



Facultad de Ciencias Económicas y Empresariales
ICADE

**¿LOS ESTUDIANTES DE COMILLAS
SABEN QUE NO SABEN?
LA PRECISIÓN DE LA PREDICCIÓN Y
EL EFECTO DUNNING-KRUGER**

Autor: María Reguero Cabada

Director: Javier Márquez Vigil

MADRID | Marzo 2025

RESUMEN

La capacidad de predecir el propio rendimiento académico es una habilidad fundamental para los estudiantes, ya que influye directamente en la planificación del estudio, la toma de decisiones y la confianza en uno mismo. Sin embargo, no siempre existe una correspondencia precisa entre las expectativas de los alumnos y sus resultados reales. Este fenómeno puede estar relacionado con el efecto Dunning-Kruger, según el cual los estudiantes con bajo rendimiento tienden a sobreestimar sus capacidades, mientras que los de alto rendimiento tienden a subestimarse.

El presente trabajo tiene como objetivo principal analizar si los estudiantes universitarios son capaces de evaluar con precisión su desempeño académico y si esta capacidad mejora a lo largo del tiempo. Para ello, se ha realizado un estudio empírico con una muestra de estudiantes de alto rendimiento de una asignatura de contabilidad en una universidad madrileña. Durante el cuatrimestre, los estudiantes realizaron nueve predicciones sobre su nota final, lo que permitió un análisis longitudinal de la evolución en la precisión de sus estimaciones.

Los resultados muestran que, al inicio del curso, los estudiantes no predicen con precisión sus calificaciones, aunque esta capacidad mejora con el tiempo. Además, se observa la presencia del efecto Dunning-Kruger y se analiza el posible impacto del género, sin encontrar diferencias significativas. Este estudio busca aportar evidencia sobre cómo los estudiantes perciben su rendimiento y sentar las bases para futuras investigaciones que exploren estrategias para mejorar la autoevaluación académica.

Palabras clave: Autoevaluación, Efecto Dunning-Kruger, Predicción académica, Estudiantes universitarios, Rendimiento académico

ABSTRACT

The ability to predict one's own academic performance is a key skill for students, as it directly affects study planning, decision-making, and self-confidence. However, there is often a mismatch between students' expectations and their actual results. This phenomenon may be related to the Dunning-Kruger effect, which suggests that low-performing students tend to overestimate their abilities, while high-performing students tend to underestimate them.

The main objective of this study is to analyze whether university students are able to accurately assess their academic performance and whether this ability improves over time. To address this, an empirical study was conducted with a sample of high-performing students enrolled in an accounting course at a university in Madrid. Throughout the semester, students made nine predictions of their final grade, allowing a longitudinal analysis of the evolution in the accuracy of their estimates.

The results show that students are not able to accurately predict their grades at the beginning of the course, although this ability improves over time. Additionally, the Dunning-Kruger effect is observed, and the possible influence of gender is examined, with no significant differences found. This study contributes to a better understanding of how students perceive their performance and lays the groundwork for future research on strategies to improve academic self-assessment.

Key words: Self-assessment, Dunning-Kruger effect, Academic prediction, University students, Academic performance

ÍNDICE

| | |
|---|----|
| 1. INTRODUCCIÓN | 6 |
| 2. MARCO TEÓRICO | 8 |
| 2.1. LA AUTOEVALUACIÓN EN LA EDUCACIÓN SUPERIOR | 8 |
| 2.1.1. Introducción, aproximación histórica a la autoevaluación | 8 |
| 2.1.2. Concepto de autoevaluación | 9 |
| 2.1.3. Tipología | 11 |
| 2.1.4. Beneficios de la autoevaluación | 12 |
| 2.1.5. Conclusión | 13 |
| 2.2. LA PREDICCIÓN DEL DESEMPEÑO Y EL EFECTO DUNNING- KRUGER | 14 |
| 3. REVISIÓN DE LA LITERATURA | 17 |
| 4. METODOLOGÍA | 23 |
| 4.1. METODOLOGÍA DEL CURSO | 23 |
| 4.2. METODOLOGÍA ESTADÍSTICA | 25 |
| 5. RESULTADOS Y DISCUSIÓN | 28 |
| 5.1. CAPACIDAD DE PREDICCIÓN Y SU EVOLUCIÓN CON EL TIEMPO | 28 |
| 5.1.1. Regresión en tres momentos clave | 28 |
| 5.1.2. Regresiones individuales para cada test | 30 |
| 5.1.3. Conclusión | 31 |
| 5.2. PREDICCIÓN Y RENDIMIENTO | 32 |
| 5.3. PREDICCIÓN Y GÉNERO | 34 |
| 6. CONCLUSIONES | 37 |
| Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado | 40 |
| BIBLIOGRAFÍA | 41 |
| ANEXO | 46 |

ÍNDICE DE GRÁFICOS Y TABLAS

| | |
|---|----|
| Tabla 1. Descriptivo datos | 24 |
| Gráfico 1. Porcentaje de hombres y mujeres en la muestra | 25 |
| Tabla 2. Resultados del modelo 1 | 27 |
| Tabla 3. Análisis de colinealidad imperfecta grave | 29 |
| Tabla 4. Resultados del modelo 2 | 29 |
| Tabla 5. Resultados del modelo 3 | 31 |
| Gráfico 2. Comparativa evolución del coeficiente en alto y bajo rendimiento | 32 |
| Tabla 6. Resultados del modelo 4 | 33 |
| Tabla 7. Tabla de correlaciones entre notas predichas y nota final | 45 |
| Tabla 8. Descripción de las variables utilizadas en el análisis | 46 |

1. INTRODUCCIÓN

El presente estudio tiene como objetivo analizar la capacidad de predicción del rendimiento académico de los estudiantes y la posible presencia del efecto Dunning-Kruger en este contexto. En el proceso de aprendizaje, los individuos atraviesan diferentes fases. Inicialmente, no son conscientes de lo que desconocen, luego comienzan a reconocer sus limitaciones, posteriormente logran una competencia consciente mediante la práctica constante y, finalmente, adquieren una competencia inconsciente en la que sus habilidades se ejecutan de manera automática. Sin embargo, diversos estudios han demostrado que la percepción del conocimiento y de las propias capacidades no siempre es precisa, estando influenciada por distintas variables. Una de las más relevantes es el efecto Dunning-Kruger, que describe la tendencia de los individuos con menor competencia a sobrestimar sus capacidades, mientras que aquellos con mayor competencia tienden a subestimarse.

Ser capaz de evaluar correctamente el propio desempeño es clave tanto en el ámbito académico como profesional, ya que influye en la toma de decisiones, la planificación del estudio y la confianza en las propias habilidades. Si los estudiantes no logran predecir con precisión sus resultados, pueden caer en una planificación ineficaz o en una autopercepción errónea que afecte negativamente su rendimiento. Aunque se han realizado numerosos estudios sobre este fenómeno, la mayoría se basan en mediciones puntuales, lo que dificulta analizar la evolución de la capacidad de predicción a lo largo del tiempo. Para superar esta limitación, el presente trabajo adopta un enfoque longitudinal, permitiendo observar si la precisión en la autoevaluación mejora conforme avanza el cuatrimestre.

Determinar si la capacidad de predicción mejora con el tiempo y si variables como el rendimiento académico o el género influyen en esta habilidad resulta fundamental para comprender mejor los mecanismos que subyacen a la autopercepción del conocimiento. En este sentido, el estudio busca aportar evidencia empírica sobre cómo estas variables afectan la precisión de las estimaciones de los estudiantes respecto a su propio desempeño. Además, aunque este estudio se centra en el diagnóstico del problema, sería interesante explorar en futuras investigaciones cómo pueden implementarse estrategias pedagógicas para mejorar la autoevaluación de los estudiantes y mitigar los efectos del sesgo cognitivo.

En cuanto a la metodología empleada, se ha realizado un experimento piloto en una asignatura del área de contabilidad en una universidad mediana de Madrid, tomando como

muestra a estudiantes de alto rendimiento, dado que la nota de corte para acceder a esta institución es elevada. El estudio sigue un enfoque cuantitativo basado en el análisis de datos obtenidos a través de nueve formularios aplicados a lo largo del cuatrimestre, en los que los estudiantes predijeron su nota final en la asignatura. Posteriormente, se realizaron análisis mediante regresiones lineales para evaluar la precisión de sus predicciones y su evolución en el tiempo.

Los resultados muestran que los estudiantes no son capaces de predecir con exactitud su calificación en las primeras etapas del curso, aunque su capacidad de predicción mejora con el tiempo. Se confirma la presencia del efecto Dunning-Kruger, evidenciando que los estudiantes con mejor rendimiento tienden a subestimarse, mientras que aquellos con menor rendimiento sobrestiman sus resultados. Asimismo, se ha analizado la influencia del género en esta capacidad de predicción, sin encontrar evidencia concluyente que confirme la hipótesis de que las mujeres subestiman sus habilidades en mayor medida que los hombres.

El trabajo se estructura en diferentes secciones. En primer lugar, se analiza la autoevaluación y la predicción del rendimiento académico, estudiando los factores que influyen en la capacidad de los estudiantes para estimar su desempeño y la forma en que esta habilidad evoluciona con el tiempo. A continuación, se expone una revisión de la literatura en la que se presentan estudios previos que han abordado esta cuestión en distintos contextos. Posteriormente, se describe la metodología utilizada en la investigación, detallando el diseño del experimento, la muestra seleccionada y las herramientas de análisis empleadas. Seguidamente, se presentan los resultados obtenidos y su discusión, donde se analizan los datos y se comparan con los hallazgos de estudios previos. Finalmente, se presentan las conclusiones del estudio, destacando sus aportes y sugiriendo posibles líneas de investigación futura.

2. MARCO TEÓRICO

2.1. LA AUTOEVALUACIÓN EN LA EDUCACIÓN SUPERIOR

La autoevaluación se ha consolidado en las últimas décadas como una herramienta clave en el ámbito universitario, especialmente a partir de los cambios estructurales impulsados por el Espacio Europeo de Educación Superior. Este apartado examina en profundidad el concepto de autoevaluación dentro del contexto de la educación superior, realizando una aproximación histórica a su evolución, analizando sus distintas definiciones teóricas, tipologías y modelos de aplicación, así como los beneficios que ofrece para el desarrollo académico del estudiante. Con ello, se pretende ofrecer un marco conceptual sólido que permita comprender su relevancia en los actuales procesos de enseñanza y aprendizaje.

2.1.1. Introducción, aproximación histórica a la autoevaluación

La autoevaluación, según la Real Academia Española, se define como la "evaluación que alguien hace de sí mismo o de algún aspecto o actividad propios". En el ámbito educativo, este concepto se refiere a la capacidad del estudiante para valorar su propio desempeño y regular su aprendizaje. En este trabajo, nos centraremos en la autoevaluación educacional, analizando la precisión con la que los estudiantes pueden medir su rendimiento académico y la relación entre este proceso y su desempeño en la educación superior.

Para comprender mejor la relevancia actual de la autoevaluación, es fundamental analizar su evolución histórica. A lo largo de los años, la educación superior ha experimentado transformaciones significativas, particularmente a partir de la década de 1990, cuando comenzaron a implementarse cambios estructurales orientados a modernizar la enseñanza universitaria. Entre estos hitos destacan la Declaración de la Sorbona en 1998 y la Declaración de Bolonia en 1999, que impulsaron la armonización de los sistemas educativos en Europa y promovieron un modelo de enseñanza basado en la movilidad de los estudiantes, la estandarización académica, la competitividad y la diversidad.

La Declaración de Bolonia (Ministros Europeos, 1999) sentó las bases para la creación del Espacio Europeo de Educación Superior (EEES), cuyo objetivo principal era garantizar la empleabilidad de los graduados mediante un sistema académico más homogéneo y centrado en el desarrollo de competencias clave. Con este cambio, el modelo de aprendizaje pasó de estar basado en la enseñanza tradicional, centrada en el docente, a un enfoque donde el

estudiante es el principal protagonista de su propio aprendizaje (Galván-Cardoso y Siado-Ramos, 2021). En este nuevo paradigma, la autoevaluación adquiere un papel fundamental, ya que permite a los alumnos desarrollar la autonomía y el pensamiento crítico necesarios para enfrentar con éxito los retos del ámbito laboral y profesional.

Uno de los hitos más relevantes en la consolidación de la autoevaluación como una estrategia clave en la educación superior fue la Recomendación del Parlamento Europeo y el Consejo de 2006 sobre competencias clave para el aprendizaje permanente. Este documento enfatizó la importancia de que los estudiantes desarrollen la capacidad de organizar su propio aprendizaje, evaluar su desempeño y buscar estrategias de mejora de manera autónoma. A partir de esta recomendación, la autoevaluación se convirtió en un componente esencial de los programas educativos, no solo en Europa, sino también en otros sistemas de enseñanza que adoptaron metodologías similares.

Además, con el auge de la digitalización y la educación a distancia, el papel de la autoevaluación ha cobrado aún mayor relevancia en la gestión del aprendizaje. Herramientas tecnológicas como plataformas de e-learning, cuestionarios interactivos y sistemas de retroalimentación han facilitado que los estudiantes puedan evaluar su progreso de manera más efectiva y personalizada. La implementación de estos recursos ha reforzado la importancia de la autoevaluación no solo como un mecanismo de valoración académica, sino también como una estrategia de aprendizaje que fomenta la reflexión crítica y la mejora continua (Herrera y Rojas, 2024).

En resumen, la autoevaluación ha evolucionado de ser una práctica complementaria a convertirse en una herramienta esencial dentro de los sistemas educativos modernos. Su integración en la educación superior responde a la necesidad de formar profesionales autónomos, críticos y adaptables a los cambios constantes del entorno académico y laboral (Muñoz Landázuri et al., 2024).

2.1.2. Concepto de autoevaluación

El concepto de autoevaluación ha sido abordado desde diversas perspectivas a lo largo de los años, dando lugar a diferentes enfoques teóricos y metodológicos. De manera general, la autoevaluación se entiende como la capacidad del individuo para valorar su propio desempeño y reflexionar sobre su proceso de aprendizaje (Calatayud Salom, 2019).

Santos Guerra (1994) define la autoevaluación como un proceso de reflexión, autocrítica y análisis que fomenta hábitos de pensamiento crítico en los estudiantes. Desde otra perspectiva, Boud (1995) la concibe como la participación activa del estudiante en la identificación de estándares y criterios aplicables a su trabajo, permitiéndole juzgar hasta qué punto ha alcanzado dichos estándares.

Boud y Brew (1995) establecen una distinción fundamental entre la reflexión y la autoevaluación. Mientras que la reflexión es un proceso exploratorio que no siempre implica emitir un juicio evaluativo, la autoevaluación se enfoca en la emisión de juicios concretos sobre el aprendizaje alcanzado. Esta diferencia radica en que la autoevaluación requiere la aplicación de criterios objetivos que permitan medir el nivel de logro de los objetivos educativos.

Por otro lado, Panadero Calderón (2011) distingue dos vertientes teóricas dentro de la autoevaluación. La primera es una visión pedagógica alineada con la evaluación formativa, en la que el estudiante recibe retroalimentación constante y ajusta su aprendizaje en función de ella. La segunda perspectiva se basa en la autorregulación del aprendizaje, donde el estudiante asume un rol activo en la planificación, monitoreo y ajuste de su propio proceso educativo.

De acuerdo con estos enfoques, la autoevaluación es una herramienta de gran valor en la educación, ya que permite desarrollar la capacidad de autoobservación y autorreflexión, promoviendo una mayor autonomía en el aprendizaje. Sin embargo, diversos estudios han demostrado que la autoevaluación no se ha integrado de manera generalizada en la educación superior. Investigaciones como las de Álvarez Rojo et al. (2011) en España y Cabrales Salazar (2010) en Colombia han evidenciado que los sistemas de evaluación universitaria siguen anclados en métodos tradicionales, con un predominio de la evaluación sumativa sobre la formativa. Esta situación ha limitado la implementación efectiva de la autoevaluación, generando una brecha entre la teoría y la práctica educativa.

Un punto de debate en la literatura académica es si la autoevaluación debe considerarse un método de evaluación sumativo o formativo. Esta distinción proviene de la clasificación realizada por Scriven en 1967. La evaluación sumativa se realiza al finalizar el proceso de aprendizaje con el propósito de determinar si el estudiante ha alcanzado los objetivos propuestos (Olmos, 2008). En cambio, la evaluación formativa ocurre durante el proceso de

aprendizaje y se enfoca en mejorar tanto el proceso como los resultados, permitiendo ajustes constantes en la enseñanza y el aprendizaje (Olmos, 2008).

Boud y Falchikov (1989) sostienen que la autoevaluación puede ser tanto formativa como sumativa. Su carácter formativo radica en que permite al estudiante identificar sus áreas de mejora y optimizar su aprendizaje de manera continua. A su vez, puede considerarse sumativa cuando contribuye a la calificación final del estudiante. Sin embargo, esta visión ha sido cuestionada por otros autores, quienes argumentan que los estudiantes no siempre son precisos al autoevaluarse y asignarse calificaciones objetivas (Hassan et al., 2014; González-Betancor et al., 2017; Blackwood, 2013). Debido a esta falta de precisión, algunos investigadores sugieren que la autoevaluación no debería utilizarse como un criterio determinante en la evaluación académica final, sino más bien como una herramienta complementaria que favorezca el aprendizaje y la metacognición.

En conclusión, la autoevaluación es una estrategia clave en la educación superior que permite a los estudiantes reflexionar sobre su propio aprendizaje y mejorar su desempeño. Si bien existen diversas posturas sobre su aplicación y utilidad, la evidencia sugiere que su implementación debe estar acompañada de un marco pedagógico adecuado que garantice su efectividad y objetividad en el ámbito académico.

2.1.3. Tipología

Existen diversos modelos de autoevaluación en la educación superior, los cuales varían en función del grado de autonomía y responsabilidad otorgado al estudiante. Taras (2010) identifica cinco modelos principales que van desde aquellos en los que el docente sigue teniendo un papel central hasta aquellos en los que el estudiante asume completamente la gestión de su evaluación.

El modelo de automarcado (Taras, 2010) es el más básico y consiste en que el estudiante compara su trabajo con un modelo de respuesta proporcionado por el profesor, identificando sus errores y asignándose una calificación. Aunque este modelo puede parecer mecánico, fomenta la autoobservación y el reconocimiento de errores, lo que contribuye al aprendizaje.

En el modelo de estándar de sonido (Taras, 2010), se ofrece al estudiante una descripción de un trabajo de calidad media junto con dos ejemplos de trabajos de diferente nivel sin especificar cuál es superior. De este modo, el estudiante debe evaluar su propio desempeño en

relación con estos ejemplos y desarrollar un criterio analítico para juzgar la calidad de su trabajo.

El modelo estándar (Taras, 2010) permite que el estudiante autoevalúe su trabajo antes de entregarlo al profesor, siguiendo ciertos criterios previamente establecidos. Este enfoque le otorga la posibilidad de mejorar su trabajo antes de la evaluación formal, promoviendo la revisión crítica y la mejora continua.

En el modelo de autoevaluación con retroalimentación integrada del tutor (Taras, 2010), los estudiantes reciben comentarios tanto de sus profesores como de sus compañeros antes de emitir su propia autoevaluación. La interacción con otras perspectivas enriquece su análisis y les ayuda a desarrollar una visión más objetiva de su rendimiento.

Finalmente, en el modelo de diseño de contratos de aprendizaje (Taras, 2010), los estudiantes gestionan completamente su proceso de evaluación, estableciendo sus propios criterios y objetivos. En este caso, el profesor solo interviene cuando el estudiante solicita asistencia, otorgándole total autonomía sobre su aprendizaje y evaluación.

Taras (2010) también distingue entre autoevaluación formal e informal. La autoevaluación formal implica que el proceso influye directamente en la calificación final del estudiante, mientras que la informal se enfoca en la reflexión y mejora personal sin que afecte la nota. Este último enfoque permite mayor libertad y autenticidad en la evaluación del estudiante, fomentando el aprendizaje autorregulado sin la presión de una calificación numérica.

Estos enfoques evidencian la importancia de adaptar los modelos de autoevaluación a las necesidades y características de cada entorno educativo. La selección del modelo adecuado depende del nivel de autonomía que se desea fomentar en los estudiantes, así como de los objetivos pedagógicos que se persigan en el proceso de enseñanza-aprendizaje.

2.1.4. Beneficios de la autoevaluación

La autoevaluación fomenta la independencia en el aprendizaje, permitiendo que los estudiantes asuman un rol activo en su formación en lugar de depender exclusivamente de la evaluación del docente. Según Panadero et al. (2015) y Hill (2016), a diferencia de los métodos tradicionales en los que el profesor asigna una calificación sin mayor reflexión por

parte del estudiante, esta estrategia promueve la autoconciencia sobre fortalezas y debilidades, ayudando a mejorar el método de estudio y optimizar el rendimiento académico.

También desarrolla el pensamiento crítico, ya que obliga a los alumnos a analizar con mayor profundidad su desempeño y a realizar juicios fundamentados sobre su progreso. Hill (2016) destaca que este proceso fortalece la capacidad de autorregulación y autoanálisis, permitiendo ajustar expectativas y ser más realistas con los avances. González-Betancor et al. (2017) afirman que esta competencia es fundamental en el mundo profesional actual, ya que potencia la capacidad de cuestionar, evaluar y perfeccionar el propio trabajo de manera objetiva.

Por otro lado, la autoevaluación contribuye a un aprendizaje más eficaz, facilitando una comprensión más profunda de los contenidos, ya que el estudiante se involucra activamente en su proceso educativo. Leach (2012) señala que, al identificar qué aspectos necesitan mejorar, los alumnos pueden ajustar sus estrategias de estudio, lo que se traduce en un aprendizaje más sólido y duradero. Además, Nulty (2011) explica que este proceso permite a los estudiantes ser conscientes de sus fortalezas y debilidades, lo que les ayuda a reorientar su forma de trabajo y obtener mejores resultados en futuras evaluaciones.

En definitiva, la autoevaluación impulsa la autonomía, fomenta el pensamiento crítico y mejora la calidad del aprendizaje, convirtiéndose en una herramienta esencial para la educación actual.

2.1.5. Conclusión

La autoevaluación es una herramienta fundamental en la educación superior, ya que promueve la autonomía, el pensamiento crítico y la mejora continua del aprendizaje. A lo largo del tiempo, ha evolucionado desde una práctica marginal hasta convertirse en un elemento central dentro del Espacio Europeo de Educación Superior. A pesar de los debates sobre su precisión como método de evaluación, su impacto en el desarrollo académico y profesional de los estudiantes es innegable.

2.2. LA PREDICCIÓN DEL DESEMPEÑO Y EL EFECTO DUNNING- KRUGER

Una vez contextualizada la autoevaluación desde un enfoque histórico y teórico, así como sus tipologías y beneficios en el ámbito educativo, en este apartado se profundizará en la

predicción del desempeño y el efecto Dunning-Kruger que esta directamente relacionado con este fenómeno.

Este trabajo analiza la capacidad de los alumnos para predecir su propio desempeño y evalúa la precisión de sus autoevaluaciones. La predicción o autoevaluación del desempeño es un proceso en el que los estudiantes, antes de realizar una tarea o examen, intentan anticipar su resultado en función de sus expectativas de dificultad, sus experiencias previas y ciertos rasgos de su carácter.

A lo largo de la literatura, se han identificado diversos factores que afectan la precisión de esta predicción. En particular, un fenómeno ampliamente estudiado que explica la imprecisión en la autoevaluación de los estudiantes es el efecto Dunning-Kruger, el cual impacta directamente en la capacidad de predecir con exactitud el propio rendimiento.

El efecto Dunning-Kruger describe la tendencia de las personas con bajo rendimiento a sobrestimar sus propias habilidades. Según Dunning (2011) y Kruger y Dunning (1999), esta sobreestimación no es homogénea en toda la población, sino que ocurre principalmente en los individuos menos informados y con menor capacidad real. Este fenómeno se explica a través de una doble maldición: los individuos con bajo rendimiento no solo tienen dificultades para ejecutar una tarea de manera correcta, sino que además carecen de la capacidad para reconocer su bajo desempeño.

La raíz del efecto Dunning-Kruger radica en que las mismas habilidades necesarias para desempeñar bien una tarea son las que permiten evaluar con precisión el propio rendimiento. Como explican Kruger y Dunning (1999), “se necesita habilidad para ver la falta de habilidad”, lo que significa que los individuos menos competentes no pueden identificar sus errores ni comparar de manera objetiva su desempeño con el de sus pares. Como resultado, creen que su rendimiento es mucho mejor de lo que realmente es.

Los datos empíricos han respaldado esta hipótesis en múltiples estudios, reflejando un patrón consistente de sobreestimación en diferentes contextos del mundo real. Por ejemplo, Dunning et al. (2003) encontraron que los estudiantes con peor rendimiento en exámenes académicos sobreestimaban drásticamente sus puntuaciones, tanto en términos absolutos como en comparación con sus compañeros. Ehrlinger et al. (2008) y Park y Santos-Pinto (2010) replicaron estos hallazgos en contextos como torneos de debate o de ajedrez y Mynttinen et al. (2009) y Hodges et al. (2001) en el contexto de conductores que no aprueban su examen

de conducir y y estudiantes de medicina que llevan a cabo una entrevista clínica con un paciente estandarizado en una simulación de comunicación de malas noticias, respectivamente.

Mientras que los individuos con menor rendimiento tienden a sobreestimarse, aquellos con alto desempeño experimentan un sesgo opuesto: aunque juzgan con precisión su propio desempeño, lo subestiman en comparación con los demás. Este fenómeno se debe al efecto de consenso falso (Ross et al., 1977), por el cual las personas asumen erróneamente que los demás poseen habilidades similares a las suyas. Como resultado, los más competentes sobreestiman el rendimiento de los demás y, al no percibirse como excepcionalmente hábiles, subestiman lo especial o único que es su desempeño dentro del grupo (Kruger y Dunning, 1999; Ehrlinger et al., 2008).

En este sentido, mientras que los errores de autoevaluación en los individuos con bajo desempeño se reflejan en una sobreestimación de sus propias habilidades, en los individuos con alto rendimiento los errores se centran en la percepción de las capacidades ajenas, llevándolos a subestimar su ventaja comparativa (Kruger y Dunning, Estudio 4; Hodges et al., 2001).

Se podría pensar que la experiencia ayuda a mejorar la autoevaluación y a reducir el efecto Dunning-Kruger. Sin embargo, los estudios presentan resultados contradictorios. Algunos investigadores han sugerido que, con el tiempo, las personas aprenden a evaluar mejor sus habilidades y reducen su sesgo de autoevaluación (Krajč y Ortman, 2008). En particular, se ha encontrado que el efecto Dunning-Kruger puede atenuarse con la experiencia, pero para que esto ocurra, se requiere retroalimentación estructurada e incentivos que guíen la corrección de la autoevaluación (Miller y Geraci, 2011).

No obstante, la evidencia sugiere que, aunque la experiencia y la retroalimentación pueden mejorar la autoevaluación, no siempre son suficientes para corregir completamente el sesgo Dunning-Kruger (Hacker et al., 2000; Ferraro, 2010). Esto se debe a que la habilidad metacognitiva necesaria para evaluar el propio desempeño no siempre se desarrolla únicamente con la repetición de tareas.

Por todo esto, el efecto Dunning-Kruger demuestra que las personas menos capacitadas no solo son incompetentes, sino que además no pueden reconocer su incompetencia. Esto se debe a que las mismas habilidades necesarias para rendir bien son las que permiten evaluar

correctamente el propio desempeño. En contraste, los individuos más competentes tienden a subestimarse debido al efecto de consenso falso, ya que creen erróneamente que su desempeño es más común de lo que realmente es. Finalmente, la experiencia por sí sola no corrige este sesgo, ya que la autoevaluación requiere de una habilidad metacognitiva que las personas con bajo rendimiento inicialmente no poseen.

3. REVISIÓN DE LA LITERATURA

La autoevaluación en la educación superior ha sido objeto de múltiples estudios que buscan comprender su precisión y los factores que la afectan. La diversidad de enfoques en la investigación refleja la complejidad del fenómeno, ya que la capacidad de los estudiantes para evaluar su propio desempeño está influenciada por diversas variables, como el rendimiento académico, el género, la familiaridad con la tarea evaluada y el impacto de la autoevaluación en la calificación final. A partir de la revisión de nueve estudios desarrollados en diferentes contextos educativos y geográficos, se pueden identificar patrones comunes, contradicciones en los hallazgos y enfoques metodológicos diferenciados que han permitido analizar este fenómeno desde distintas perspectivas. En este sentido, la presente investigación parte de la hipótesis de que los estudiantes no son capaces de predecir con precisión su calificación final, especialmente en etapas tempranas del proceso de aprendizaje.

Los estudios revisados han sido realizados en distintos países, incluyendo España (González-Betancor et al., 2017; Bolívar-Cruz et al., 2018; Tejeiro et al., 2012; Gómez-Puerta et al., 2019), Hungría (Máté y Darabos, 2017), Colombia (Cuesta-Melo et al., 2022), Estados Unidos (Fitzgerald et al., 2003), Dinamarca (Pallisgaard Støve, 2020) y Singapur y Países Bajos (Lew et al., 2010). A pesar de la diversidad geográfica y cultural, estos estudios coinciden en investigar la precisión de la autoevaluación en educación superior, aunque con enfoques y metodologías diferentes.

En cuanto al enfoque metodológico, la gran mayoría de los estudios han utilizado cuestionarios y encuestas de autoevaluación, donde los estudiantes debían estimar su rendimiento antes o después de una prueba académica. Estas autoevaluaciones fueron comparadas con las calificaciones reales otorgadas por los profesores o, en algunos casos, con la evaluación de pares. Este enfoque ha sido predominante en estudios como el de González-Betancor et al. (2017) en España, que evaluó la autoevaluación en presentaciones orales, Máté y Darabos (2017) en Hungría, que comparó la precisión de la autoevaluación en exámenes digitales y en papel, y Gómez-Puerta et al. (2019), quienes aplicaron una metodología similar al analizar la relación entre las predicciones de los estudiantes sobre sus calificaciones y sus resultados reales.

No obstante, algunos estudios han empleado enfoques cualitativos para obtener una visión más profunda del proceso de autoevaluación. En el estudio realizado en Colombia (Cuesta-Melo et al., 2022), se utilizó un diseño cualitativo-descriptivo, basado en diarios reflexivos, entrevistas y discusiones grupales, con el objetivo de analizar cómo los estudiantes perciben la autoevaluación y qué factores subjetivos influyen en su precisión. Este enfoque permitió identificar creencias y actitudes hacia la autoevaluación que los estudios cuantitativos no pueden captar con la misma profundidad.

A continuación, se analizarán los principales factores que influyen en la precisión de la autoevaluación, como el género, el rendimiento académico, la familiaridad con la tarea evaluada y el impacto de la autoevaluación en la calificación final. Si bien ciertos patrones han sido identificados en la literatura, los resultados no siempre son concluyentes, lo que refleja la complejidad del fenómeno.

Uno de los factores más recurrentes en la literatura sobre autoevaluación es el género, ya que diversas teorías y estudios previos han sugerido que hombres y mujeres pueden diferir en su percepción del propio rendimiento. En muchos estudios, se ha observado la tendencia de que los hombres tienden a sobrestimar sus capacidades, mientras que las mujeres suelen subestimarse. Estas diferencias podrían estar relacionadas con factores como la confianza en sí mismos, la socialización de género y la percepción del éxito y el fracaso (González-Betancor et al., 2017).

Por ejemplo, algunos estudios en psicología educativa han sugerido que las mujeres suelen ser más autocríticas en la evaluación de sus propias habilidades, lo que podría explicar por qué en algunos estudios se reporta que ellas se autoevalúan por debajo de su rendimiento real. Esto podría estar vinculado al llamado "síndrome del impostor", que sugiere que algunas personas, especialmente mujeres, tienden a dudar de su competencia incluso cuando tienen buenos resultados (Quevedo Lezama, 2023).

Esta tendencia se observa en el estudio de González-Betancor et al. (2017) en España, donde los hombres sobrestimaban su desempeño en presentaciones orales, mientras que las mujeres tendían a evaluarse por debajo de sus capacidades reales. En el estudio sobre presentaciones orales en España, también se identificó que las mujeres basaban su autoevaluación en su percepción de autoeficacia, mientras que los hombres eran más influenciados por incentivos

externos, lo que refuerza la idea de que las mujeres pueden ser más conservadoras al evaluarse.

No obstante, esta tendencia no es universal y varios estudios han reportado que no existen diferencias significativas entre hombres y mujeres en la precisión de la autoevaluación. En el artículo de Maté y Darabos (2017) y en el estudio con estudiantes de fisioterapia en Dinamarca (Pallisgaard Støve, 2020), no se encontraron diferencias de género en la precisión de la autoevaluación. Esto plantea la posibilidad de que el impacto del género varíe según el tipo de tarea evaluada.

En tareas más subjetivas, como presentaciones orales o trabajos de escritura, las diferencias de género pueden ser más notorias, ya que la evaluación del propio desempeño depende en parte de la confianza en sí mismo. Sin embargo, en pruebas más objetivas y técnicas, como exámenes de anatomía o matemáticas, el género puede no desempeñar un papel significativo en la autoevaluación, ya que las respuestas correctas o incorrectas son más evidentes. Estas diferencias encontradas en la literatura conducen a plantear la hipótesis de que el género influye en la predicción del propio desempeño, siendo las mujeres más propensas a subestimar sus capacidades y los hombres más tendentes a sobreestimarlas.

Otro factor ampliamente analizado ha sido la relación entre rendimiento académico y precisión en la autoevaluación. La hipótesis predominante en la literatura es que los estudiantes con mejor desempeño académico tienden a ser más precisos en su autoevaluación, ya que poseen un mayor nivel de autoconciencia sobre sus fortalezas y debilidades.

En la Universidad de Debrecen, Hungría (Maté y Darabos, 2017) y en Republic Polytechnic, Singapur (Lew et al., 2010), se encontró que los estudiantes con notas más altas eran menos propensos a sobrevalorarse y realizaban autoevaluaciones más precisas. Esto se debe a que estos estudiantes suelen tener más experiencia en procesos de evaluación y poseen mejores habilidades metacognitivas, lo que les permite evaluar con mayor realismo su desempeño.

Estos hallazgos están estrechamente relacionados con el efecto Dunning-Kruger, según el cual los estudiantes con bajo rendimiento académico no solo obtienen peores resultados, sino que también sobreestiman su desempeño de manera significativa. Esto ocurre porque carecen de la habilidad necesaria para evaluar correctamente sus propias competencias, lo que genera una percepción errónea de sus capacidades.

En esta línea, el estudio de Gómez-Puerta et al. (2019) confirmó que los estudiantes con menor rendimiento tendían a sobrestimar sus calificaciones esperadas, mientras que aquellos con mejor desempeño eran más precisos en sus predicciones o incluso llegaban a subestimarse levemente. Este estudio respalda la idea de que la precisión en la autoevaluación mejora a medida que aumenta el rendimiento académico, en parte debido al desarrollo de mayores competencias metacognitivas.

Este patrón observado respalda una de las hipótesis principales del presente trabajo: el rendimiento académico influye en la precisión de la predicción, de modo que los estudiantes con bajo rendimiento tienden a sobrestimarse, mientras que aquellos con alto rendimiento tienden a infraestimarse, en línea con el efecto Dunning-Kruger.

Sin embargo, el estudio longitudinal de la Universidad de Michigan (Fitzgerald et al., 2003) mostró que esta relación no es uniforme y varía según el tipo de evaluación. En este caso, los estudiantes fueron más precisos en exámenes teóricos, pero tendieron a sobrestimar su desempeño en pruebas clínicas OSCE. Esto sugiere que la familiaridad con la tarea evaluada es un factor determinante en la precisión de la autoevaluación.

Este hallazgo es clave, ya que indica que la autoevaluación no es una habilidad fija, sino que está influenciada por la experiencia previa, el contexto de evaluación y el nivel de competencias metacognitivas. Esto refuerza la necesidad de entrenar a los estudiantes en autoevaluación desde etapas tempranas para que puedan mejorar progresivamente su precisión y evitar sesgos como el efecto Dunning-Kruger. A partir de estos planteamientos, se formula una última hipótesis: la capacidad de predicción de los estudiantes mejora progresivamente a lo largo del cuatrimestre, a medida que acumulan experiencia y familiaridad con la asignatura.

Por otro lado, algunos estudios han comparado la autoevaluación en distintos escenarios para analizar el impacto de factores externos. En la Universidad de Las Palmas de Gran Canaria (España), se investigó si la precisión de la autoevaluación variaba cuando esta influía en la calificación final (Bolívar-Cruz et al., 2018). Se encontró que los estudiantes tendían a sobrevalorarse más cuando sabían que su autoevaluación afectaría su nota, lo que sugiere que la autoevaluación puede volverse menos precisa cuando está vinculada a incentivos externos.

Otro estudio innovador es el de la Universidad de Debrecen (Maté y Darabos, 2017), que analizó si la precisión de la autoevaluación variaba dependiendo del formato del examen

(digital o en papel). Se encontró que los exámenes digitales en Moodle permitían una autoevaluación más precisa, probablemente debido a la retroalimentación inmediata, lo que facilitaba una mejor calibración del desempeño percibido por los estudiantes.

Otro enfoque diferencial es el del estudio longitudinal de la Universidad de Michigan (Fitzgerald et al., 2003), que realizó mediciones repetidas de la autoevaluación durante tres años, permitiendo evaluar si la precisión mejoraba con el tiempo o si permanecía constante. Los resultados de este estudio fueron clave para demostrar que la precisión de la autoevaluación no mejora automáticamente con la práctica, sino que depende en gran medida de la familiaridad del estudiante con la tarea evaluada.

A pesar de los avances en la investigación sobre autoevaluación, una de las principales limitaciones identificadas es la escasez de estudios longitudinales. La mayoría de las investigaciones han adoptado un enfoque transversal, evaluando la autoevaluación en un único momento sin considerar su evolución a lo largo del tiempo. Si bien estos estudios permiten identificar correlaciones entre factores como género, rendimiento académico y precisión en la autoevaluación, no pueden determinar cómo estos factores cambian a medida que los estudiantes adquieren más experiencia en autoevaluarse.

El único estudio longitudinal en la revisión fue el realizado en la Universidad de Michigan (Fitzgerald et al., 2003), el cual analizó la evolución de la autoevaluación en estudiantes de medicina a lo largo de tres años. Los resultados mostraron que la precisión de la autoevaluación se mantenía relativamente estable durante los primeros dos años, pero disminuía en el tercer año, cuando los estudiantes enfrentaban nuevas formas de evaluación más complejas, como los exámenes clínicos OSCE. Este hallazgo es particularmente relevante, ya que sugiere que la autoeficacia en la autoevaluación no se desarrolla de manera automática con la experiencia, sino que puede verse afectada por la familiaridad con la tarea evaluada.

Esto plantea interrogantes sobre si la autoevaluación puede mejorarse con entrenamiento o si su precisión depende de factores contextuales. Por ejemplo, sería útil investigar si los estudiantes que reciben formación específica en autoevaluación desde el inicio de su formación universitaria muestran una evolución diferente en su precisión a lo largo del tiempo en comparación con aquellos que no reciben entrenamiento formal.

Además, no se han realizado estudios longitudinales que examinen cómo las diferencias de género en la autoevaluación evolucionan con el tiempo. Dado que algunos estudios han encontrado que las mujeres tienden a subestimarse y los hombres a sobreestimarse, sería interesante evaluar si estas diferencias se reducen, aumentan o desaparecen a medida que los estudiantes avanzan en sus estudios y adquieren más confianza en sus habilidades.

Otra posible línea de investigación longitudinal podría centrarse en diferencias entre disciplinas. No está claro si la evolución de la autoevaluación es similar en carreras como ciencias, ingeniería o medicina, donde las evaluaciones suelen ser más objetivas, en comparación con disciplinas como humanidades o artes, donde la autoevaluación puede depender en mayor medida de juicios subjetivos.

Finalmente, es importante considerar el impacto de la digitalización en la precisión de la autoevaluación a lo largo del tiempo. En la Universidad de Debrecen (Maté y Darabos, 2017), se encontró que los exámenes digitales en Moodle mejoraban la precisión de la autoevaluación, lo que plantea la posibilidad de que las plataformas tecnológicas ayuden a los estudiantes a mejorar su autoevaluación de manera más progresiva si se integran de forma sistemática en la enseñanza.

En conclusión, la falta de estudios longitudinales limita la comprensión de la evolución de la autoevaluación y plantea la necesidad de futuras investigaciones que analicen cómo factores como el entrenamiento, la experiencia acumulada y la tecnología pueden influir en la precisión de la autoevaluación en distintos contextos educativos. En particular, el efecto Dunning-Kruger resalta la importancia de estudiar cómo los estudiantes con menor rendimiento pueden mejorar su capacidad de autoevaluarse a lo largo del tiempo. Diseñar estudios longitudinales que analicen estos aspectos permitiría no solo comprender mejor la evolución de la precisión en la autoevaluación, sino también desarrollar estrategias más efectivas para reducir este sesgo y fortalecer la autoevaluación como herramienta de aprendizaje en educación superior.

4. METODOLOGÍA

En esta sección se describe la metodología utilizada en el estudio, abordando tres aspectos fundamentales. En primer lugar, se explica el sistema de evaluación seguido en el curso, detallando los criterios empleados por el profesor para calificar a los estudiantes y la estructura de la evaluación continua. A continuación, se expone el procedimiento de recopilación de datos, especificando cómo se han obtenido las predicciones de calificaciones realizadas por los alumnos a lo largo del semestre. Finalmente, se presenta la metodología estadística empleada para el análisis de los datos, incluyendo la clasificación de los estudiantes según su desempeño académico y la aplicación del modelo de mínimos cuadrados ordinarios (MCO) mediante regresiones lineales para evaluar la relación entre la autoevaluación de los alumnos y su rendimiento real.

4.1. METODOLOGÍA DEL CURSO

El estudio se llevó a cabo con alumnos del grado en Administración y Dirección de Empresas en la Universidad Pontificia Comillas, matriculados en una asignatura del área contable durante el curso académico 2024-2025. Esta asignatura se impartió en el primer semestre, con una duración de 14 semanas. La metodología docente combinó dos sesiones presenciales de dos horas por semana con dos horas adicionales de tutoría voluntaria, disponibles a solicitud de los estudiantes.

La calificación final de esta asignatura se calcula por la siguiente fórmula:

$$CF = 60\% EF + 20\% EI + 20\% EC$$

donde el examen final (EF) representa el 60% de la calificación final (CF), mientras que el examen intermedio (EI) y la evaluación continua (EC) representan el 20% y el 20%, respectivamente.

La evaluación continua se mide y analiza semanalmente a través del sistema Work in class assessment tool (WICAT) (Márquez Vigil, et al. 2023) diseñado por el equipo docente de Contabilidad. Este sistema tiene en cuenta tres dimensiones principales: asistencia, participación en clase y Breves Tareas Frecuentes (BTF).

Desde el inicio del curso, cada estudiante comienza con una puntuación base de 50 sobre 100 en la evaluación continua, lo que equivale a un aprobado. La evolución de esta calificación

depende de su desempeño a lo largo del semestre. La asistencia y la participación en clase permiten sumar puntos, mientras que las faltas de asistencia, el uso indebido de dispositivos electrónicos o cualquier interrupción en el aula pueden restar puntos.

La participación en clase se valora tanto por intervenciones voluntarias, como responder preguntas del profesor o aportar a un debate, como por intervenciones no voluntarias, cuando el profesor selecciona a un estudiante para resolver un ejercicio. Al final del semestre, la calificación en este apartado puede llegar hasta un máximo de 100 puntos, e incluso superar esa cifra en casos excepcionales. Posteriormente, esta puntuación se ajusta a una escala sobre 10 para su integración en la evaluación continua.

Además de la asistencia y participación, otro factor clave en la evaluación continua es la realización de Breves Tareas Frecuentes (BTF). Estas tareas incluyen diferentes actividades diseñadas para reforzar el aprendizaje y fomentar la implicación del estudiante. Entre ellas, se encuentran los cuestionarios semanales, que pueden ser individuales o en pareja, y permiten o no el uso de material de clase según la dinámica establecida. Estos cuestionarios incluyen un 80 % de preguntas sobre temas previamente estudiados y un 20 % sobre el nuevo contenido, con el objetivo de consolidar el aprendizaje y fomentar la preparación previa de la materia. A lo largo del curso, se realizaron 9 cuestionarios.

En estos cuestionarios, se incorpora una pregunta específica en la que los alumnos deben predecir cuál será su calificación final en la asignatura. Esta predicción se recoge en cada cuestionario a lo largo del semestre, lo que permite analizar la evolución de la autoevaluación de los estudiantes en relación con su desempeño real.

Otras actividades incluidas en las Breves Tareas Frecuentes son el Escape Room en Excel, basado en la resolución de problemas contables mediante desbloqueo de hojas de cálculo, y el trabajo en grupo sobre las cuentas anuales de una empresa real. En este último, los estudiantes deben analizar la información financiera de una empresa extraída de la base de datos SABI y responder preguntas relacionadas con indicadores clave como el Fondo de Maniobra, el ROE y la posición de equilibrio financiero.

Al final del semestre, la puntuación obtenida en las Breves Tareas Frecuentes se transforma en una escala de -2 a +2 puntos, los cuales se suman o restan a la calificación obtenida en la asistencia y participación en clase. De este modo, los estudiantes que no participan activamente en clase y tienen un rendimiento deficiente en las tareas solo pueden obtener

puntuaciones en función de los exámenes, lo que limita su evaluación continua a valores cercanos a 5 sobre 10. En contraste, los estudiantes con alto rendimiento pueden mejorar su calificación hasta 9 o 10, e incluso superar esta puntuación en algunos casos.

Este sistema de evaluación permite medir no solo el conocimiento adquirido, sino también la implicación y el esfuerzo de los estudiantes a lo largo del semestre. Como se mencionó anteriormente, la recopilación de datos sobre la autoevaluación de los alumnos se realiza mediante la pregunta incluida en los cuestionarios semanales, en la que se les solicita que estimen su calificación final en la asignatura.

4.2. METODOLOGÍA ESTADÍSTICA

Tal como se indicó en la sección de metodología del curso, durante el semestre se realizaron nueve cuestionarios en los que se evaluaron los conocimientos de los estudiantes. Además de las preguntas sobre la materia, cada cuestionario incluía un apartado en el que los alumnos debían predecir su calificación final en la asignatura. De esta manera, se obtuvo información clave sobre su capacidad de autoevaluación.

La muestra analizada en este estudio está compuesta por 46 alumnos, de los cuales 27 son hombres y 19 son mujeres. Para cada estudiante se recopilaron los siguientes datos: género, calificación del examen intersemestral, calificación del examen final, calificación final y las notas predichas en cada uno de los cuestionarios realizados a lo largo del semestre.

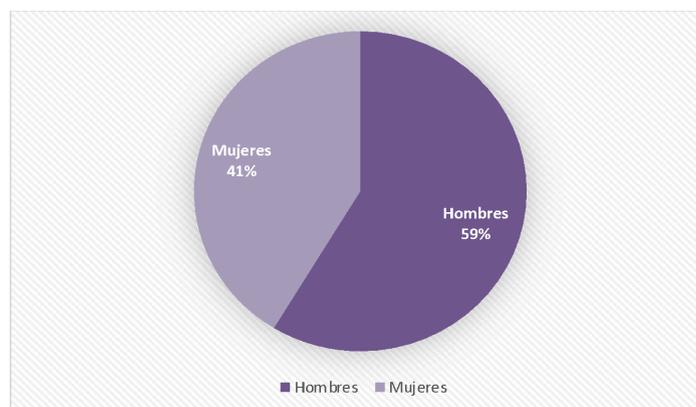
Tabla 1. Descriptivo datos

| n=46 | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 | Test 7 | Test 8 | Test 9 | Nota Final |
|-------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|------------|
| Media | 8.42 | 8.05 | 7.90 | 7.81 | 7.69 | 7.60 | 7.47 | 7.51 | 7.51 | 5.93 |
| Desviación típica | 0.97 | 0.96 | 0.93 | 1.01 | 1.00 | 1.01 | 0.96 | 1.12 | 1.14 | 1.62 |

Nota: Aunque la muestra total está compuesta por 46 estudiantes, el número de participantes varió en cada autoevaluación, ya que no todos asistieron a todas las sesiones.

Gráfico 1. Porcentaje de hombres y mujeres en la muestra

Fuente: Elaboración propia



A partir de estos datos, se ha realizado una clasificación de los estudiantes en función de su desempeño académico. Para ello, las calificaciones finales se han dividido en tres cuartiles, lo que ha permitido identificar dos grupos de especial interés. El primer grupo corresponde a los estudiantes de alto desempeño, aquellos cuya calificación final se sitúa por encima del tercer cuartil, es decir, en el rango superior de la distribución de notas. Por otro lado, el segundo grupo está compuesto por los estudiantes de bajo desempeño, cuyas calificaciones finales se encuentran por debajo del primer cuartil, lo que indica un rendimiento académico significativamente inferior al de sus compañeros.

Además, se ha definido una nueva variable denominada “diferencia test n”, que representa la diferencia entre la calificación predicha en el cuestionario correspondiente y la calificación final obtenida en la asignatura. Se ha creado una variable específica de este tipo para cada uno de los nueve cuestionarios realizados durante el semestre.

Para el análisis de los datos obtenidos, se ha trabajado con el modelo de mínimos cuadrados ordinarios (MCO) mediante regresiones lineales. Este modelo es una de las técnicas estadísticas más utilizadas en el análisis de relaciones entre variables cuantitativas. Su objetivo principal es encontrar la línea de mejor ajuste que minimice la suma de los errores cuadráticos entre los valores observados y los valores estimados por el modelo, permitiendo así realizar predicciones con base en los datos disponibles.

Matemáticamente, el modelo de regresión lineal simple se expresa como:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

donde Y_i representa la variable dependiente, es decir, aquella que se busca predecir o explicar. Por otro lado, X_i es la variable independiente, la cual se emplea para realizar la predicción. El parámetro β_0 , conocido como intercepto, indica el valor esperado de Y cuando X es igual a cero, reflejando así el punto en el que la línea de regresión corta el eje vertical. El coeficiente β_1 mide la relación entre ambas variables y señala cuánto cambia Y cuando X aumenta en una unidad. Finalmente, ε_i representa el término de error, el cual recoge la variabilidad en la variable dependiente que no puede ser explicada por la variable independiente.

Este modelo permite analizar la influencia de la variable independiente sobre la variable dependiente y determinar si existe una relación significativa entre ambas. En este estudio, el modelo de mínimos cuadrados ordinarios ha sido utilizado para evaluar la relación entre la autoevaluación de los estudiantes y su desempeño real, permitiendo analizar si los estudiantes son capaces de predecir su calificación final con precisión y si esta capacidad mejora con el tiempo. Además, facilita el análisis de sesgos sistemáticos en la autoevaluación a lo largo del semestre, como la tendencia de algunos estudiantes a sobrestimar o subestimar su desempeño.

5. RESULTADOS Y DISCUSIÓN

En esta sección se presentan los resultados del análisis cuantitativo con el fin de contrastar las hipótesis formuladas en la revisión de la literatura. En concreto, se parte de las siguientes afirmaciones: (H1) los estudiantes no son capaces de predecir con precisión su calificación final; (H2) la capacidad de predicción mejora progresivamente a lo largo del cuatrimestre; (H3) el rendimiento académico influye en la precisión de la predicción, de modo que los estudiantes con bajo rendimiento tienden a sobreestimarse y los de alto rendimiento a subestimarse; y (H4) el género influye en la autoevaluación, siendo las mujeres más propensas a subestimar su desempeño.

5.1. CAPACIDAD DE PREDICCIÓN Y SU EVOLUCIÓN CON EL TIEMPO

Para dar respuesta a las dos primeras hipótesis, se han desarrollado dos modelos distintos. En primer lugar, se realiza una regresión lineal en tres momentos clave del semestre para observar diferencias generales en la capacidad de predicción. Posteriormente, se llevan a cabo nueve regresiones, una por cada test realizado, con el fin de analizar su evolución de forma más precisa y detallada.

5.1.1. Regresión en tres momentos clave

Se ha utilizado un modelo de mínimos cuadrados ordinarios en el que la variable dependiente es la nota final, mientras que las variables independientes corresponden a las predicciones de los estudiantes sobre su calificación en tres momentos clave del semestre:

- Test 1: Primera semana del curso.
- Test 4: Mitad del semestre.
- Test 9: Última semana antes del examen final.

Tabla 2. Resultados del modelo 1

| | Coefficiente | Desviación típica | Estadístico t | Valor p |
|--------|--------------|-------------------|---------------|---------|
| Const | 1.993 | 1.897 | 1.051 | 0.301 |
| Test 1 | 0.034 | 0.290 | 0.118 | 0.907 |

| | | | | |
|--------|--------|-------|--------|----------|
| Test 4 | -0.827 | 0.356 | -2.321 | 0.027 |
| Test 9 | 1.348 | 0.284 | 4.752 | 4.09e-05 |

| | | | |
|------------|-------|------------------|---------|
| R-cuadrado | 0.432 | R-cuadrado corre | 0.379 |
| F (3, 32) | 8.128 | Valor p (de F) | 0.00036 |

Para que una variable sea considerada significativa, su valor p debe ser inferior a 0.05. Si el valor p es mayor, la variable no tiene un impacto estadísticamente relevante en la nota final.

El primer test (Test 1) no resulta significativo (p-valor = 0.907), lo que indica que al inicio del semestre los estudiantes no son capaces de predecir con precisión su calificación final.

En la mitad del semestre (Test 4), el p-valor (0.027) cae por debajo del umbral de 0.05, lo que parecería que los estudiantes comienzan a mejorar en su capacidad de predicción, aunque con cierta incertidumbre. Sin embargo, el signo negativo del coeficiente indica justamente lo contrario: a medida que aumenta la nota predicha por el alumno (Test 4) disminuye la nota final. Por tanto, la situación parece aún peor que en el test 1. No obstante, cuando se lleva a cabo un análisis individual para cada test (véase tabla 4) observamos que esta variable deja de ser significativa. Esto indica que el resultado obtenido es una significancia espuria, probablemente debido al tamaño muestral. Debido a la presencia de valores faltantes, la estimación de este modelo exige la eliminación de 11 registros. Sin embargo, al realizar la regresión solo con la variable test 4, la muestra es mayor, y el coeficiente deja de ser significativo. La conclusión, por tanto, es que debemos considerar que esta variable realmente no es relevante: en este momento del curso, los alumnos no son capaces de predecir adecuadamente su nota.

En la última semana (Test 9), el p-valor está muy por debajo de 0.05(4.09e-05), y el coeficiente es positivo, lo que confirma que, en este punto, los estudiantes pueden predecir en cierta medida la nota que obtendrán.

El R-cuadrado del modelo es 0.432, con un R-cuadrado ajustado de 0.379, lo que indica que el modelo explica aproximadamente el 43.2% de la variabilidad en la nota final. Además, el

contraste de significatividad global (prueba F) arroja un p-valor de 0.000365, que al estar por debajo de 0.05 confirma que el modelo es globalmente significativo y contiene al menos una variable relevante.

Tabla 3. Análisis de colinealidad imperfecta grave

| | |
|--------|-------|
| Test 1 | 1.877 |
| Test 4 | 3.030 |
| Test 9 | 2.221 |

En relación con la posible colinealidad imperfecta grave, dado que todos los VIF (factores de inflación de la varianza) están por debajo de 5, diremos que el modelo no tiene un problema de colinealidad imperfecta grave.

Por último, la prueba de contraste de heterocedasticidad de White arroja un p-valor de 0.647, lo que indica que no hay evidencia de heterocedasticidad en el modelo al estar este por encima de 0.05. Por lo tanto, el supuesto de homocedasticidad se cumple.

5.1.2. Regresiones individuales para cada test

Para evaluar mejor la evolución de la precisión en la predicción de los estudiantes, se realizaron nueve modelos de mínimos cuadrados ordinarios (MCO) por separado. En cada uno de estos modelos, la variable dependiente sigue siendo la nota final, mientras que la variable independiente corresponde a la predicción realizada en cada test.

Tabla 4. Resultados del modelo 2

| | Coeficiente | Valor p | R-cuadrado |
|--------|-------------|---------|------------|
| Test 1 | 0.247 | 0.352 | 0.022 |
| Test 2 | 0.207 | 0.439 | 0.015 |
| Test 3 | 0.461 | 0.086 | 0.076 |

| | | | |
|---------------|--------------|---------------|--------------|
| Test 4 | 0.221 | 0.353 | 0.021 |
| Test 5 | 0.511 | 0.045 | 0.104 |
| Test 6 | 0.346 | 0.170 | 0.049 |
| Test 7 | 0.653 | 0.049 | 0.140 |
| Test 8 | 0.347 | 0.121 | 0.058 |
| Test 9 | 0.767 | 0.0003 | 0.298 |

Como se observa en la tabla, el p-valor supera el umbral de 0.05 en la mayoría de las variables, lo que indica que no son estadísticamente significativas. Se han resaltado en negrita aquellas que sí lo son.

En el caso de los test 5 y 7, su p-valor está por debajo de 0.05, lo que indica que son estadísticamente significativas. Sin embargo, dado que el valor es apenas ligeramente inferior al umbral, matemáticamente se pueden considerar significativas, pero una mínima variación en los datos podría hacer que dejaran de serlo.

Por otro lado, la variable más significativa es la del test 9, cuyo p-valor está muy por debajo de 0.05, lo que confirma que solo en este momento los estudiantes logran predecir con cierta precisión su nota final.

5.1.3. Conclusión

Ambos análisis muestran que, durante todo el semestre, los estudiantes no tienen una capacidad significativa para predecir su nota final. A medida que avanza el curso, su capacidad de predicción mejora levemente, alcanzando su punto máximo en la última semana antes del examen final, el único momento en que parece que las predicciones sí están significativamente asociadas con la nota final. Estos resultados permiten corroborar las dos primeras hipótesis.

Estos resultados son consistentes con algunos estudios previos, como el de Fitzgerald et al. (2003), que subrayan la importancia de la familiaridad con la tarea y el contexto de

evaluación. Mientras que dicho estudio mostró que la precisión en la autoevaluación no mejora automáticamente con el tiempo, nuestros resultados sugieren que, al menos en el contexto de esta asignatura, la experiencia acumulada sí contribuye a una mayor exactitud, especialmente en etapas finales del proceso de aprendizaje.

5.2. PREDICCIÓN Y RENDIMIENTO

La tercera hipótesis de este análisis plantea que el nivel de rendimiento de los estudiantes influye en la precisión con la que predicen sus calificaciones. Para analizar esta relación, se emplea un modelo de mínimos cuadrados ordinarios donde la variable dependiente, "diferencia test n", representa la diferencia entre la calificación estimada en cada prueba a lo largo del semestre y la nota final obtenida. Un valor positivo indica que el estudiante se ha sobrevalorado, mientras que un valor negativo refleja una infravaloración.

Como variables independientes, se consideran "alto desempeño" y "bajo desempeño". Se define como alto desempeño a los estudiantes que se encuentran por encima del tercer cuartil y como bajo desempeño a aquellos por debajo del primer cuartil en la nota final de la asignatura. Los alumnos con rendimiento medio (ubicados entre los cuartiles 1 y 3, es decir, el 50% de la clase) se establecen como nivel base, contra la que se compararán las otras dos variables: dado que esta variable es categórica, una de sus categorías debe excluirse del modelo, en este caso, la de desempeño medio.

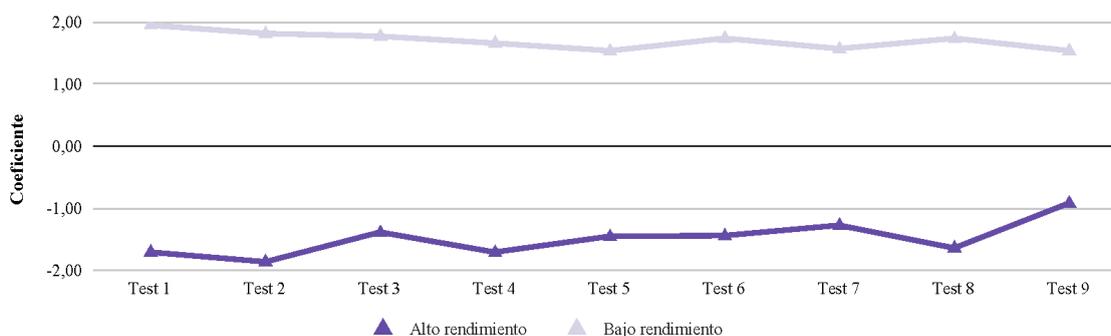
Tabla 5. Resultados del modelo 3

| | Alto rendimiento coeficiente | Valor p | Bajo rendimiento coeficiente | Valor p | R2 |
|--------|---------------------------------|---------|---------------------------------|----------|-------|
| Test 1 | -1.698 | 0.0003 | 1.953 | 6.50e-05 | 0.596 |
| Test 2 | -1.855 | 0.0001 | 1.815 | 0.0003 | 0.572 |
| Test 3 | -1.386 | 0.0028 | 1.773 | 0.0002 | 0.520 |
| Test 4 | -1.709 | 0.0006 | 1.663 | 0.0009 | 0.504 |

| | | | | | |
|--------|--------|--------|-------|--------|-------|
| Test 5 | -1.454 | 0.0025 | 1.535 | 0.0027 | 0.470 |
| Test 6 | -1.440 | 0.0037 | 1.741 | 0.0006 | 0.506 |
| Test 7 | -1.270 | 0.0264 | 1.569 | 0.0048 | 0.531 |
| Test 8 | -1.633 | 0.0008 | 1.732 | 0.0006 | 0.518 |
| Test 9 | -0.911 | 0.028 | 1.540 | 0.0006 | 0.452 |

Gráfico 2. Comparativa evolución del coeficiente en alto y bajo rendimiento

Fuente: Elaboración propia



En primer lugar, el p-valor de todas las variables se encuentra por debajo de 0.05, lo que indica que son estadísticamente significativas.

Analizando los coeficientes, se observa que los estudiantes de bajo desempeño presentan valores positivos, lo que significa que tienden a sobrevalorar su rendimiento al predecir sus calificaciones (en comparación con el nivel base, rendimiento medio). En cambio, los alumnos de alto desempeño muestran coeficientes negativos, lo que indica que tienden a infravalorarse (en comparación con el nivel base, rendimiento medio). Lo realmente interesante es que este patrón se repite a lo largo de absolutamente todo el curso, de principio a fin.

Este patrón es un reflejo del efecto Dunning-Kruger, según el cual las personas con menor competencia suelen sobreestimar sus capacidades, mientras que aquellos con mayor conocimiento o habilidad suelen ser más autocríticos.

Asimismo, se aprecia que, con el paso de las semanas, los estudiantes de bajo rendimiento mejoran ligeramente la precisión de sus predicciones, pasando el coeficiente de 1.953 a 1.540. No obstante, esta corrección es mucho más pronunciada en los alumnos de alto rendimiento. En su caso, la diferencia entre la nota predicha y la obtenida pasa de -1.698 en el primer test a -0.911 en el último, evidenciando un ajuste progresivo en sus estimaciones.

A la luz de los resultados obtenidos, se pueden identificar patrones diferenciados según el nivel de rendimiento académico de los estudiantes. Por un lado, los estudiantes de alto rendimiento tienden a infravalorarse, pero logran ajustar mejor sus predicciones con el tiempo. Por otro, los alumnos de bajo rendimiento sobrevaloran sus calificaciones y, aunque también mejoran ligeramente la precisión de sus estimaciones, lo hacen en menor medida que sus compañeros con mejor desempeño.

Si bien todos los alumnos tienden a sobrestimar su nota final, el error es menor al aumentar el rendimiento: los estudiantes con peor rendimiento (por debajo del Q1) sobrestiman su desempeño más que los estudiantes de rendimiento medio (entre el Q1 y el Q3), y estos a su vez sobrestiman su desempeño más que los alumnos de rendimiento alto (por encima del Q3).

Estos hallazgos confirman la tercera hipótesis y además, los resultados coinciden con estudios previos como los de Máté y Darabos (2017), Lew et al. (2010) y Gómez-Puerta et al. (2019), que también encontraron que los estudiantes con bajo rendimiento tienden a sobrestimarse, mientras que aquellos con mejor desempeño son más precisos o incluso tienden a subestimarse levemente.

5.3. PREDICCIÓN Y GÉNERO

En este apartado, se busca contrastar la hipótesis de que el género influye en la precisión con la que los estudiantes predicen su calificación.

Para ello, se realizaron nueve modelos de mínimos cuadrados ordinarios (MCO), en los que la variable dependiente es la "diferencia test n", al igual que en el modelo anterior sobre rendimiento y precisión en la predicción. Sin embargo, en este caso, la variable independiente es el género (mujer), que toma el valor de 1 si el estudiante es mujer y 0 si es hombre.

Tabla 6. Resultados del modelo 4

| | Mujer coeficiente | Valor p | R-cuadrado |
|--------|-------------------|---------|------------|
| Test 1 | 0.588 | 0.292 | 0.028 |
| Test 2 | 0.124 | 0.828 | 0.001 |
| Test 3 | 0.155 | 0.767 | 0.002 |
| Test 4 | 0.114 | 0.835 | 0.001 |
| Test 5 | 0.164 | 0.757 | 0.003 |
| Test 6 | 0.202 | 0.715 | 0.004 |
| Test 7 | 0.314 | 0.612 | 0.010 |
| Test 8 | 0.314 | 0.563 | 0.008 |
| Test 9 | -0.142 | 0.7609 | 0.003 |

El p-valor es superior a 0.05 en todos los casos, lo que indica que la variable género (mujer) no es estadísticamente significativa en ninguno de los modelos.

En conclusión, los resultados no permiten corroborar la cuarta hipótesis, dado que el género no tiene un impacto significativo en la precisión de la predicción de notas. No se confirma, por tanto, la idea de que las mujeres tienden sistemáticamente a infravalorar sus capacidades. Este hallazgo coincide con estudios como los de Máté y Darabos (2017) y Pallisgaard Støve (2020), que tampoco hallaron diferencias significativas entre hombres y mujeres. Sin embargo, contrasta con investigaciones como las de González-Betancor et al. (2017) y Quevedo Lezama (2023), donde se observaba una tendencia de las mujeres a subestimarse, atribuida a factores como la autoeficacia o el síndrome del impostor. Por tanto, nuestros resultados matizan la literatura existente, sugiriendo que la influencia del género puede depender del tipo de tarea o del contexto evaluativo específico.

6. CONCLUSIONES

Este trabajo ha permitido confirmar que la autoevaluación ha pasado de ser una práctica complementaria a consolidarse como una herramienta relevante en la educación superior. Su integración en las metodologías actuales responde a la necesidad de formar estudiantes autónomos, reflexivos y con capacidad para adaptarse a contextos cambiantes, tanto en el ámbito académico como profesional. La autoevaluación favorece una implicación activa del estudiante en su proceso de aprendizaje, facilitando la identificación de sus propias fortalezas y debilidades.

Los resultados del estudio reflejan que los estudiantes, especialmente en las primeras semanas del cuatrimestre, no predicen con precisión su calificación final. A medida que avanza el curso, la experiencia y la familiaridad con la asignatura contribuyen a mejorar, de forma progresiva, esa capacidad de predicción. Esta evolución confirma que la autoevaluación no es una competencia fija ni se desarrolla de forma automática, sino que requiere práctica, conocimiento de la tarea evaluada y habilidades metacognitivas.

También se ha confirmado que el rendimiento académico influye directamente en la precisión de la autoevaluación. Los estudiantes con menor rendimiento tienden a sobreestimarse, mientras que los de mejor rendimiento presentan predicciones más ajustadas o tienden a subestimarse ligeramente. Este patrón se corresponde con el efecto Dunning-Kruger, según el cual los estudiantes con bajo rendimiento no solo obtienen peores resultados, sino que también carecen de las habilidades necesarias para evaluar su propio desempeño de forma realista. Por el contrario, los estudiantes con mejor desempeño muestran una percepción más ajustada e incluso más modesta, posiblemente influenciada por el efecto de consenso falso.

Respecto al género, los datos obtenidos no permiten confirmar que existan diferencias significativas entre hombres y mujeres en la precisión de sus autoevaluaciones. Esto contradice algunos estudios previos que señalaban una tendencia de las mujeres a infravalorar su rendimiento. En este caso, la variable género no ha mostrado un efecto relevante, lo que indica que su influencia podría depender del tipo de tarea evaluada, el enfoque docente o el grado de madurez académica. Por tanto, conviene seguir explorando esta dimensión en futuras investigaciones para esclarecer su posible impacto.

La principal aportación de este estudio reside en el seguimiento longitudinal realizado en una asignatura concreta a lo largo de todo el cuatrimestre, lo que ha permitido observar la evolución de la precisión de la autoevaluación en relación con los resultados reales. Este enfoque difiere de los estudios transversales habituales y ofrece una perspectiva dinámica del fenómeno. El uso de análisis estadísticos, como las regresiones lineales, ha permitido contrastar de forma rigurosa las hipótesis planteadas en la revisión de la literatura.

En cuanto al significado de los hallazgos, se confirma que la autoevaluación puede ser una estrategia útil siempre que esté acompañada de un entorno pedagógico adecuado. La práctica sistemática, el acompañamiento por parte del profesorado y el uso de herramientas que favorezcan la reflexión del estudiante pueden mejorar la precisión en las predicciones y fomentar un aprendizaje más consciente y autorregulado.

Los resultados obtenidos tienen implicaciones prácticas para el ámbito educativo. El profesorado universitario puede utilizar esta información para diseñar actividades que promuevan una autoevaluación efectiva desde el inicio del curso. Además, la evidencia empírica generada respalda la importancia de incluir el desarrollo de competencias metacognitivas en los planes de estudio. También puede ser útil para los propios estudiantes, ya que les permite tomar conciencia de los sesgos que pueden afectar a su autoevaluación y trabajar de forma más consciente sobre su desempeño académico.

En términos de valor añadido, este trabajo contribuye con evidencia empírica al debate sobre la autoevaluación en educación superior, en un contexto poco explorado: estudiantes de alto rendimiento en una universidad española. Se ha demostrado que la experiencia acumulada y el rendimiento académico influyen en la precisión de la autoevaluación, mientras que el género no parece desempeñar un papel determinante en este caso. Estos resultados permiten matizar algunas conclusiones previas y aportan una visión más contextualizada del fenómeno.

Desde un punto de vista práctico, los resultados pueden utilizarse para diseñar estrategias que mejoren la precisión de la autoevaluación, como el uso de plataformas digitales con retroalimentación inmediata, actividades de reflexión y programas de formación metacognitiva desde etapas tempranas del aprendizaje universitario.

Este trabajo también presenta algunas limitaciones. La muestra está compuesta por estudiantes con un nivel académico elevado, lo que puede reducir la generalización de los

resultados. Además, el enfoque cuantitativo impide recoger percepciones subjetivas sobre el proceso de autoevaluación, y al centrarse en una única asignatura, se limita la variedad de contextos analizados. Estos factores deben tenerse en cuenta a la hora de interpretar los resultados.

A partir de estas limitaciones, se plantean distintas líneas de investigación futuras. Sería útil replicar este estudio en diferentes titulaciones, con estudiantes de distintos niveles y perfiles académicos, y en universidades con características diversas. También sería interesante analizar si un entrenamiento específico en autoevaluación contribuye a mejorar su precisión y evaluar si las posibles diferencias de género se mantienen, se reducen o desaparecen a lo largo del proceso formativo. Por último, resulta relevante explorar cómo el uso de tecnologías digitales y plataformas de aprendizaje influye en la evolución de la capacidad de autoevaluarse, especialmente en contextos híbridos o virtuales.

En definitiva, este trabajo pone de relieve la importancia de la autoevaluación como herramienta formativa, aporta datos empíricos relevantes para su análisis y plantea nuevas preguntas que pueden guiar futuras investigaciones en este campo. Más allá de confirmar o rechazar hipótesis, este estudio pretende abrir vías de mejora en la práctica educativa y en la formación del estudiantado universitario.

¿Los alumnos de ICADE saben que no saben? No es posible generalizar. Este estudio demuestra que aquellos con menor rendimiento, efectivamente, no saben que no saben. Sin embargo, los estudiantes con mejor desempeño parecen no saber que saben. Esta paradoja subraya la complejidad de la autoevaluación y sugiere que el conocimiento de uno mismo no siempre avanza al mismo ritmo que el conocimiento académico.

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

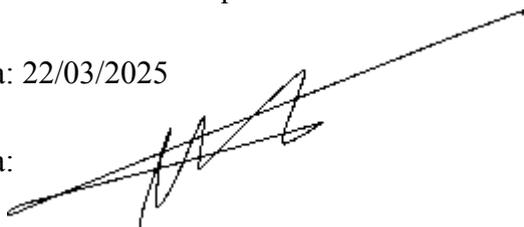
Por la presente, yo, María Reguero Cabada, estudiante de Derecho. y Administración y Dirección de empresas de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "¿Los estudiantes de comillas saben que no saben? La precisión de la predicción y el efecto Dunning-Kruger", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. Brainstorming de ideas de investigación: Utilizado para idear y esbozar posibles áreas de investigación.
2. Crítico: Para encontrar contra-argumentos a una tesis específica que pretendo defender.
3. Metodólogo: Para descubrir métodos aplicables a problemas específicos de investigación.
4. Constructor de plantillas: Para diseñar formatos específicos para secciones del trabajo.
5. Corrector de estilo literario y de lenguaje: Para mejorar la calidad lingüística y estilística del texto.
6. Revisor: Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
7. Traductor: Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 22/03/2025

Firma:



BIBLIOGRAFÍA

- Blackwood, T. (2013). Business undergraduates' knowledge monitoring accuracy: how much do they know about how much they know?. *Teaching in Higher Education*, 18(1), 65-77. <https://doi.org/10.1080/13562517.2012.694100>
- Bolívar-Cruz, A., Verano-Tacoronte, D., & Galván-Sánchez, I. (2018). Do self-efficacy, incentives and confidence in public speaking influence how students self-assess?/¿Influyen la autoeficacia, los incentivos y la confianza para hablar en público en cómo se autoevalúan los estudiantes?. *Culture and Education*, 30(3), 528-555. <https://doi.org/10.1080/11356405.2018.1488420>
- Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher education*, 18(5), 529-549. <https://doi.org/10.1007/BF00138746>
- Boud, D. (1995). *Enhancing Learning Through Self-assessment* (1st ed.). Routledge. <https://doi.org/10.4324/9781315041520>
- Boud, D., & Brew, A. (1995). Developing a typology for learner self-assessment practices. *Research and development in Higher Education*, 18(1), 130-135. https://www.researchgate.net/publication/327350493_Developing_a_typology_for_learner_self-assessment_practices
- Cardoso, A. P. G., & Ramos, E. S. (2021). Educación Tradicional: Un modelo de enseñanza centrado en el estudiante. *Cienciamatria*, 7(12), 962-975. <https://dialnet.unirioja.es/servlet/articulo?codigo=7915387>
- Cuesta-Melo, C. H., Lucero-Zambrano, M. A., & Herrera-Mosquera, L. (2022). The Influence of Self-Assessment on the English Language Learning Process of students from a public university in Colombia. *Colombian Applied Linguistics Journal*, 24(1), 89-104. <https://doi.org/10.14483/22487085.17673>
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current directions in psychological science*, 12(3), 83-87. <https://doi.org/10.1111/1467-8721.01235>

- Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In *Advances in experimental social psychology* (Vol. 44, pp. 247-296). Academic Press. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational behavior and human decision processes*, 105(1), 98-121. <https://doi.org/10.1016/j.obhdp.2007.05.002>
- Ferraro, P. J. (2010). Know thyself: Competence and self-awareness. *Atlantic Economic Journal*, 38, 183-196. <https://doi.org/10.1007/s11293-010-9226-2>
- Fitzgerald, J. T., White, C. B., & Gruppen, L. D. (2003). A longitudinal study of self-assessment accuracy. *Medical education*, 37(7), 645-649. <https://doi.org/10.1046/j.1365-2923.2003.01567.x>
- Gómez-Puerta, M., Bueno Vargas, M. J., Casero Martínez, A., Chiner, E., Fajardo Bravo, I., Melero-Pérez, P., ... & Gómez Martínez, Á. (2019). Identificación de la presencia de sesgo cognitivo derivado del efecto de Dunning-Kruger en estudiantes universitarios. <http://hdl.handle.net/10045/99588>
- González-Betancor, S. M., Bolívar-Cruz, A., & Verano-Tacoronte, D. (2017). Self-assessment accuracy in higher education: The influence of gender and performance of university students. *Active learning in higher education*, 20(2), 101-114. <https://doi.org/10.1177/1469787417735604>
- Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, 92(1), 160. <https://doi.org/10.1037/0022-0663.92.1.160>
- Hassan, O. A., Fox, A., & Hannah, G. (2014). Self-and peer-assessment: Evidence from the accounting and finance discipline. *Accounting education*, 23(3), 225-243. <https://doi.org/10.1080/09639284.2014.905259>
- Hill, T. (2016). Do accounting students believe in self-assessment?. *Accounting Education*, 25(4), 291-305. <https://doi.org/10.1080/09639284.2016.1191271>

- Hodges, B., Regehr, G., & Martin, D. (2001). Difficulties in recognizing one's own incompetence: novice physicians who are unskilled and unaware of it. *Academic Medicine*, 76(10), S87-S89. https://journals.lww.com/academicmedicine/fulltext/2001/10001/Difficulties_in_Recognizing_One_s_Own.29.aspx
- Krajč, M., Ortmann, A., & Ryvkin, D. (2008). How to undo biased self-assessments. <https://d1wqtxts1xzle7.cloudfront.net/30815487/>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://holub.com/goodies/incompetence.pdf>
- Landázuri, M. M., Aguilar, G. F. C., & Alcívar, I. A. M. (2024). TRANSFORMACIÓN CURRICULAR EN SOCIOLOGÍA PARA POTENCIAR EL DESARROLLO INTEGRAL DE PROFESIONALES EN LA UNIVERSIDAD DE GUAYAQUIL. *Revista Científica y Arbitrada de Ciencias Sociales y Trabajo Social: Tejedora*. ISSN: 2697-3626, 7(16), 185-202. <https://doi.org/10.56124/tj.v7i16.012>
- Leach, L. (2012). Optional self-assessment: some tensions and dilemmas. *Assessment & Evaluation in Higher Education*, 37(2), 137-147. <https://doi.org/10.1080/02602938.2010.515013>
- Lew, M. D., Alwis, W. A. M., & Schmidt, H. G. (2010). Accuracy of students' self-assessment and their beliefs about its utility. *Assessment & Evaluation in Higher Education*, 35(2), 135-156. <https://doi.org/10.1080/02602930802687737>
- Lezama, C. R. Q. (2023). El síndrome del impostor en relación con la enseñanza: limitantes, factores y consecuencias para alcanzar un aprendizaje integral. *Revista Científica Disciplinarias*, 2(3), 25-48. <https://doi.org/10.71727/disciplinarias.v2i3.134>
- Márquez Vigil, J., Lazcano Benito, L. I., Bada Olan, M. D. C., & Arroyo Barrigüete, J. L. (2023). Class participation and feedback as enablers of student academic performance. <https://doi.org/10.1177/21582440231177298>

- Máté, D., & Darabos, É. (2017). Measuring the accuracy of self-assessment among undergraduate students in higher education to enhance competitiveness. *Journal of Competitiveness*, 9(2). <https://www.cjournal.cz/files/251.pdf>
- Melo Rojas, G. M., & Hernández Herrera, C. A. (2024). Percepción estudiantil de la responsabilidad social universitaria en la gestión organizacional: un estudio en una unidad académica del Instituto Politécnico Nacional. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 14(28). <https://doi.org/10.23913/ride.v14i28.1803>
- Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. *Metacognition and Learning*, 6, 303-314. <https://doi.org/10.1007/s11409-011-9083-7>
- Ministros Europeos. (1999). Declaración de Bolonia. Declaración conjunta de los ministros europeos de educación. Bolonia.
- Mynttinen, S., Sundström, A., Vissers, J., Koivukoski, M., Hakuli, K., & Keskinen, E. (2009). Self-assessed driver competence among novice drivers—a comparison of driving test candidate assessments and examiner assessments in a Dutch and Finnish sample. *Journal of safety research*, 40(4), 301-309. <https://doi.org/10.1016/j.jsr.2009.04.006>
- Nulty, D. D. (2011). Peer and self-assessment in the first year of university. *Assessment & evaluation in higher education*, 36(5), 493-507. <https://doi.org/10.1080/02602930903540983>
- Olmos Migueláñez, S. (2008). Evaluación formativa y sumativa de estudiantes universitarios: aplicación de las tecnologías a la evaluación educativa. <https://gedos.usal.es/handle/10366/18453>
- Panadero, E. (2011). *Ayudas instruccionales a la autoevaluación y la autorregulación: evaluación de la eficacia de los guiones de autoevaluación frente a la de las rúbricas* (Doctoral dissertation, Universidad Autónoma de Madrid). <https://repositorio.uam.es/handle/10486/10832>

- Panadero, E., Brown, G. T., & Strijbos, J. W. (2015). The future of student self- assessment: A review of known unknowns and potential directions. *Educational psychology review*, 28(4), 803-830. <https://doi.org/10.1007/s10648-015-9350-2>
- Park, Y. J., & Santos-Pinto, L. (2010). Overconfidence in tournaments: Evidence from the field. *Theory and Decision*, 69, 143-166. <https://doi.org/10.1007/s11238-010-9200-0>
- ROJO, V. Á., CARMONA, M. T. P., SANTERO, J. R., GORDILLO, J. J. T., & ORTEGA, M. S. (2011). Análisis de la participación del alumnado universitario en la evaluación de su aprendizaje. *Revista Española de Pedagogía*, 69(250), 401–425. <http://www.jstor.org/stable/23766699>
- Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of experimental social psychology*, 13(3), 279-301. [https://doi.org/10.1016/0022-1031\(77\)90049-X](https://doi.org/10.1016/0022-1031(77)90049-X)
- Salazar, O. C. (2010). Evaluación de los aprendizajes en estudios de postgrado bajo los principios de la evaluación participativa. *Educación y Desarrollo Social*, 4(1), 10-37. <https://dialnet.unirioja.es/servlet/articulo?codigo=5386184>
- Salom, M. A. C. (2019). Una oportunidad para avanzar hacia la evaluación auténtica en Educación Física. *Retos: nuevas tendencias en educación física, deporte y recreación*, (36), 259-265. <https://dialnet.unirioja.es/servlet/articulo?codigo=7260912>
- Santos Guerra, M. A. (1994). Entre bastidores. El lado oculto de la organización escolar. *Málaga: Aljibe*. <https://cbibliotecavirtual.wordpress.com/wp-content/uploads/2017/07/organizacion-escolar.pdf>
- Støve, M. P. (2020). Physiotherapy students' self-assessment of performance—Are there gender differences in self-assessment accuracy?. *Physiotherapy Research International*, 26(1), e1878. <https://doi.org/10.1002/pri.1878>
- Taras, M. (2010). Student self-assessment: processes and consequences. *Teaching in higher education*, 15(2), 199-209. <https://doi.org/10.1080/13562511003620027>

Tejeiro, R. A., Gómez-Vallecillo, J. L., Romero, A. F., Pelegrina, M., Wallace, A., & Emberley, E. (2012). La autoevaluación sumativa en la enseñanza superior: implicaciones de su inclusión en la nota final. *Electronic journal of research in educational psychology*, 10(2), 789-812.
<https://www.redalyc.org/articulo.oa?id=293123547014>

ANEXO

Tabla 7. Tabla de correlaciones entre notas predichas y nota final

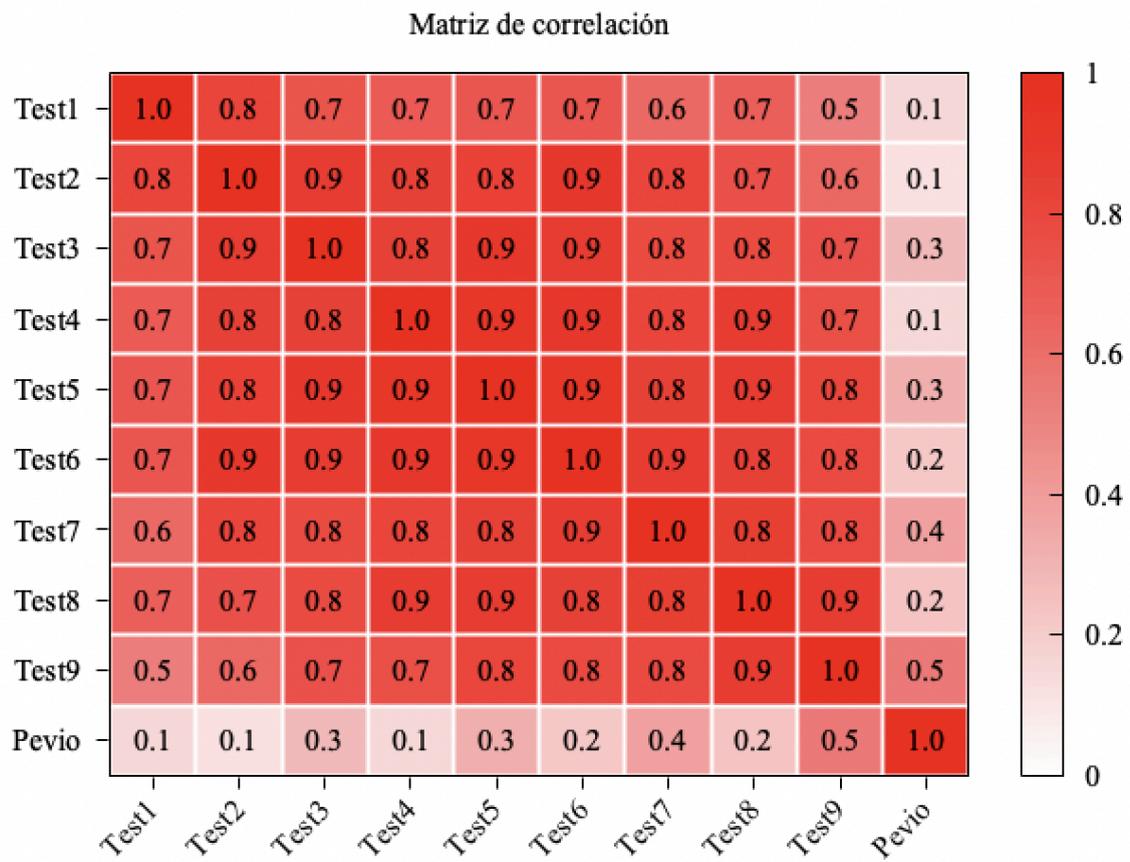


Tabla 8. Descripción de las variables utilizadas en el análisis

| Variable | Descripción |
|---------------------|--|
| NOTA FINAL | Nota real obtenida por el estudiante en la asignatura al finalizar el semestre (variable dependiente en varios modelos). |
| TEST 1 - TEST 9 | Nota predicha por el estudiante en cada uno de los nueve test realizados a lo largo del semestre, antes del examen final. |
| DIFERENCIA TEST n | Diferencia entre la nota predicha en el test n y la nota final obtenida. Un valor positivo indica sobreestimación; un valor negativo, infravaloración. |
| ALTO DESEMPEÑO | Variable categórica que toma valor 1 si el estudiante se encuentra por encima del tercer cuartil de la nota final; 0 en caso contrario. |
| BAJO DESEMPEÑO | Variable categórica que toma valor 1 si el estudiante se encuentra por debajo del primer cuartil de la nota final; 0 en caso contrario. |
| DESEMPEÑO MEDIO | Categoría de referencia (entre Q1 y Q3) frente a la cual se comparan los grupos de alto y bajo desempeño. No se incluye explícitamente en los modelos. |
| MUJER | Variable que toma valor 1 si el estudiante es mujer y 0 si es hombre. |