del profesorado no conlleva necesariamente la aparición de sesgo cognitivo favorable hacia los alumnos. Por el contrario, la proyección de expectativas negativas por parte del profesorado sí que implica la aparición de sesgo cognitivo, aunque no influye en

que sea favorable o desfavorable. Dicho de otro modo, ningún profesor con diferencia nula entre las correcciones afirma proyectar expectativas negativas sobre los alumnos, por lo que se concluye que para la reducción del sesgo cognitivo en la corrección de contenido es necesario evitar la proyección de expectativas malas sobre el alumnado.

Finalmente, se concluye que los alumnos que se ven más perjudicados en sus correcciones son los que más afinidad presentan con el profesorado, mientras que los alumnos beneficiados tienden a mostrar menos afinidad. En cuanto al sesgo de afinidad en su versión negativa, es decir, cuando existen discrepancias entre alumno y profesor, los resultados de las encuestas presentan valores muy bajos, por lo que se sospecha que no son fiables y que los profesores no han contestado a la pregunta de forma fidedigna.

Con todo, se da respuesta a la pregunta de investigación inicial y a los dos grandes objetivos planteados en esta investigación: se determina el grado de sesgo cognitivo presente en la evaluación de contenido en Educación Secundaria y se relaciona con los distintos tipos de sesgo, evaluando la relevancia e influencia de cada uno de ellos.

Las implicaciones derivadas de esta investigación afectan principalmente al proceso evaluativo en sí mismo, al profesorado y al alumnado. En primer lugar, se cuestiona la objetividad del proceso evaluativo y la necesidad de adoptar mecanismos a nivel escolar para fomentar una evaluación más justa e imparcial. En segundo lugar, se evidencia la necesidad de formar e informar al profesorado sobre la existencia de los sesgos cognitivos, inconscientes por definición en el proceso educativo, pero que se pueden hacer conscientes mediante la sensibilización y la formación continua en este ámbito. También se pone de manifiesto la importancia de implementar estrategias que reduzcan o anulen los distintos tipos de sesgo, tales como el uso de rúbricas, las correcciones ciegas o la temporalización en la corrección. Por último, pero no menos importante, es fundamental que el alumno perciba la evaluación de su trabajo como un proceso justo que refleja verdaderamente su desempeño académico.

5.2 Limitaciones de la investigación

Aunque los resultados y conclusiones son relevantes, es importante no perder de vista los límites dentro de los cuales se circunscribe la investigación y, por ende, los resultados y conclusiones mismas. Al realizar el trabajo final de máster en el centro donde se desarrollan las prácticas curriculares, el tiempo limitado para realizar la investigación juega un papel fundamental. Tanto las dobles correcciones como la encuesta se desarrollan en un espacio de tiempo muy breve, de dos o tres meses, por lo que es complicado obtener una muestra de alumnos suficientemente grande como para poder extrapolar los resultados a nivel poblacional. No obstante, se obtiene una muestra de 100 alumnos y 200 correcciones.

Otro factor que ha limitado en gran medida la investigación es el número de profesorado implicado en las dobles correcciones. Si bien el número de alumnos es elevado, el número de profesores es bastante reducido (solo 5 participantes), por lo que puede haber una influencia desproporcionada de algún profesor.

También cabe destacar que la investigación obvia alumnos con adaptaciones curriculares o sistemas de puntuación especial, por lo que la muestra puede ser sesgada, especialmente en el grupo de alumnos con nota previa inferior a cinco.

5.3 Líneas prospectivas

En cuanto a las líneas prospectivas, un primer paso puede consistir en solventar las limitaciones de la presente investigación, esto es:

- Aumentar la muestra de los alumnos (hasta 384, como mínimo).
- Aumentar la muestra del profesorado participante en la investigación.
- Incluir en la muestra a alumnos con adaptaciones curriculares.

Además, sería interesante realizar el estudio en otros centros educativos, dado que el centro donde se realiza la investigación destaca especialmente por el vínculo entre alumnado y profesorado, por lo que puede haber más influencia en algunos sesgos, como el de afinidad o el de confirmación.

A nivel metodológico, también podría resultar positivo escoger otro criterio para determinar el rendimiento académico previo, como la nota media del curso anterior o la nota media del primer trimestre del curso en cuestión (no de la asignatura en cuestión, sino la media entre todas las asignaturas). Solventar la obtención de información fiable en cuanto a nivel de discrepancia entre alumno y profesor también es un reto que se podría tratar en investigaciones posteriores.

Sea cual sea la línea de investigación futura, es primordial seguir estudiando el efecto del sesgo cognitivo en la corrección de contenido. No solamente por la justicia en sí misma, sino porque la evaluación recibida a lo largo de los años, a nivel académico, puede ser un referente y un reflejo en el modo de evaluar los propios resultados y logros personales a lo largo de la vida. Solo así el alumno se podrá forjar una identidad acorde con su realidad, capaz de ofrecer al mundo un abanico de potencialidades que deberán materializarse mediante el trabajo y el esfuerzo diario.

6 Bibliografía

- Allen, M. y Coole, H. (2012). Experimenter Confirmation Bias and the Correction of Science Misconceptions. *Journal of Science Teacher Education*, 23(4), 387-405.
- Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, 41(3), 258–290. <https://doi.org/10.1037/h0055756>
- Briceño, C.E. (2024). Incidencia del Síndrome de Burnout sobre el desempeño en educación superior. *Revista Educación y Sociedad*, 5(9), 3-14.
<https://doi.org/10.53940/reys.v5i9.166>
- Bhargava, S. y Fisman R. (2014). Contrast Effects in Sequential Decisions: Evidence from Speed Dating. *The Review of Economics and Statistics*, 96(3), 444–457.
doi: https://doi.org/10.1162/REST_a_00416
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.
- Clark, P. y Zygmunt, E. (2014). A close encounter with personal bias: Pedagogical implications for teacher education. *Journal of Negro Education*, 83(2), 147–161. <https://doi.org/10.7709/jnegroeducation.83.2.0147>
- Clínica Universidad de Navarra. (2023). Sesgo cognitivo. En *Diccionario Médico*. Recuperado en 11/05/2025 de <https://www.cun.es/diccionario-medico/terminos/sesgo-cognitivo#:~:text=Se%20trata%20de%20una%20desviaci%C3%B3n,la%20salud%20y%20el%20bienestar>
- Del Valle, M. (2013). Una aproximación a los paradigmas de evaluación cuantitativa vs evaluación cualitativa. *Didáctica y Educación*, 4(3), 109-122.
- Díaz-Barriga, F. y Hernández, G. (2002). *Estrategias docentes para un aprendizaje significativo: Una interpretación constructivista* (2ª ed.). McGraw-Hill.
- Epley, N. y Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science*, 17(4), 311-318.
<https://doi.org/10.1111/j.1467-9280.2006.01704.x>
- Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255-278.
<https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Fiske, S. y Taylor, S. (1991). *Social cognition* (2ª ed.). McGraw-Hill.
- Furnham, A. y Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1), 35–42. <https://doi.org/10.1016/j.socec.2010.10.008>

- Goldbach, C., Sickmann, J. y Pitz, T. (2021). Sequential decision bias – evidence from grading exams. *Applied Economics*, 54(32), 3727–3739.
<https://doi.org/10.1080/00036846.2021.1976390>
- Gómez Díaz, D. y Martínez López, J.M. (1997). La evaluación en ciencias sociales: la historia económica. *Boletín del Instituto de Estudios Almerienses*, 15, 133-146.
- Heredia, A. (2009). Dos formas diferenciadas de evaluación didáctica: evaluación normativa para seleccionar a los alumnos y evaluación criterial para el dominio del conocimiento básico. *Bordón* 61(4), 39-48.
<https://dialnet.unirioja.es/servlet/articulo?codigo=3109891>
- Joy, M. C. (2025, 27 de febrero). First come, better served: Reality of sequential grading biases, GPA disparities. *The Tulane Hullabaloo*.
<https://tulanehullabaloo.com/68854/data/first-come-better-served-reality-of-sequential-grading-biases-gpa-disparities/>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Keoma, S. (2021). *La influencia de los sesgos cognitivos en la utilización de la tipografía digital* [Tesis de Maestría no publicada]. Universidad Nacional de Colombia.
- Klayman, J. (1995). Varieties of confirmation bias. *Psychological Bulletin*, 118(3), 338-355.
<https://doi.org/10.1037/0033-2909.118.3.338>
- Martínez Pérez, A. (2010). El síndrome del Burnout. Evolución conceptual y estado actual de la cuestión. *Vivat Academia*, 112, 42-80.
- Massolo, A. y Traversi, M. (2021). ¿Es posible mitigar sesgos cognitivos? Un análisis crítico de diversas propuestas para reducir el sesgo a mi favor. *Prometeica – Revista de Filosofía y Ciencias*, 23, 60-76.
- Mora-Vicarioli, F. (2019). Estado del arte de la evaluación de los aprendizajes en la modalidad del e-learning desde la perspectiva de evaluar para aprender: precisiones conceptuales. *Revista electrónica Calidad en la Educación Superior*, 10(1), 58-95.
- Nauts, S., Langner, O., Huijsmans, I., Vonk, R. y Wigboldus, D. H. J. (2014). Forming impressions of personality. *Social Psychology*, 45(3), 153–163.
<https://doi.org/10.1027/1864-9335/a000179>
- Nisbett, R. E. y Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Prentice-Hall.
- Nisbett, R. E. y Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35(4), 250–256.
<https://doi.org/10.1037/0022-3514.35.4.250>

- Ravela, P. (2013). La evaluación de y en la educación (I). *Boletín del Instituto Nacional de Evaluación Educativa*, 1-4.
- Navaridas, F. (2013). *Procesos y contextos educativos: nuevas perspectivas para la práctica docente*. Genuve Ediciones.
- Rosales, C. (1990). *Evaluar es reflexionar sobre la enseñanza* (3ª ed.). Narcea.
- Rosenthal, R. y Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. Holt, Rinehart & Winston.
- Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159-179.
<https://doi.org/10.1080/02602930801956059>
- Simon, H. A. (1947). *Administrative behavior: A study of decision-making processes in administrative organizations* (2ª ed.). Free Press.
- Slovic, P. (1987). The perception of risk. *Science*, 236(4799), 280-285.
- Sola, T. et al. (2020). *Investigación Educativa e Inclusión. Retos actuales en la sociedad del siglo XXI*. Dykinson.
- Stufflebeam, D. y Shinkfield, A. (1987). *Evaluación sistemática. Guía teórica y práctica*. Paidós Ibérica.
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4(1), 25–29. <https://doi.org/10.1037/h0071663>
- Todd, P. y Gigerenzer, G. (2000). Précis of Simple heuristics that make us smart. *Behavioral and Brain Sciences*, 23, 727-780.
- Tversky, A. y Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131. <https://doi.org/10.1126/science.185.4157.1124>

ANEXOS

Anexo 1. Tabla de dimensiones del cuestionario

Tabla 16

Tabla de dimensiones del cuestionario.

Variable	Dimensión	Descripción	Número	Preguntas	Tipo de opción	Filtro
-	Datos de identificación		0	Nombre del profesor	Abierta	
			1	Nombre del alumno	Abierta	
VD	Sesgo de anclaje	Tiene en mente un ancla general	2	¿Recuerdas cuál fue la calificación del primer trimestre del alumno?	Dicotómica	Sí
			3	Escribela	Numérica	
		Tiene en mente un ancla concreta	4	¿Recuerdas alguna calificación destacada (buena o mala) en el 1r o 2º trimestre?	Dicotómica	
VD	Efecto halo/cuerno	Existe un efecto halo sobre el alumno	5	¿Reconoces alguna característica general del alumno que destaque positivamente?	Opción múltiple excluyente	
		Existe un efecto cuerno sobre el alumno	6	¿Reconoces alguna característica general del alumno que destaque negativamente?	Opción múltiple excluyente	
VD	Sesgo de comparación	Comparación del estudiante en público	7	¿Has puesto como ejemplo al estudiante en algún trabajo/explicación...?	Opción múltiple excluyente	
		Comparación del estudiante en privado	8	¿Has utilizado algún examen/prueba del alumno para compararlos con otros alumnos?	Opción múltiple excluyente	
VD	Sesgo de confirmación	Expectativa proyectada de un estudiante bueno	9	¿Has pensado alguna vez "va a sacar muy buena nota" antes de corregir el contenido?	Dicotómica	
		Expectativa proyectada de un estudiante malo	10	¿Has pensado alguna vez "va a sacar muy mala nota" antes de corregir el contenido?	Dicotómica	
VD	Sesgo de afinidad	Medición de la conexión o afinidad existente	11	¿Cómo calificarías tu relación con el alumno?	Numérica gradual	
		Temas comunes de interés	12	¿Tienes afinidad con el alumno en algún tema particular?	Dicotómica	
		Temas comunes de discrepancia	13	¿Tienes discrepancias con el alumno en algún tema particular?	Dicotómica	

Nota. Tabla de elaboración propia.

Anexo 2. Resultados de la doble corrección

Tabla 17

Resultados de la doble corrección.

Estudiante	Profesor	Calificación máxima Puntos	Correcciones				Diferencia %
			1 ^a		2 ^a		
			Puntos	%	Puntos	%	
1			0,6	80,0	0,6	80,0	0,0
2			0,8	100,0	0,8	100,0	0,0
3			0,5	66,7	0,4	53,3	13,3
4	1	0,75	0,8	100,0	0,8	100,0	0,0
5			0,2	26,7	0,2	26,7	0,0
6			0,6	80,0	0,7	86,7	-6,7
7			0,6	80,0	0,7	86,7	-6,7
8			0,8	100,0	0,7	93,3	6,7
9			0	0,0	0	0,0	0,0
10			0,5	100,0	0,25	50,0	50,0
11			0,3	60,0	0	0,0	60,0
12			0,5	100,0	0,5	100,0	0,0
13			0,5	100,0	0,5	100,0	0,0
14			0,3	60,0	0,25	50,0	10,0
15			0,3	60,0	0,25	50,0	10,0
16			0,4	80,0	0,4	80,0	0,0
17			0,2	40,0	0,25	50,0	-10,0
18			0,4	80,0	0,4	80,0	0,0
19			0,5	100,0	0,4	80,0	20,0
20	2	0,5	0,4	80,0	0,4	80,0	0,0
21			0,4	80,0	0,4	80,0	0,0
22			0,25	50,0	0,25	50,0	0,0
23			0,5	100,0	0,4	80,0	20,0
24			0,25	50,0	0,25	50,0	0,0
25			0,5	100,0	0,4	80,0	20,0
26			0,2	40,0	0,2	40,0	0,0
27			0	0,0	0,25	50,0	-50,0
28			0,5	100,0	0,5	100,0	0,0
29			0,5	100,0	0,5	100,0	0,0
30			0,25	50,0	0,25	50,0	0,0
31			0,25	50,0	0,5	100,0	-50,0

Estudiante	Profesor	Calificación máxima	Correcciones				Diferencia
			1ª		2ª		
			Puntos	%	Puntos	%	
32			1	100,0	1	100,0	0,0
33			0,5	50,0	0,3	30,0	20,0
34			1	100,0	0,9	90,0	10,0
35			1	100,0	0,6	60,0	40,0
36			1	100,0	0,6	60,0	40,0
37			0,8	80,0	0,9	90,0	-10,0
38			1	100,0	0,8	80,0	20,0
39			0,5	50,0	0,3	30,0	20,0
40			1	100,0	1	100,0	0,0
41			1	100,0	0,6	60,0	40,0
42			0,6	60,0	0,6	60,0	0,0
43			1	100,0	0,7	70,0	30,0
44	3	1	0,8	80,0	0,4	40,0	40,0
45			0,6	60,0	0,6	60,0	0,0
46			0,5	50,0	0,5	50,0	0,0
47			0,5	50,0	0,3	30,0	20,0
48			1	100,0	0,7	70,0	30,0
49			1	100,0	1	100,0	0,0
50			0,6	60,0	0,3	30,0	30,0
51			0,7	70,0	0,3	30,0	40,0
52			0,7	70,0	0,6	60,0	10,0
53			1	100,0	0,6	60,0	40,0
54			1	100,0	0,7	70,0	30,0
55			1	100,0	0,6	60,0	40,0
56			0,8	80,0	1	100,0	-20,0
57			0,5	50,0	0,3	30,0	20,0
58			0,4	80,0	0,5	100,0	-20,0
59			0,5	100,0	0,5	100,0	0,0
60			0,5	100,0	0,2	40,0	60,0
61			0,5	100,0	0,5	100,0	0,0
62			0,5	100,0	0,5	100,0	0,0
63			0,2	40,0	0,3	60,0	-20,0
64			0,5	100,0	0,5	100,0	0,0
65			0,2	40,0	0,3	60,0	-20,0
66			0,5	100,0	0,5	100,0	0,0
67			0,5	100,0	0,5	100,0	0,0
68	4	0,5	0,5	100,0	0,5	100,0	0,0
69			0,5	100,0	0,5	100,0	0,0
70			0,4	80,0	0,5	100,0	-20,0
71			0,25	50,0	0,3	60,0	-10,0
72			0,5	100,0	0,5	100,0	0,0
73			0,5	100,0	0,3	60,0	40,0
74			0,25	50,0	0	0,0	50,0
75			0,5	100,0	0,5	100,0	0,0
76			0	0,0	0	0,0	0,0
77			0,5	100,0	0,5	100,0	0,0
78			0,4	80,0	0,5	100,0	-20,0

Estudiante	Profesor	Calificación máxima Puntos	Correcciones				Diferencia %
			1ª		2ª		
			Puntos	%	Puntos	%	
79			0,8	40,0	0,7	35,0	5,0
80			1,4	70,0	1,8	90,0	-20,0
81			0,8	40,0	1,1	55,0	-15,0
82			1	50,0	1,5	75,0	-25,0
83			0,8	40,0	0,3	15,0	25,0
84			1,1	55,0	1,3	65,0	-10,0
85			1,6	80,0	2	100,0	-20,0
86			2	100,0	0,6	30,0	70,0
87			1	50,0	1	50,0	0,0
88			1,2	60,0	1,2	60,0	0,0
89	5	2	1	50,0	0,8	40,0	10,0
90			1,1	55,0	1,5	75,0	-20,0
91			1	50,0	1	50,0	0,0
92			1,6	80,0	1,3	65,0	15,0
93			1	50,0	0,5	25,0	25,0
94			1,8	90,0	1,3	65,0	25,0
95			1	50,0	0,8	40,0	10,0
96			1,2	60,0	1,8	90,0	-30,0
97			1	50,0	0,5	25,0	25,0
98			1	50,0	1	50,0	0,0
99			1	50,0	1,8	90,0	-40,0
100			1,6	80,0	1,5	75,0	5,0

Nota: Tabla de elaboración propia.

Anexo 3. Respuestas del cuestionario

Tabla 18

Respuestas del cuestionario.

Estudiante	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13
1	Sí	9	Sí	Sí	Sí	Sí	Sí	Sí	No	5	Sí	No
2	Sí	7,5	Sí	Sí	No	No	No	Sí	No	5	Sí	Sí
3	Sí	3,7	No	Sí	Sí	No	No	No	Sí	5	Sí	No
4	Sí	8,8	Sí	Sí	Sí	No	No	Sí	No	5	Sí	Sí
5	Sí	6	No	No	Sí	No	No	No	No	3	No	Sí
6	Sí	4,4	Sí	Sí	Sí	Sí	No	No	Sí	4	Sí	Sí
7	Sí	8,6	Sí	Sí	Sí	No	No	Sí	No	5	Sí	No
8	Sí	3,2	Sí	Sí	Sí	No	No	No	Sí	5	Sí	Sí
9												
10	Sí	5	Sí	Sí	Sí	No	No	No	No	4	No	No
11	Sí	7	Sí	Sí	No	Sí	No	Sí	No	5	Sí	No
12												
13												
14	Sí	6	Sí	Sí	No	Sí	No	Sí	No	3	No	No
15	Sí	6	Sí	Sí	No	Sí	No	Sí	No	5	Sí	No
16												
17	Sí	4	Sí	Sí	No	No	No	No	No	3	Sí	No
18												
19	Sí	6	Sí	Sí	No	Sí	No	Sí	No	4	No	No
20												
21												
22												
23	Sí	5	Sí	Sí	No	No	No	No	No	5	Sí	No
24												
25	Sí	3	Sí	Sí	Sí	No	No	No	No	3	No	No
26												
27	Sí	2	Sí	No	Sí	No	No	No	Sí	4	No	No
28												
29												
30												
31	Sí	7	Sí	Sí	No	Sí	No	Sí	No	5	Sí	No

Estudiante	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13
32	No		Sí	Sí	No	No	No	Sí	No	4	Sí	No
33	No		Sí	Sí	Sí	Sí	Sí	No	Sí	4	No	No
34	No		No	No	No	No	No	No	No	3	No	No
35	No		No	Sí	Sí	No	No	Sí	No	3	No	No
36	No		Sí	Sí	No	Sí	No	No	No	3	No	No
37	No		Sí	Sí	No	No	No	No	No	4	Sí	No
38	No		Sí	Sí	No	No	No	No	No	4	No	No
39	No		Sí	No	No	No	No	No	Sí	3	No	No
40	No		Sí	Sí	No	No	Sí	Sí	No	3	No	No
41	No		Sí	Sí	No	No	No	No	Sí	3	No	No
42	No		Sí	Sí	No	Sí	No	Sí	No	4	Sí	No
43	No		No	Sí	No	No	No	Sí	No	4	Sí	No
44	No		No	Sí	No	No	No	No	No	4	No	No
45	No		No	No	No	No	No	No	No	3	No	No
46	No		Sí	Sí	No	No	No	No	No	5	Sí	No
47	No		Sí	No	Sí	No	No	No	Sí	4	No	No
48	No		Sí	Sí	No	Sí	Sí	Sí	No	5	Sí	No
49	No		No	Sí	No	No	No	No	No	4	Sí	No
50	No		No	Sí	No	No	No	Sí	No	3	No	No
51	No		Sí	Sí	No	No	No	No	No	4	Sí	No
52	No		No	No	No	No	No	Sí	No	3	No	No
53	No		No	Sí	Sí	No	No	No	Sí	4	No	No
54	No		No	No	No	No	No	No	No	3	No	No
55	No		Sí	Sí	No	Sí	Sí	Sí	No	5	Sí	No
56	No		No	No	No	No	No	Sí	No	3	No	No
57	No		No	Sí	No	No	No	No	No	4	Sí	No

Estudiante	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13
58	No		Sí	Sí	Sí	No	No	No	No	3	No	No
59	No		Sí	Sí	No	No	No	Sí	No	3	No	No
60	No		Sí	No	Sí	No	No	No	Sí	3	No	No
61	No		Sí	Sí	Sí	Sí	Sí	Sí	No	5	Sí	No
62	No		Sí	Sí	No	No	No	Sí	No	3	No	No
63	No		No	No	No	No	No	No	No	3	No	No
64	No		Sí	Sí	No	Sí	Sí	Sí	No	4	Sí	No
65	No		No	No	Sí	No	No	No	Sí	2	No	No
66	No		No	Sí	Sí	No	No	No	No	3	Sí	No
67	No		No	Sí	Sí	No	No	Sí	No	3	No	No
68	No		No	Sí	Sí	No	No	No	No	3	No	No
69	No		Sí	Sí	Sí	Sí	Sí	Sí	No	2	No	No
70	No		No	Sí	Sí	No	No	No	No	3	Sí	No
71	No		No	Sí	Sí	No	No	No	No	3	Sí	No
72	No		No	No	Sí	No	No	No	No	3	No	No
73	No		No	Sí	Sí	No	No	No	No	3	No	No
74	No		Sí	Sí	No	No	No	Sí	No	4	No	No
75	No		No	No	Sí	No	No	No	No	3	No	No
76	No		No	No	Sí	No	No	No	No	2	No	No
77	No		No	Sí	No	No	No	No	No	3	No	No
78	No		Sí	Sí	No	No	Sí	Sí	No	3	No	No
79	No		Sí	Sí	Sí	No	No	Sí	No	3	No	No
80	No		Sí	Sí	Sí	No	No	Sí	Sí	4	Sí	No
81	No		Sí	Sí	No	No	No	Sí	No	3	No	No
82	No		No	Sí	Sí	No	No	No	Sí	3	No	No
83	Sí	4	Sí	Sí	Sí	No	No	No	Sí	4	Sí	Sí
84	No		No	Sí	Sí	No	No	No	Sí	4	Sí	No
85	No		Sí	Sí	Sí	No	Sí	Sí	No	5	Sí	No
86	Sí	5	Sí	Sí	Sí	No	No	No	No	4	Sí	No
87												
88												
89	Sí	5	No	Sí	Sí	No	No	No	Sí	3	No	No
90	No		Sí	Sí	No	No	No	Sí	No	4	Sí	No
91												
92	No		No	Sí	Sí	No	No	Sí	No	3	No	No
93	No		No	Sí	Sí	No	No	No	Sí	3	No	No
94	No		No	Sí	Sí	No	No	Sí	No	3	Sí	Sí
95	No		No	Sí	Sí	No	No	Sí	Sí	2	No	No
96	No		No	Sí	Sí	No	No	Sí	No	4	No	No
97	No		Sí	Sí	Sí	No	No	No	No	3	No	No
98												
99	No		No	Sí	Sí	No	No	Sí	No	3	No	No
100	No		Sí	Sí	Sí	No	No	Sí	No	3	Sí	No

Nota: Tabla de elaboración propia.