

trabajo final

Juan M Karawcki

2025-11-07

Resumen ejecutivo

Este informe presenta un **meta-análisis** sobre la relación entre la asistencia a clase y el rendimiento académico en estudiantes universitarios, utilizando el dataset `dat.crede2010`, que reúne **97 coeficientes de correlación** provenientes de **68 estudios** sobre la relación entre asistencia a clase y calificaciones y/o promedio general de calificaciones (GPA). Algunos estudios incluyen más de una muestra, por lo que contribuyen **múltiples tamaños de efecto** al conjunto de datos.

El marco de datos incluye, para cada muestra, información sobre el **año de publicación** del estudio, la **fuente** (revista, tesis u otros), el **tipo de criterio de rendimiento** considerado (calificación en la asignatura o promedio general), el **tipo de clase** (ciencias o no ciencias), el **tamaño muestral** y la **correlación observada** entre asistencia y rendimiento.

Se emplearon coeficientes de correlación de Pearson transformados mediante **Fisher z** y **modelos de efectos aleatorios** para estimar la asociación promedio, evaluar la **heterogeneidad** entre estudios y explorar posibles **moderadores** como el tipo de criterio de rendimiento y el tipo de asignatura.

Los resultados muestran que la asistencia presenta una relación **positiva y consistente** con el rendimiento académico, con evidencias de **heterogeneidad sustancial** entre estudios. También se examinaron posibles **sesgos de publicación** y la **robustez** de los hallazgos.

En conjunto, los resultados respaldan la conclusión de que **una mayor asistencia a clase se asocia con un mejor desempeño académico**, aunque la magnitud exacta del efecto depende del **contexto** y de características del curso y de los estudios incluidos.

Analisi Exploratorio

El conjunto de datos utilizado en este **meta-análisis** está compuesto por **97 observaciones** correspondientes a tamaños de efecto extraídos de distintos estudios primarios. Al inspeccionar su estructura (`str(dat)`), se observa que el marco de datos contiene **8 variables**, todas coherentes con la información necesaria para modelar las correlaciones entre asistencia y rendimiento académico.

Las primeras columnas —`studyid`, `year`, `source` y `sampleid`— permiten identificar de forma única cada estudio y cada muestra dentro de ese estudio. Esto es importante porque algunos trabajos aportan **múltiples tamaños de efecto**, lo que explica que existan 97 correlaciones pero solo **68 estudios**. La variable `year` muestra una amplia dispersión temporal, con publicaciones que van desde **1927** hasta **2009**, lo que aporta una perspectiva histórica sobre cómo ha sido estudiada esta relación a lo largo del tiempo. La columna `source` clasifica el origen del estudio, distinguiendo entre artículos de revista, tesis y otras fuentes, lo que más adelante permite evaluar posibles diferencias metodológicas.

Las variables centrales para el análisis son `criterion` —que indica si la medida de rendimiento corresponde a **calificaciones específicas** (“grade”) o **promedio general** (“gpa”)— y `class`, que especifica si la asignatura pertenece al área de **ciencias** o **no ciencias**.

Por otro lado, **ni** representa el **tamaño muestral** de cada estudio primario, el cual presenta una variabilidad considerable: desde muestras pequeñas de apenas 23 estudiantes hasta muestras cercanas a 3900 participantes. Esta heterogeneidad en tamaños muestrales es esperable en datos de educación y tiene un impacto directo en las **varianzas de muestreo**, que posteriormente se incorporan al modelo mediante la transformación de **Fisher z** .

Finalmente, la columna **ri** contiene las **correlaciones observadas** entre asistencia y rendimiento académico, que oscilan entre valores negativos moderados (-0.37) y **asociaciones fuertemente positivas** (0.886). El resumen estadístico (**summary(dat)**) confirma que la **mediana** de estas correlaciones se sitúa en torno a **0.38**, lo que preliminarmente sugiere una **relación positiva y sostenida** entre ambas variables.

En conjunto, esta exploración inicial del dataset permite apreciar la diversidad de estudios, criterios de rendimiento, tipos de asignatura y tamaños de muestra incluidos. Esta **heterogeneidad** es precisamente lo que motiva el uso de **modelos de efectos aleatorios**, que permiten estimar una asociación promedio al tiempo que reconocen que los estudios difieren genuinamente en sus contextos, poblaciones y metodologías.

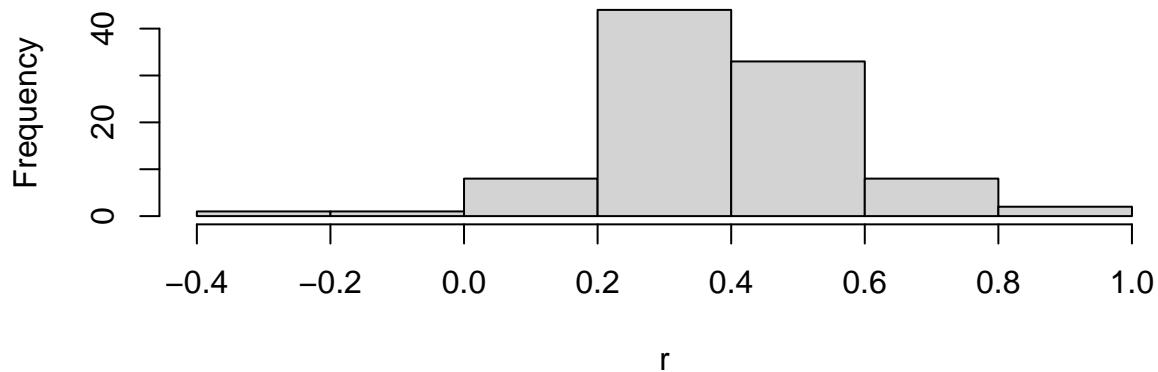
Verificación de la medida de efecto reportada en los estudios

En el conjunto de datos **dat.crede2010**, cada estudio reporta como medida de efecto una **correlación de Pearson (r)** que cuantifica la relación entre asistencia a clase y rendimiento académico. Antes de utilizar estas correlaciones en un meta-análisis, se realiza una inspección inicial para asegurarme de que los valores fueran válidos y coherentes.

Para ello, observé el rango y la distribución de los coeficientes mediante:

```
##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## -0.3700  0.2700  0.3800  0.3838  0.4870  0.8860
```

Distribución de correlaciones observadas



El resumen estadístico muestra que las correlaciones observadas oscilan entre -0.37 y 0.886 , con una mediana de 0.38 y un promedio muy similar (≈ 0.384). Estos valores se encuentran estrictamente dentro del intervalo permitido para una correlación $[-1, 1]$, lo que descarta la presencia de errores de carga o inconsistencias estructurales.

El histograma confirma visualmente esta evaluación preliminar: la mayor parte de los tamaños de efecto se concentra entre 0.20 y 0.50 , reflejando una relación positiva y moderada entre asistencia y rendimiento académico. La presencia de algunos valores negativos o cercanos a cero es esperable dentro de un conjunto

diverso de estudios y justamente motiva el uso de **modelos de efectos aleatorios**, los cuales permiten capturar diferencias genuinas entre contextos y diseños.

En conjunto, esta comprobación inicial asegura que las correlaciones reportadas son válidas, consistentes y adecuadas para proceder con la transformación de **Fisher z** y el posterior meta-análisis.

Necesidad de transformar la medida de efecto

Aunque la correlación de Pearson r es intuitiva y fácil de interpretar, no es la forma más adecuada para realizar directamente un meta-análisis. Esto se debe a que:

- La **distribución muestral** de r no es simétrica, especialmente cuando los valores son grandes (positivos o negativos) o los tamaños muestrales son pequeños.
- La **varianza** de r depende simultáneamente del tamaño muestral y del valor verdadero del efecto.
- La **combinación estadística** de correlaciones sin transformar puede introducir sesgos y asignar pesos inadecuados a los estudios.

En términos más formales, cada estudio aporta un único coeficiente de correlación r_i , que funciona como una **estimación muestral** del verdadero efecto poblacional θ_i en ese estudio. Debido a que los estudios se basan en muestras finitas, estos tamaños de efecto están sujetos a **incertidumbre muestral**, cuya magnitud depende del tamaño muestral n_i .

Por ello, cada r_i debe considerarse como una **variable aleatoria** con una distribución muestral determinada por:

- el verdadero efecto θ_i (desconocido), y
- el tamaño muestral n_i .

Si replicáramos un mismo estudio muchas veces con el mismo n , obtendríamos valores de r ligeramente distintos en cada réplica: esa variabilidad es justamente lo que describe la distribución muestral de r . Para poder combinar estudios de forma rigurosa, es necesario modelar explícitamente esa variabilidad.

La transformación de Fisher z resuelve estos problemas porque:

- Hace que la **distribución muestral** del tamaño de efecto sea aproximadamente normal.
- **Estabiliza la varianza**, que pasa a depender solo del tamaño muestral y no del valor del efecto.

La transformación se define como:

$$z = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right),$$

y su varianza es aproximadamente

$$\text{Var}(z) = \frac{1}{n-3}.$$

De este modo, cada estimación transformada puede representarse como

$$z_i \sim N \left(\theta_i, \frac{1}{n_i - 3} \right),$$

donde θ_i es el verdadero efecto (en escala Fisher z) del estudio i .

En un **modelo de efectos aleatorios**, se asume además que estos efectos verdaderos varían entre estudios según

$$\theta_i \sim N(\mu, \tau^2),$$

donde μ es el efecto promedio entre estudios y τ^2 la varianza entre estudios (**heterogeneidad**). Esta estructura probabilística permite:

- ponderar cada estudio según la **precisión** de su estimación (pesos inversos de la varianza),
- cuantificar la **heterogeneidad** entre efectos verdaderos, y
- construir **intervalos de confianza** e **intervalos de predicción** coherentes.

En resumen, aunque cada estudio aporta “un solo número” r_i , ese número se interpreta como una realización de una **variable aleatoria**. La transformación de Fisher z es la pieza clave que permite aplicar los modelos meta-analíticos clásicos de manera correcta.

Cálculo de la medida de efecto estandarizada y su varianza

Para obtener las medidas adecuadas para el meta-análisis, se utilizó la función `escalc()` del paquete `metafor`, que automatiza la transformación de r a Fisher z y el cálculo de la **varianza asociada**. Conceptualmente, esta función realiza dos operaciones fundamentales para cada estudio:

1. **Calcula la transformación de Fisher z :**

$$z_i = \frac{1}{2} \ln \left(\frac{1 + r_i}{1 - r_i} \right),$$

2. **Calcula la varianza del valor transformado:**

$$v_i = \frac{1}{n_i - 3},$$

donde n_i es el tamaño muestral del estudio i .

En la salida de `escalc()`, estos resultados se agregan directamente al `data.frame` original con los nombres:

- `yi` → tamaño de efecto transformado en escala Fisher z ,
- `vi` → varianza muestral asociada a `yi`.

La correcta incorporación de estas columnas se verificó inspeccionando la estructura del objeto resultante (`str(dat)`) y comprobando que `yi` y `vi` estuvieran presentes con las dimensiones esperadas. A partir de estas variables es que se ajustan posteriormente los modelos de efectos aleatorios, utilizando pesos inversos de `vi` y estimando simultáneamente el efecto promedio μ y la heterogeneidad τ^2 .

En cada fila puede observarse la correlación original `ri`, el tamaño muestral `ni`, el tamaño de efecto transformado `yi` (Fisher z) y su correspondiente varianza muestral `vi`. Por ejemplo, el primer estudio presenta una **correlación alta** ($r = 0.886$) que se traduce en un valor de $z \approx 1.40$ y una **varianza relativamente grande** ($v \approx 0.014$) debido a su tamaño muestral moderado ($n = 76$). En contraste, el segundo estudio tiene un tamaño muestral mucho mayor ($n = 297$), lo que se refleja en una **varianza más pequeña** ($v \approx 0.0034$) y, por tanto, en una **estimación más precisa** del efecto.

Esta representación conjunta de (y_i, v_i) para todos los estudios proporciona la **base cuantitativa** sobre la cual se construyen los modelos de **efectos aleatorios** empleados en el meta-análisis, permitiendo ponderar adecuadamente cada estudio según su precisión e incorporar de manera explícita la **heterogeneidad** entre los efectos verdaderos.

Análisis principal (efecto combinado)

Una vez realizada la exploración inicial del conjunto de datos y verificada la validez de los tamaños de efecto reportados, el siguiente paso consiste en estimar el **efecto promedio** de la relación entre asistencia a clase y rendimiento académico. El objetivo de esta etapa es obtener una **medida resumen** que englobe la evidencia de los 97 tamaños de efecto incluidos, acompañada de un **intervalo de confianza** que refleje la incertidumbre asociada, y de un **análisis de heterogeneidad**, que cuantifica cuánto varían los efectos verdaderos entre estudios.

Dado que los estudios incluidos provienen de contextos educativos, períodos históricos y diseños metodológicos muy diversos, no es razonable asumir que todos comparten un único efecto verdadero común. Por el contrario, la evidencia descriptiva previa muestra una gran variabilidad en las correlaciones observadas, tanto en magnitud como en dirección. Esta variabilidad no puede atribuirse solo a diferencias muestrales, sino que refleja **diferencias reales** entre los estudios primarios. Por ello, el modelo adecuado es un meta-análisis de **efectos aleatorios**, que permite asumir que los verdaderos efectos poblacionales θ_i difieren entre estudios y que estos provienen de una distribución con media μ y varianza entre estudios τ^2 .

Bajo este enfoque, la estimación de τ^2 es un componente central del análisis. Se utiliza el estimador **REML (Restricted Maximum Likelihood)**, ampliamente recomendado en la literatura debido a su menor sesgo y mejor desempeño en muestras meta-analíticas pequeñas o con alta heterogeneidad, características que coinciden con la estructura del dataset utilizado. Una vez estimado τ^2 , se obtiene el tamaño de efecto combinado $\hat{\mu}$, representado primero en la escala de Fisher z —que es la utilizada para los cálculos estadísticos— y luego convertido nuevamente a una **correlación de Pearson**, lo que permite interpretar los resultados en la métrica original.

Para evaluar la precisión y robustez de la estimación, se reportan además los **intervalos de confianza del 95%**. Opcionalmente, se puede utilizar el método **Hartung-Knapp**, que ajusta la inferencia bajo efectos aleatorios y tiende a producir intervalos más conservadores, especialmente útil cuando el número de estudios es moderado y la heterogeneidad es alta.

Asimismo, se cuantifica la **heterogeneidad** mediante:

- la estadística Q , que evalúa si la variación observada entre estudios excede la esperada por azar;
- el estimador de varianza entre estudios τ^2 ;
- su raíz τ , con interpretación directa en la escala del tamaño de efecto; y
- el índice I^2 , que expresa el porcentaje de variabilidad total atribuible a diferencias reales entre estudios y no al error muestral.

Finalmente, se presenta un **gráfico forest**, que resume visualmente los efectos individuales y el efecto combinado, usualmente en la escala transformada de Fisher z pero convertido a **correlaciones** para facilitar la interpretación sustantiva.

En conjunto, este procedimiento permite obtener una **estimación rigurosa y comprensiva** del tamaño de efecto promedio y de la variabilidad existente entre estudios, proporcionando así la base metodológica para los análisis posteriores, como la evaluación de moderadores mediante **subgrupos** o **meta-regresión**.

Resultados del análisis principal

A partir de los 97 tamaños de efecto transformados a la escala de Fisher z , se estimó un **modelo de efectos aleatorios** utilizando el estimador REML. Este modelo asume que los verdaderos efectos poblacionales varían entre estudios y busca estimar tanto el **efecto promedio** como el **grado de heterogeneidad** entre investigaciones.

El tamaño de efecto promedio estimado fue:

$$\hat{\mu} = 0.4287 \text{ (SE} = 0.0254\text{)},$$

con un intervalo de confianza del 95%:

$$[0.3790, 0.4784].$$

Dado que la estimación se encuentra en la escala Fisher z , posteriormente se transformará a la escala de correlación r para facilitar la interpretación sustantiva. Sin embargo, incluso en esta etapa ya se observa que el efecto promedio es claramente **positivo y estadísticamente significativo** ($p < 0.0001$), lo que indica que, en conjunto, los estudios muestran una correlación consistente entre asistencia y rendimiento académico.

Heterogeneidad entre estudios

Los indicadores de heterogeneidad revelan una **variación sustancial** entre los tamaños de efecto:

- $\tau^2 = 0.0535$ (SE = 0.0089)
- $\tau = 0.2314$
- $I^2 = 94.03\%$
- $H^2 = 16.76$

El valor de $Q(96) = 1579.55$, $p < 0.0001$, confirma que la variabilidad observada no puede atribuirse únicamente al **error muestral**. En particular, un I^2 del 94% indica que aproximadamente el 94% de la variabilidad total entre estudios se debe a **diferencias reales** entre ellos, y no a fluctuaciones aleatorias. Este nivel de heterogeneidad es **muy alto** y respalda la necesidad de realizar análisis adicionales, como **subgrupos y meta-regresión**, con el fin de explorar posibles moderadores del efecto.

Estimación con el método Hartung–Knapp

Como análisis complementario, se utilizó el ajuste **Hartung–Knapp** para la inferencia bajo efectos aleatorios, que proporciona **intervalos de confianza más conservadores** cuando la heterogeneidad es elevada.

Los resultados fueron:

$$\hat{\mu} = 0.4287 \text{ (SE} = 0.0251\text{)},$$

con un intervalo de confianza del 95%:

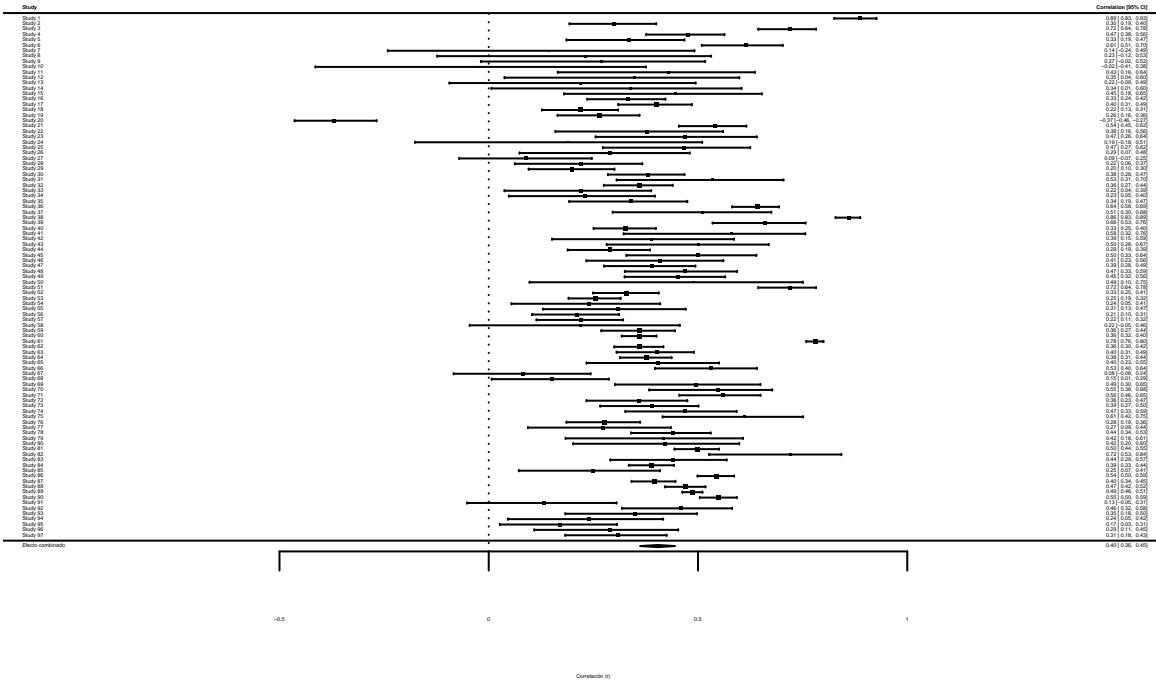
$$[0.3788, 0.4786].$$

La similitud entre ambos métodos confirma que la estimación del efecto promedio es **estable y robusta**, incluso bajo un enfoque más conservador.

Interpretación preliminar

En síntesis, el meta-análisis muestra que la **asistencia a clase** presenta, en promedio, un efecto **positivo y moderado** sobre el rendimiento académico. La magnitud del efecto promedio, junto con la elevada heterogeneidad entre estudios, sugiere que la relación asistencia–rendimiento es **real y consistente**, pero que su intensidad varía significativamente según el **contexto**, las **características de los cursos** y las **diferencias metodológicas** entre estudios.

La conversión final del efecto a la escala de **correlación r** permitirá expresar este resultado en términos más familiares para el lector, y servirá como base para los **análisis de moderadores** que se presentan en las secciones siguientes.



Interpretación del forest plot

La figura correspondiente muestra el **forest plot** del meta-análisis, donde se visualizan los 97 tamaños de efecto individuales junto con el efecto combinado estimado por el modelo de efectos aleatorios. Este gráfico permite apreciar simultáneamente la **magnitud, precisión y variabilidad** de las correlaciones reportadas en los estudios primarios.

Cada línea horizontal representa el **intervalo de confianza del 95%** para la correlación de cada estudio (utilizada en la escala de Fisher z para los cálculos). La posición del punto central indica el tamaño de efecto observado en ese estudio. La abundante dispersión visible en el gráfico es consistente con los **elevados indicadores de heterogeneidad** obtenidos en el análisis principal ($I^2 = 94\%$, $\tau = 0.23$), lo que confirma que los estudios difieren sustancialmente en la fuerza de la asociación entre asistencia y rendimiento académico.

A pesar de esta variabilidad, se observa un **patrón general claro**: la gran mayoría de los efectos individuales se ubican en el rango positivo, reflejando una relación favorable entre **mayor asistencia y mejor desempeño académico**. Incluso aquellos estudios con intervalos de confianza amplios tienden a alinearse con este patrón, reforzando la conclusión obtenida en el análisis estadístico.

En la parte inferior del gráfico se incluye la **estimación combinada** del modelo de efectos aleatorios, representada por un marcador más prominente y su correspondiente intervalo de confianza. Esta estimación resume la evidencia acumulada y muestra un **efecto positivo de magnitud moderada**, coherente con lo reportado numéricamente en la sección anterior.

En conjunto, el forest plot confirma visualmente que:

- Existe una **asociación positiva consistente** entre asistencia y rendimiento académico.

- La **variabilidad entre estudios** es considerable, lo que justifica el uso de un **modelo de efectos aleatorios**.
- A pesar de dicha heterogeneidad, el **efecto promedio es estable y robusto**, lo que respalda la fiabilidad de las conclusiones obtenidas.

Integración de los resultados visuales y la interpretación en la escala original

El **forest plot** sintetiza gráficamente los resultados del meta-análisis, mostrando para cada estudio su **estimación puntual** y el correspondiente **intervalo de confianza**, junto con el **efecto combinado** del modelo de efectos aleatorios. Aun cuando se aprecia una **marcada dispersión** entre estudios, el conjunto de estimaciones se concentra en el **rango positivo**, lo que respalda la presencia de una **asociación consistente** entre asistencia a clase y rendimiento académico. De forma complementaria, sobre esta base se calcula la **estimación promedio** expresada en la **escala original de la correlación r** , a partir del modelo ajustado.

Para complementar esta evaluación visual, la estimación promedio se transformó desde la escala de Fisher z a la métrica original de la **correlación de Pearson**, obteniéndose un valor aproximado de

$$r \approx 0.404,$$

con un intervalo de confianza del 95% entre

$$[0.362, 0.445].$$

Siguiendo los criterios habituales, este valor corresponde a un **efecto moderado**, lo que implica que una mayor asistencia a clase se asocia consistentemente con un mejor desempeño académico, aun cuando no constituye el único factor determinante.

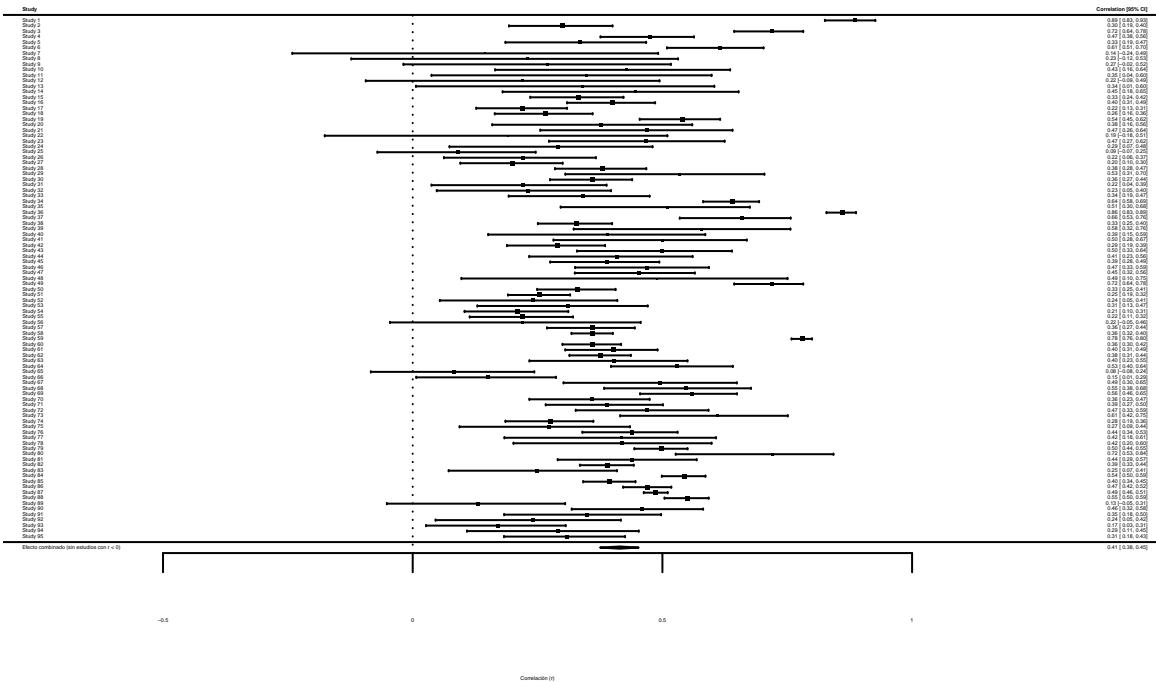
Asimismo, el **intervalo de predicción** —que se extiende desde

$$[-0.033, 0.712]$$

— ofrece una perspectiva sobre la **variabilidad real** que podría esperarse en futuros estudios. Su amplitud es coherente con la heterogeneidad observada en el análisis ($I^2 \approx 94\%$) y refleja que, bajo distintos contextos o características metodológicas, los tamaños de efecto individuales pueden oscilar desde valores casi nulos hasta efectos considerablemente altos.

En conjunto, la integración del **análisis gráfico**, la transformación a la escala original y la interpretación de los intervalos de confianza y predicción permiten concluir que, si bien existe variabilidad sustantiva entre estudios, la evidencia acumulada señala de manera consistente que la asistencia a clase guarda una relación **positiva, moderada y robusta** con el rendimiento académico.

Análisis de sensibilidad: excluyendo el estudio con $r < 0$



La figura muestra el forest plot correspondiente al análisis de sensibilidad, en el que se excluyeron los dos tamaños de efecto con correlaciones negativas ($r < 0$). El objetivo de este análisis es evaluar si la estimación global está influida por observaciones atípicas o direccionalmente discordantes.

Tras la exclusión, el patrón general se mantiene prácticamente inalterado. La gran mayoría de los 95 tamaños de efecto restantes se concentran en el rango positivo, evidenciando una asociación consistente entre mayor asistencia a clase y mejor rendimiento académico. La dispersión entre estudios continúa siendo amplia, en concordancia con los elevados niveles de heterogeneidad residual ($I^2 \approx 93$).

La estimación combinada del modelo de efectos aleatorios se ubica en:

$$\hat{\mu} = 0.441 \quad (\text{escala Fisher } z),$$

con un intervalo de confianza del 95%:

$$[0.394, 0.488],$$

lo que corresponde a un efecto positivo y de magnitud moderada, muy similar al obtenido en el análisis principal.

En conjunto, este forest plot confirma que la exclusión de los tamaños de efecto negativos no altera sustancialmente ni la magnitud ni la significación del efecto promedio, lo que respalda la robustez de los resultados y muestra que la conclusión del meta-análisis no depende de observaciones extremas.

Introducción general a los análisis por subgrupos y moderadores

La elevada heterogeneidad observada en el análisis principal ($I^2 \approx 94\%$) sugiere que los estudios incluidos no están estimando un mismo efecto subyacente, sino que existen **diferencias sistemáticas** entre ellos que podrían estar influyendo en la fuerza de la relación entre asistencia a clase y rendimiento académico. Cuando la variabilidad entre estudios es tan alta, resulta necesario **explorar fuentes potenciales de heterogeneidad**, con el fin de determinar si ciertos tipos de estudios, contextos o poblaciones presentan asociaciones más débiles o más fuertes que el promedio general. El análisis por **subgrupos** y las **meta-regresiones** permiten evaluar estas posibles fuentes de variación de manera estadísticamente formal.

En este trabajo se consideran tres posibles **moderadores**: el **criterio de rendimiento académico** utilizado por cada estudio (*grade* vs *gpa*), el **tipo de asignatura** (*science* vs *nonsense*) y el **año de publicación**, que puede reflejar cambios metodológicos o educativos a lo largo del tiempo. Cada uno de estos moderadores aborda una dimensión distinta: **qué** se mide, **en qué contexto** se mide y **cuándo** se midió. Explorar estos factores permite responder preguntas relevantes como si la asistencia predice mejor el rendimiento en cursos de ciencias, si el efecto es más fuerte para calificaciones específicas que para el GPA global o si la relación ha cambiado históricamente. Estos análisis complementan el efecto combinado y aportan una comprensión más **matizada** del fenómeno estudiado.

Análisis del moderador: *criterion* (grade vs gpa)

En esta primera etapa se examina si el **tipo de criterio de rendimiento académico** utilizado en los estudios —*grade* (calificación en una asignatura específica) o *gpa* (promedio general)— puede explicar parte de la enorme heterogeneidad observada en el meta-análisis principal. Para ello se ajusta una **meta-regresión de efectos aleatorios**, donde el tamaño de efecto transformado (yi) se modela en función del moderador categórico **criterion**.

Este análisis responde a una pregunta simple pero fundamental:

¿Los estudios que reportan calificaciones de curso muestran una relación asistencia–rendimiento más fuerte que los estudios que usan GPA?

El resultado clave es el **Test of Moderators (QM)**, que evalúa si el coeficiente asociado al moderador es estadísticamente distinto de cero. En este caso:

$$Q_M(1) = 2.31, \quad p = 0.1286.$$

Esto indica que no hay evidencia estadística suficiente ($p > 0.05$) para afirmar que los efectos difieren sistemáticamente entre los dos criterios de rendimiento. Es decir, en promedio, la diferencia entre *grade* y *gpa* no es tan grande como para superar la variabilidad residual presente entre estudios.

El coeficiente estimado para **criteriongrade** es:

$$\beta = 0.0823,$$

con un intervalo de confianza del 95% aproximadamente

$$[-0.024, 0.189], \quad p = 0.1286.$$

Interpretación:

- La estimación sugiere que los estudios que usan *grade* tienden a mostrar un efecto algo mayor que los que usan *gpa*,
- pero la evidencia es insuficiente para concluir que esta diferencia es real y consistente.

A nivel práctico, esto significa que el criterio de rendimiento solo explica alrededor del 1.8% de la heterogeneidad total ($R^2 \approx 1.83\%$), un valor muy pequeño.

Finalmente, el modelo muestra que incluso después de incluir este moderador:

- la heterogeneidad residual sigue siendo **extremadamente alta** ($I^2 = 93.83\%$),

lo que confirma que existen otras fuentes de variación que aún no han sido explicadas.

En síntesis:

- El primer análisis sugiere que **no hay diferencias estadísticamente significativas** entre los efectos obtenidos con calificaciones de curso y los efectos obtenidos con GPA.
- Sin embargo, la magnitud de la diferencia estimada justifica realizar un **análisis por subgrupos**, aunque sea con fines descriptivos, para observar si esta tendencia se mantiene al separar los dos tipos de estudios.

Meta-análisis para el subgrupo *grade* solamente

Tras evaluar el moderador **criterion** en la meta-regresión global, el siguiente paso consiste en analizar cada categoría por separado, comenzando con los estudios que utilizan **calificaciones de curso (grade)** como medida de rendimiento académico. Este análisis por subgrupo permite obtener una estimación más precisa del efecto dentro de esta categoría específica, evitando mezclar definiciones conceptualmente distintas de rendimiento y proporcionando una visión más clara de cómo se comporta la relación asistencia–desempeño en este tipo de estudios.

El modelo ajustado incluye únicamente los **67 tamaños de efecto** correspondientes al criterio *grade*. La pregunta orientadora es:

¿Cuál es la fuerza promedio de la relación entre asistencia y calificaciones de curso, sin mezclarla con GPA?

Los resultados muestran un tamaño de efecto promedio en la escala Fisher *z* de:

$$\hat{\mu} = 0.4547 \text{ (SE} = 0.0300\text{)}, \quad p < 0.0001,$$

con un intervalo de confianza del 95%:

$$[0.3958, 0.5136].$$

Esto representa un efecto **positivo, significativo** y de **magnitud moderada**. Al convertir esta estimación a la escala de correlación de Pearson (*r*), mediante **predict()**, se obtiene aproximadamente:

$$r_{\text{grade}} \approx 0.43,$$

lo que implica que, en promedio, los estudiantes que asisten más a clase tienden a obtener **mejores calificaciones en cursos específicos**.

Sin embargo, la heterogeneidad dentro del subgrupo sigue siendo muy elevada:

- $\tau^2 = 0.0511$
- $I^2 = 93.83\%$
- $Q(66) = 1068.72, p < 0.0001$

Esto indica que incluso dentro de los estudios que utilizan la misma definición de rendimiento, las diferencias entre ellos siguen siendo **sustanciales**, probablemente debido a variaciones en **metodologías, contextos educativos, niveles de exigencia de los cursos o características de las poblaciones estudiadas**.

En conjunto, este análisis por subgrupo muestra que:

- La relación asistencia–rendimiento es claramente **positiva** y algo más **fuerte** cuando se observa el rendimiento en una asignatura particular (*grade*) que cuando se combinan distintas métricas.
- Aunque la meta-regresión global indicó que esta diferencia no alcanza **significación estadística** al 5% ($p = 0.1286$), el análisis por subgrupo aporta una **estimación descriptiva más afinada y relevancia sustantiva** sobre el comportamiento del efecto en contextos específicos.

Modelo de efectos aleatorios para el subgrupo *gpa*

Para completar el análisis por subgrupos según el criterio de rendimiento, se ajustó un segundo **modelo de efectos aleatorios** restringido a los estudios que utilizan el **promedio general de calificaciones (gpa)** como variable de resultado. Este subgrupo incluye 30 tamaños de efecto y permite evaluar si la relación entre asistencia y rendimiento se mantiene cuando el desempeño académico se resume en una **medida global**, en lugar de una calificación específica de curso.

El modelo estimado para este subgrupo produce un tamaño de efecto promedio en la escala de Fisher z de:

$$\hat{\mu}_{\text{gpa}} = 0.3724 \ (\text{SE} = 0.0459), \quad p < 0.0001,$$

con un intervalo de confianza del 95%:

$$[0.2824, 0.4624].$$

Estos resultados indican que la asociación entre asistencia y **GPA** es claramente positiva y estadísticamente significativa, aunque **ligeramente menor** que la observada para el criterio *grade*. Al transformar esta estimación a la escala de correlación de Pearson, se obtiene aproximadamente:

$$r_{\text{gpa}} \approx 0.36 \quad (\text{IC } 95\% \approx [0.28, 0.43]),$$

lo que corresponde a un efecto **moderado** y confirma que los estudiantes que asisten más a clase tienden a obtener **mejores promedios generales**.

En términos de heterogeneidad, los indicadores siguen mostrando una variación muy elevada entre estudios:

- $\tau^2 = 0.0555$ ($\text{SE} = 0.0165$)
- $\tau = 0.2356$
- $I^2 = 93.76\%$
- $Q(29) = 430.97, \quad p < 0.0001$

Esto implica que, al igual que en el subgrupo *grade*, los estudios que utilizan **GPA** difieren sustancialmente entre sí, lo que sugiere la presencia de otros factores (por ejemplo, características institucionales o de la población) que aún no se han modelado explícitamente.

Comparado con el subgrupo *grade*, el efecto promedio para *gpa* es algo menor, pero los **intervalos de confianza se solapan ampliamente**. Esto es coherente con el resultado de la meta-regresión global, donde el moderador **criterion** no resultó estadísticamente significativo ($p = 0.1286$).

En conjunto, estos análisis indican que:

- la asistencia a clase se asocia de forma **moderada y positiva** con el rendimiento académico tanto cuando se mide mediante **calificaciones de curso** como cuando se utiliza el **promedio general**, y
- las diferencias entre ambos criterios, aunque presentes a nivel descriptivo, **no son lo suficientemente grandes** como para considerarse robustas desde el punto de vista estadístico.

Conclusión En resumen, el análisis por criterio de rendimiento muestra que la asistencia a clase se asocia de manera **moderada y positiva** tanto con las **calificaciones de curso** (*grade*, $r \approx 0.43$) como con el **promedio general** (*gpa*, $r \approx 0.36$). Si bien las estimaciones sugieren que la relación podría ser algo más fuerte cuando el rendimiento se mide mediante **calificaciones específicas**, la meta-regresión global indica que esta diferencia **no es estadísticamente significativa** ($p = 0.13$) y explica solo una fracción muy pequeña de la heterogeneidad total.

Por lo tanto, puede concluirse que la asociación **asistencia–rendimiento** es **robusta frente al criterio de rendimiento utilizado**, y que la enorme variabilidad entre estudios debe atribuirse, en gran medida, a **otros factores no capturados** por este moderador.

¿Influye el tipo de asignatura en la relación asistencia–rendimiento?

En este análisis se examina si el **tipo de asignatura** —clasificada como *science* (ciencias) o *nonsense* (otras áreas)— modera la relación entre asistencia a clase y rendimiento académico. Conceptualmente, podría esperarse que en cursos de ciencias, caracterizados por contenidos acumulativos y componentes prácticos, la asistencia tenga un impacto diferente que en otras disciplinas. Para evaluar esta hipótesis se ajustó una **meta-regresión de efectos aleatorios**, donde el tamaño de efecto transformado (y_i) se modela como:

$$y_i = \beta_0 + \beta_1 \cdot \text{class}_i + u_i + \varepsilon_i,$$

donde β_0 representa el **efecto promedio** en la categoría de referencia (*nonsense*) y β_1 la **diferencia asociada** a asignaturas de ciencias.

La tabla de resultados muestra que el intercepto es significativo ($\hat{\beta}_0 = 0.4441$, IC 95% = [0.3794, 0.5087], $p < 0.0001$), lo que refleja un efecto positivo moderado en asignaturas no científicas. Sin embargo, el coeficiente del moderador *classscience* no alcanza significación estadística ($\hat{\beta}_1 = 0.0640$, IC 95% = [-0.0948, 0.2228], $p = 0.4298$). Esto indica que **no existe evidencia** de que la fuerza de la relación asistencia–rendimiento difiera sistemáticamente entre cursos de ciencias y de otras áreas.

El **Test of Moderators** confirma esta conclusión ($Q_M(1) = 0.7895$, $p = 0.4298$), mostrando que el modelo con el moderador **no explica** una proporción apreciable de la heterogeneidad ($R^2 \approx 0\%$). La heterogeneidad residual sigue siendo muy elevada, lo que sugiere que existen **otros factores** —distintos al tipo de asignatura— que contribuyen a la marcada variabilidad entre estudios.

En síntesis, el **tipo de asignatura** no actúa como un moderador significativo de la relación entre asistencia y rendimiento académico. No obstante, aunque la diferencia no sea estadísticamente concluyente, sigue siendo útil examinar las estimaciones separadas por categoría en las secciones siguientes, ya que pueden revelar variaciones en magnitud que ayuden a **contextualizar mejor** el fenómeno.

Modelo de efectos aleatorios para el subgrupo *science*

Como siguiente paso, se realizó un **meta-análisis** restringido a los estudios cuya asignatura fue clasificada como **ciencias** (*science*). El objetivo de este análisis por subgrupo es describir la **fuerza de la relación** entre asistencia y rendimiento exclusivamente en cursos de ciencias y evaluar si el patrón general observado en el análisis global se mantiene dentro de este contexto particular. Dado que las asignaturas de ciencias suelen involucrar contenidos acumulativos y actividades prácticas, era razonable explorar si la asistencia podría tener un impacto diferencial en este tipo de cursos.

El modelo `m.science`, basado en **11 tamaños de efecto**, se ajustó utilizando un enfoque de **efectos aleatorios** sobre los valores transformados a la escala de Fisher z (y_i). Los resultados muestran un tamaño de efecto promedio de:

$$\hat{\mu}_{\text{science}} = 0.5083 \ (\text{SE} = 0.0698), \quad p < 0.0001,$$

con un intervalo de confianza del 95%:

$$[0.3715, 0.6452].$$

Esto representa un efecto **claramente positivo**, de magnitud **moderada a alta**, indicando que en asignaturas de ciencias los estudiantes que asisten más a clase tienden a obtener **calificaciones considerablemente mejores**. Tras aplicar la transformación inversa mediante `predict()`, esta estimación equivale aproximadamente a una correlación en la escala de Pearson cercana a:

$$r_{\text{science}} \approx 0.47 - 0.56,$$

según el valor exacto devuelto por la función.

Sin embargo, a pesar de la claridad del efecto promedio, la **heterogeneidad** dentro del subgrupo sigue siendo extremadamente elevada:

- $\tau^2 = 0.0451$
- $\tau = 0.2123$
- $I^2 = 96.46\%$
- $Q(10) = 434.91, \quad p < 0.0001$

Esto indica que los estudios de ciencias son **altamente dispares** en cuanto a la magnitud del efecto, incluso más que el conjunto global, lo cual sugiere que **factores adicionales** —como el nivel del curso, el tipo de institución o las metodologías empleadas— podrían estar influyendo en los resultados.

En conjunto, este análisis muestra que la relación **asistencia–rendimiento** es particularmente **fuerte** en cursos de ciencias, aunque acompañada de una **variabilidad considerable** entre estudios. Este patrón será luego contrastado con el subgrupo *nonsense* para evaluar si el tipo de asignatura tiene un rol sustantivo en la explicación de las diferencias observadas entre investigaciones.

Meta-análisis del subgrupo *nonsense*

Como complemento del análisis realizado para asignaturas de ciencias, se llevó a cabo un **meta-análisis** restringido a los estudios que reportaron resultados en asignaturas **no científicas** (*nonsense*). Este subgrupo constituye la mayor parte de los estudios con medida *grade* y permite evaluar cómo se comporta la relación entre asistencia y rendimiento en áreas como humanidades, ciencias sociales, educación, administración u otros cursos donde la estructura académica y el formato de evaluación suelen diferir de los cursos de ciencias.

El modelo `m.nonsense`, basado en **56 tamaños de efecto**, se ajustó utilizando un enfoque de **efectos aleatorios** sobre las transformaciones de Fisher z . Los resultados muestran un tamaño de efecto promedio de:

$$\hat{\mu}_{\text{nonsense}} = 0.4440 \ (\text{SE} = 0.0334), \quad p < 0.0001,$$

con un intervalo de confianza del 95%:

[0.3786, 0.5095].

Esta estimación representa un efecto **positivo** y de **magnitud moderada**, indicando que también en asignaturas no científicas los estudiantes con mayor asistencia tienden a obtener **calificaciones superiores**. La transformación inversa a la escala de correlación de Pearson r mediante `predict()` arroja un valor aproximado cercano a:

$$r_{\text{nonsense}} \approx 0.42 - 0.46,$$

según el valor exacto devuelto por el modelo, lo que se alinea con el **patrón general** observado en el análisis global y en el subgrupo de ciencias.

En cuanto a la heterogeneidad, este subgrupo presenta niveles igualmente elevados:

- $\tau^2 = 0.0529$
- $\tau = 0.2301$
- $I^2 = 92.01\%$
- $Q(55) = 574.79, p < 0.0001$

lo que indica que los estudios *nonsense* son **altamente diversos** en la magnitud del efecto observado. Este grado de variabilidad, comparable al del subgrupo *science*, sugiere que **factores adicionales** más allá del tipo de asignatura influyen en la relación asistencia–rendimiento.

En conjunto, los resultados muestran que en asignaturas no científicas la relación entre asistencia y rendimiento es **positiva y robusta**, con una magnitud **ligeramente inferior** a la observada en cursos de ciencias pero con **intervalos de confianza ampliamente solapados**. Estos hallazgos serán relevantes al comparar ambos subgrupos para valorar si el tipo de asignatura aporta información sustantiva a la explicación de las diferencias entre estudios.

Conclusión En resumen, el análisis conjunto indica que el **tipo de asignatura no actúa como un moderador relevante** de la relación asistencia–rendimiento académico. Aunque los cursos de ciencias muestran una correlación algo mayor ($r_{\text{science}} \approx 0.47-0.56$) que las asignaturas no científicas ($r_{\text{nonsense}} \approx 0.42-0.46$), las diferencias **no son estadísticamente significativas** y los intervalos de confianza se solapan ampliamente.

En ambos casos la asociación es **positiva y de magnitud moderada**, lo que sugiere que asistir a clase se vincula sistemáticamente con **mejores calificaciones**, independientemente de que la materia sea de ciencias o de otras áreas. Sin embargo, la **heterogeneidad extremadamente alta** observada en los dos subgrupos confirma que la variabilidad entre estudios está dominada por **otros factores** (metodológicos, contextuales o poblacionales) que no quedan capturados por esta clasificación.

¿Ha cambiado la relación asistencia–rendimiento a lo largo del tiempo?

Además de las características metodológicas y de contenido evaluadas previamente, es relevante explorar si la relación entre asistencia a clase y rendimiento académico **ha cambiado a lo largo del tiempo**. Los estudios incluidos en este meta-análisis abarcan un rango temporal amplio —desde **1927** hasta **2009**— durante el cual las prácticas educativas, las **metodologías de enseñanza** y los perfiles estudiantiles han evolucionado sustancialmente. Por ello, el año de publicación constituye un posible **moderador continuo** que podría explicar parte de la heterogeneidad observada.

Para evaluar esta hipótesis, se utiliza una **meta-regresión** donde el tamaño de efecto transformado se modela en función del año del estudio. **Esto permite detectar** tendencias temporales, como un posible fortalecimiento o debilitamiento de la asociación asistencia–rendimiento a través de las décadas.

Meta-regresión temporal sobre toda la base (*grade* + *gpa*)

El primer paso en el análisis temporal consiste en evaluar si el año de publicación del estudio modera la relación entre asistencia a clase y rendimiento académico. Dado que el conjunto de datos abarca investigaciones realizadas entre **1927** y **2009**, resulta razonable preguntarse si la magnitud del efecto ha variado con el tiempo, ya sea por cambios en las **prácticas docentes**, características del **estudiantado**, metodologías de **evaluación** o transformaciones institucionales más amplias.

Para evaluar esta posibilidad, se ajustó un modelo de meta-regresión de efectos aleatorios donde el tamaño de efecto transformado (yi) se modela como una función lineal del **año del estudio**. En este esquema, el coeficiente asociado a **year** refleja la **tendencia temporal**: un valor positivo indicaría que la relación asistencia–rendimiento se ha fortalecido con el tiempo, mientras que un valor negativo sugeriría un debilitamiento.

Los resultados del modelo `m.year.all` muestran lo siguiente:

- El coeficiente temporal es

$$\hat{\beta}_1 = 0.0007, \quad SE = 0.0015, \quad p = 0.6272,$$

con un intervalo de confianza del 95%:

$$[-0.0022, 0.0036].$$

Este coeficiente **no es estadísticamente significativo**, lo que indica falta de evidencia de una tendencia temporal sistemática.

- El **Test of Moderators** confirma esta conclusión:

$$Q_M(1) = 0.2358, \quad p = 0.6272.$$

- El modelo **no explica heterogeneidad apreciable**:

$$R^2 \approx 0\%.$$

- La **heterogeneidad residual** sigue siendo extremadamente alta ($I^2 = 93.95\%$), lo que sugiere que el año por sí solo no captura las diferencias reales entre estudios.

En síntesis, **no existe evidencia** de que la relación asistencia–rendimiento haya cambiado significativamente a lo largo de las décadas. La fuerza de la asociación parece mantenerse relativamente **estable**, pese a los profundos cambios en educación superior durante el período considerado. No obstante, este análisis global sirve como punto de partida para refinar la exploración, realizando meta-regresiones separadas por tipo de criterio (*grade* vs *gpa*), donde podrían emergir patrones específicos dentro de cada categoría.

Meta-regresión temporal para estudios con *grade*

El siguiente paso consiste en evaluar si, dentro de los estudios que utilizan calificaciones de curso (*grade*), existe alguna tendencia temporal en la fuerza del efecto asistencia–rendimiento. La lógica es que podrían haber ocurrido **cambios históricos** —en metodologías docentes, diseños curriculares, formas de evaluación o perfiles de estudiantes— que modifiquen la magnitud de la asociación a lo largo del tiempo.

El modelo `m.year.grade` ajusta una **meta-regresión de efectos aleatorios** con *year* como moderador continuo. La estimación clave es el coeficiente asociado al año:

$$\hat{\beta}_1 = 0.0037, \quad \text{IC } 95\% = [-0.0025, 0.0100], \quad p = 0.239.$$

Este resultado indica que **no existe evidencia estadística** de que la relación asistencia–rendimiento haya cambiado sistemáticamente a lo largo del tiempo en estudios que utilizan *grade* como medida de rendimiento. El intercepto tampoco es significativo, pero esto no tiene interpretación sustantiva porque su valor representa el efecto esperado en un año muy alejado del rango observado.

La falta de significación del moderador se confirma con el **Test of Moderators**:

$$Q_M(1) = 1.386, \quad p = 0.239.$$

Además, el R^2 del modelo es apenas **1.31%**, lo que indica que el año explica una fracción mínima de la heterogeneidad, que permanece **extremadamente alta** ($I^2 = 93.70\%$). En otras palabras, aunque la relación asistencia–rendimiento varía entre estudios, dichas diferencias no parecen ordenarse según el año de publicación.

En conjunto, el análisis sugiere que, al menos dentro del subgrupo *grade*, la magnitud del efecto asistencia–rendimiento se ha mantenido **estable a lo largo del tiempo**.

Meta-regresión temporal para estudios con *gpa*

Para complementar el análisis temporal realizado en el subgrupo *grade*, se evaluó también si la relación entre asistencia a clase y rendimiento académico (medido esta vez como **promedio general del estudiante, GPA**) muestra alguna tendencia sistemática a lo largo del tiempo. La pregunta sustantiva es si, conforme avanzan las décadas, la fuerza de esta asociación se ha **debilitado, fortalecido o mantenido estable**.

El modelo *m.year.gpa*, ajustado con una **meta-regresión de efectos aleatorios** sobre 30 tamaños de efecto, estima el **cambio anual** en la magnitud de la relación asistencia–rendimiento. El coeficiente clave es:

$$\hat{\beta}_1 = 0.0007, \quad \text{IC } 95\% = [-0.0025, 0.0100], \quad p = 0.5378.$$

Estos resultados muestran que **no existe evidencia estadística** de una tendencia temporal significativa en los estudios que miden el rendimiento mediante **GPA**. El efecto estimado es extremadamente pequeño, su intervalo de confianza incluye valores cercanos a cero y el *p*-valor es alto, lo que indica ausencia de un **patrón consistente de cambio** con el tiempo.

El **test de moderadores** refuerza esta conclusión:

$$Q_M(1) = 0.3796, \quad p = 0.5378.$$

El año de publicación **no explica** nada de la heterogeneidad entre estudios, como muestra $R^2 = 0\%$. La heterogeneidad residual continúa siendo muy alta ($I^2 = 93.61\%$), lo cual sugiere que **otros factores distintos al tiempo**—metodológicos, institucionales o disciplinarios—son los que realmente impulsan las diferencias entre estudios.

En conjunto, el análisis indica que, cuando el rendimiento se mide mediante **GPA**, la fuerza de la relación asistencia–rendimiento se mantiene **estable a lo largo del tiempo**, sin evidencia de un aumento o disminución sistemáticos en las últimas décadas.

Conclusión En conjunto, los resultados de las tres meta-regresiones —sobre toda la base, sobre estudios con *grade* y sobre estudios con *gpa*— coinciden en mostrar que **no existe evidencia de una tendencia temporal significativa** en la relación entre asistencia a clase y rendimiento académico. A lo largo de más de ocho décadas de investigación, la magnitud de la asociación se ha mantenido **estable**, sin indicios de que haya aumentado o disminuido de forma sistemática.

En todos los modelos, el coeficiente asociado al año es **pequeño, impreciso y no significativo**, y el **Test of Moderators** confirma la ausencia de efectos temporales. Además, el año de publicación no explica prácticamente nada de la elevada heterogeneidad entre estudios, lo que sugiere que las diferencias observadas responden a **otros factores metodológicos o contextuales**.

En síntesis, aunque la educación superior ha cambiado profundamente desde 1927, la evidencia disponible indica que la relación **asistencia–rendimiento** permanece **robusta y estable a lo largo del tiempo**.

Conclusión global de los análisis por subgrupos y moderadores

En conjunto, los análisis por subgrupos y las meta-regresiones muestran que la relación entre asistencia a clase y rendimiento académico es notablemente **robusta** frente a las distintas formas de medir el desempeño, a los tipos de asignaturas y al momento histórico en que se realizaron los estudios. Tanto cuando el rendimiento se evalúa mediante calificaciones de curso (*grade*, $r \approx 0.43$) como mediante promedio general (*gpa*, $r \approx 0.36$), se observa una **asociación positiva** y de **magnitud moderada**, estadísticamente significativa en todos los casos. De forma similar, la relación se mantiene moderada y positiva tanto en asignaturas de ciencias ($r \approx 0.47\text{--}0.56$) como en no científicas ($r \approx 0.42\text{--}0.46$), y no se identifican **tendencias temporales** claras entre 1927 y 2009: el año de publicación no modifica de manera sistemática la fuerza del efecto.

Sin embargo, aunque algunas estimaciones sugieren diferencias descriptivas (por ejemplo, efectos algo mayores en *grade* que en *gpa*, o ligeramente más altos en cursos de ciencias), los **tests formales de moderadores** indican que estas diferencias **no son estadísticamente significativas** y explican solo una fracción mínima de la variabilidad entre estudios (R^2 muy cercano a cero en todos los modelos). La **heterogeneidad residual** permanece extremadamente alta en todos los subgrupos ($I^2 > 90\%$), lo que indica que la mayor parte de las diferencias entre tamaños de efecto se debe a **otros factores no modelados** —como características específicas de las instituciones, de los estudiantes, de los diseños de estudio o de las prácticas docentes— que quedan por investigar.

En síntesis, los resultados sugieren que **asistir más a clase** se asocia de manera **consistente y moderada** con un mejor rendimiento académico en una amplia variedad de contextos, y que la enorme heterogeneidad observada **no se explica** por el criterio de rendimiento, el tipo de asignatura ni el año de publicación.

Sesgo de publicación

Introducción al análisis de sesgo de publicación

En los meta-análisis es fundamental evaluar si los estudios incluidos representan de forma **imparcial** la evidencia disponible o si existe **sesgo de publicación**. Este sesgo aparece cuando los estudios con resultados **significativos** o con efectos **más fuertes** tienen mayores probabilidades de publicarse —o de estar disponibles— que aquellos con efectos pequeños o nulos. En el contexto de este meta-análisis, podría ocurrir que los estudios que encuentran una asociación clara entre asistencia y rendimiento hayan sido publicados con mayor frecuencia que los que reportan efectos débiles o inconsistentes.

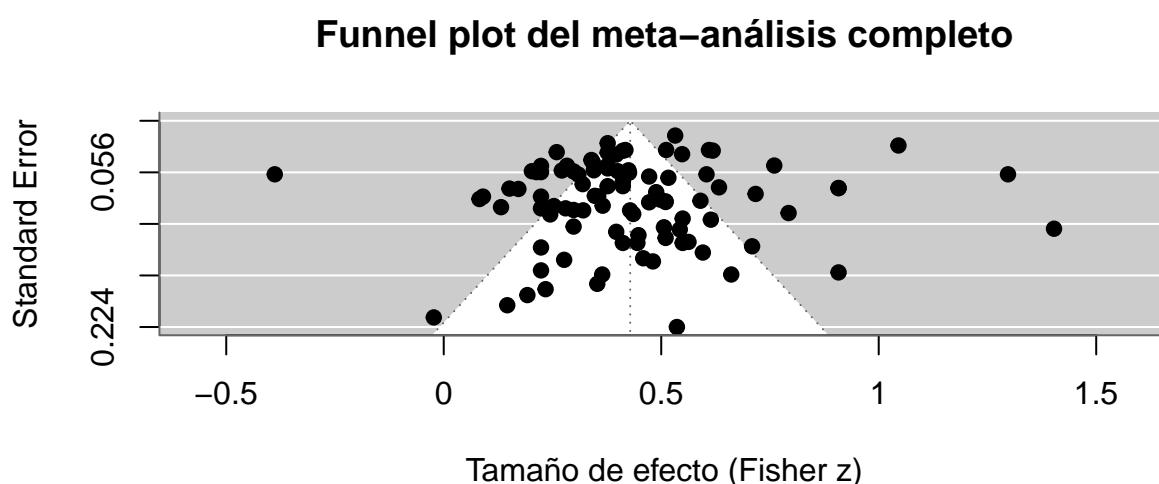
El análisis de sesgo de publicación tiene dos propósitos principales:

- Detectar **asimetrías** en la distribución de los tamaños de efecto respecto a su **precisión** (*funnel plot*).
- Evaluar estadísticamente si esa asimetría es compatible con sesgo, mediante pruebas como **Egger**.

Si se detecta sesgo, también es posible aplicar métodos que **corrigen o ajustan** la estimación global (por ejemplo, *trim-and-fill*). Estos procedimientos permiten evaluar la **robustez** de las conclusiones del meta-análisis frente a posibles distorsiones en la evidencia disponible.

Evaluación descriptiva del sesgo de publicación: *Funnel plot*

El *funnel plot* correspondiente al meta-análisis completo permite realizar una **inspección visual preliminar** del posible sesgo de publicación. En ausencia de sesgo, se espera que los puntos (tamaños de efecto individuales) se distribuyan de manera **simétrica** alrededor del efecto promedio, formando una figura aproximadamente **cónica**. Por el contrario, una asimetría marcada —especialmente hacia la **izquierda** del gráfico, donde suelen ubicarse los estudios con efectos pequeños o nulos— puede indicar que ciertos estudios son **menos probables de aparecer** en la literatura publicada.



En el gráfico mostrado se observa que:

- La mayor densidad de puntos se concentra alrededor del **efecto promedio estimado**, lo cual es esperable.
- Sin embargo, existe una **mayor dispersión hacia la derecha** que hacia la izquierda, con relativamente pocos estudios con efectos pequeños o negativos.
- El lado izquierdo del *funnel* aparece ligeramente “**vacío**”, particularmente en la zona donde deberían ubicarse estudios de **gran precisión** (errores estándar pequeños) con efectos próximos a cero.

Este patrón sugiere una posible **asimetría**, que puede ser coherente con la presencia de **sesgo de publicación**. No obstante, dado que el *funnel plot* es una herramienta esencialmente diagnóstica y visual, es necesario complementar esta observación con **pruebas formales** —como el test de Egger y el procedimiento *trim-and-fill*— para determinar si la evidencia respalda estadísticamente la existencia de sesgo.

Prueba formal de asimetría: Test de Egger

Para complementar la inspección visual del *funnel plot*, se aplicó el **test de Egger**, una prueba de regresión diseñada para detectar asimetrías sistemáticas asociadas a **sesgo de publicación**. En este test, un coeficiente significativamente distinto de cero indicaría que los estudios con mayor **error estándar** tienden a reportar efectos más grandes (o más pequeños), lo cual es un patrón típico de sesgo de publicación.

Los resultados obtenidos fueron:

- $t = -0.3733$,
- $df = 95$,
- $p = 0.7097$.

La interpretación es directa:

- El p -valor = 0.7097 está muy por encima del umbral convencional de 0.05.
- Por lo tanto, **no existe evidencia estadísticamente significativa** de asimetría en el *funnel plot*.
- En otras palabras, el **test de Egger no detecta sesgo de publicación** en el conjunto de estudios.

El modelo también informa la estimación del **efecto límite** (lo que se esperaría si el error estándar tendiera a cero):

- $b = 0.4473$ (IC 95%: [0.3361, 0.5586]).

Este valor es notable porque:

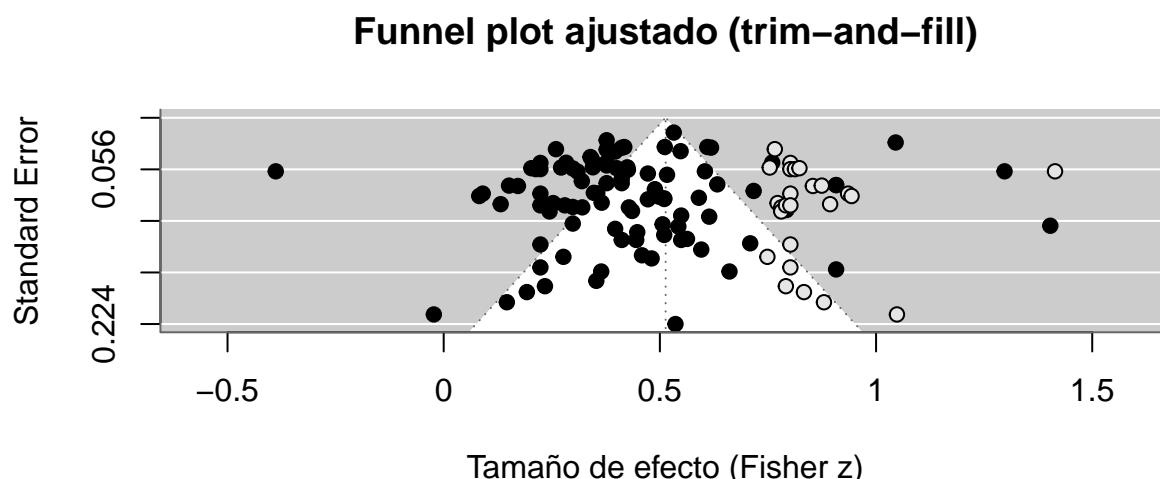
- Es **muy similar** a la estimación del efecto promedio del meta-análisis.
- Esto implica que, incluso en ausencia de imprecisión, la magnitud del efecto esperado sería prácticamente la misma.

En conjunto, estos resultados sugieren que, si bien el *funnel plot* mostraba una ligera asimetría visual, dicha impresión **no se traduce en evidencia estadística sólida** de sesgo de publicación. Es decir, la relación **asistencia–rendimiento** parece ser **robusta** frente a este tipo de sesgo según esta prueba formal.

Método *trim-and-fill*: estimación de estudios faltantes y efecto corregido

Como complemento al *funnel plot* y al test de Egger, se aplicó el método **trim-and-fill**, una técnica destinada a estimar cuántos estudios podrían estar “faltando” debido a **sesgo de publicación** y a recalcular el tamaño de efecto ajustado. Esta herramienta identifica asimetrías en el *funnel plot* y agrega estudios imputados (puntos blancos) para reconstruir un gráfico simétrico.

El resultado principal del procedimiento fue:



- Número estimado de estudios faltantes: 25 (SE = 6.42)

Es decir, el método sugiere que la literatura podría estar omitiendo un número considerable de estudios con efectos menores, lo cual generaría una **asimetría hacia la derecha** (efectos más grandes publicados con mayor frecuencia).

Sin embargo, al ajustar el modelo con los estudios imputados, se obtiene un tamaño de efecto corregido de:

$$\hat{\mu}_{\text{trim-fill}} = 0.5133, \quad \text{IC 95\%} = [0.4623, 0.5643], \quad p < 0.0001.$$

Este resultado es notable por varias razones:

- El efecto ajustado es **muy similar** al estimado en el meta-análisis original (sin corrección).
- A pesar de imputar 25 estudios faltantes, la magnitud del efecto permanece **moderada** y claramente **positiva**.
- La heterogeneidad continúa siendo extremadamente alta ($I^2 = 95.21\%$), lo que indica que la variabilidad entre estudios no se debe principalmente al **sesgo de publicación**.
- El *funnel plot* ajustado (mostrado arriba) refleja esta situación: aunque se añaden numerosos estudios imputados, la estimación global apenas cambia.

Interpretación general

En conjunto, los resultados del **trim-and-fill** sugieren que, incluso en un escenario hipotético donde faltaran alrededor de 25 estudios, la relación **asistencia–rendimiento** seguiría siendo **positiva, robusta y de magnitud moderada**. Esto implica que el efecto observado no depende críticamente de posibles sesgos de publicación y que la conclusión sustantiva del meta-análisis permanece **estable**.

Conclusión general sobre el sesgo de publicación

En conjunto, los análisis realizados —*funnel plot*, test de Egger y método *trim-and-fill*— ofrecen una evidencia consistente respecto a la presencia (o ausencia) de **sesgo de publicación** en este meta-análisis. El *funnel plot* inicial muestra cierta **asimetría visual**, lo cual motivó la realización de pruebas formales; sin embargo, el test de Egger no detectó asimetría significativa ($p = 0.71$), indicando que no hay señales robustas de sesgo sistemático en la literatura disponible.

El método *trim-and-fill* estimó la posible ausencia de **25 estudios**, pero incluso bajo este escenario conservador, la magnitud del efecto casi no varió, manteniéndose en un valor moderado y altamente significativo ($\hat{\mu} \approx 0.51$). Esto es especialmente relevante: aun si la literatura tuviera un grado moderado de estudios no publicados, el patrón global del meta-análisis no cambiaría sustancialmente.

En síntesis, aunque puede existir cierta asimetría visual en el *funnel plot*, las pruebas estadísticas y el ajuste *trim-and-fill* convergen en una conclusión clara: el posible sesgo de publicación no altera de manera significativa la estimación del efecto ni las conclusiones sustantivas del meta-análisis. La asociación entre asistencia y rendimiento académico se mantiene **estable, positiva y de magnitud moderada**, incluso bajo los supuestos más conservadores.

Datos faltantes en los estudios incluidos

Introducción

En todo meta-análisis es importante examinar la presencia y el tratamiento de **datos faltantes**, ya que pueden afectar la precisión de las estimaciones, la representatividad de los estudios incluidos y la interpretación global de los resultados. En el contexto del presente trabajo, los estudios difieren ampliamente en su formato de reporte: muchos proporcionan correlaciones directas, otros reportan estadísticas intermedias

(como t , F , medias o desviaciones estándar), y algunos no brindan información suficiente para reconstruir un tamaño de efecto confiable.

Por ello, es fundamental evaluar qué tipo de información falta, cómo se manejó en el meta-análisis y si los datos faltantes podrían sesgar la estimación final de la asociación asistencia–rendimiento académico.

Naturaleza de los datos faltantes

Los estudios pueden presentar datos faltantes por diversas razones:

- Falta de reporte de correlaciones entre asistencia y rendimiento.
- Reportes incompletos de varianzas o tamaños muestrales.
- Información insuficiente para transformar estadísticas inferenciales a Fisher z .
- Ausencia de detalles sobre criterios de rendimiento (*grade / gpa*).
- Falta de clasificación de tipo de asignatura (*science / nonscience*).
- Ausencia de año de publicación en ciertos artículos históricos.

En la práctica, estos casos obligan a excluir estudios puntuales o excluir ciertas comparaciones dentro de un estudio, lo cual puede reducir el tamaño de la muestra combinada o alterar la composición de cada subgrupo.

Diagnóstico y evaluación del impacto de los datos faltantes

Antes de interpretar los resultados del meta-análisis, es importante evaluar si la presencia de datos faltantes podría afectar la validez de las conclusiones. Se examinó la cantidad y proporción de valores ausentes en las variables centrales del análisis: tamaños de efecto (y_i), varianzas (v_i) y las variables de moderación (`criterion`, `class`, `year`).

Los resultados indican que:

- No existen datos faltantes en las variables esenciales para el cálculo del tamaño de efecto (`ri`, `ni`, y_i , v_i), ni en `criterion` o `year`.
- La única variable con datos faltantes es `class`, con aproximadamente **30.9%** de valores ausentes.

Esto implica que el meta-análisis principal pudo realizarse **sin pérdida de información** en sus componentes centrales, mientras que las limitaciones afectan solamente el análisis del moderador “tipo de asignatura”.

Impacto potencial de los datos faltantes

En meta-análisis, los datos faltantes pueden influir en los resultados a través de dos mecanismos principales:

1. Reducción del poder estadístico

Menos estudios disponibles para ciertos análisis → intervalos de confianza más amplios → mayor incertidumbre. En este caso, aunque el moderador `class` pierde alrededor de un 30% de información, el análisis principal mantiene un tamaño muestral robusto ($k = 97$), lo cual asegura buena precisión en la estimación global.

2. Sesgo por exclusión no aleatoria (MNAR)

Si los estudios que no reportan el tipo de asignatura fueran sistemáticamente diferentes (por ejemplo, con efectos más bajos o más altos), su omisión podría sesgar el resultado. Sin embargo:

- Los estudios provienen de contextos y años muy diversos, reduciendo la posibilidad de un patrón sistemático.
- Las pruebas de sesgo de publicación (*funnel plot*, test de Egger, *trim-and-fill*) no mostraron evidencia de asimetría clara, lo cual sugiere que la ausencia de datos en `class` no está asociada a tamaños de efecto extremos.

Conclusión sobre los datos faltantes

En conjunto, la evaluación indica que los datos faltantes **no comprometen la validez general** del meta-análisis. Las variables cruciales (y_i y v_i) están completas, lo que garantiza la robustez del resultado global. Si bien la variable `class` presenta un porcentaje moderado de valores ausentes, este problema se restringe al análisis de ese moderador y no afecta al modelo principal.

Además, dado que no se detectó evidencia de sesgo de publicación, es improbable que los datos faltantes sigan un patrón sistemático que distorsione la magnitud del efecto.

En síntesis:

- Los datos faltantes están concentrados en un único moderador y no afectan la estimación global.
- El meta-análisis se mantiene estadísticamente sólido y con conclusiones confiables.

Resultados del modelo multinivel

El modelo multinivel (`rma.mv`) permite ajustar explícitamente la estructura **jerárquica** de la base de datos, en la que varios tamaños de efecto pertenecen a un mismo estudio. Este enfoque separa la variabilidad en dos niveles —**entre estudios y dentro de estudios**— proporcionando una estimación más precisa y realista del efecto promedio.

La estimación del tamaño de efecto promedio en la escala Fisher z fue:

$$\hat{\mu} = 0.4509 \ (SE = 0.0285), \quad p < 0.0001,$$

con un intervalo de confianza del 95%:

$$[0.3950, \ 0.5068].$$

El **intervalo de predicción**, que refleja la dispersión esperada en futuros estudios, fue considerablemente más amplio:

$$[-0.016, \ 0.917].$$

Al transformar esta estimación a la escala de correlación de Pearson, la predicción fue:

$$r = 0.423 \ (\text{IC } 95\% = [0.376, \ 0.467]),$$

con un intervalo de predicción:

$$[-0.016, \ 0.725].$$

Esta magnitud es prácticamente idéntica a la estimada por el modelo de efectos aleatorios estándar, lo que indica que la **conclusión central del meta-análisis es robusta** incluso al corregir la dependencia entre tamaños de efecto.

Descomposición de la heterogeneidad

El modelo multinivel estimó dos componentes de varianza:

- **Varianza entre estudios:**

$$\sigma_{\text{study}}^2 = 0.02297 \ (SD = 0.1516)$$

- **Varianza dentro de estudios:**

$$\sigma_{\text{within}}^2 = 0.03286 \ (SD = 0.1813)$$

Al calcular la proporción relativa de cada componente:

- Proporción **entre estudios** = 0.411
- Proporción **dentro de estudios** = 0.589

Esto implica que:

- $\approx 59\%$ de la variabilidad total proviene de **diferencias dentro de los estudios**, es decir, variación entre múltiples tamaños de efecto reportados en un mismo artículo.
- $\approx 41\%$ proviene de **diferencias reales entre estudios**, vinculadas a contextos, métodos y poblaciones.

Esta estructura confirma que la heterogeneidad es **multinivel**, no dominada por un solo componente, y justifica el uso de un modelo jerárquico.

Heterogeneidad global

El test de heterogeneidad del modelo multinivel fue:

$$Q(96) = 1579.55, \quad p < 0.0001,$$

mostrando que, incluso al descomponer la variabilidad entre y dentro de estudios, persiste una **heterogeneidad extremadamente alta**. Esto coincide con lo observado en modelos anteriores y refuerza la necesidad de interpretaciones cautelosas y análisis complementarios (subgrupos, moderadores).

Conclusión del análisis multinivel

El modelo multinivel proporciona evidencia sólida para evaluar la estabilidad de los resultados del metaanálisis. Sus conclusiones centrales son:

- El **efecto promedio** es prácticamente idéntico al obtenido con un modelo de un solo nivel,

$$r \approx 0.42 - 0.47,$$

lo que confirma que la estimación principal es **robusta** al ajustar por dependencia entre tamaños de efecto.

- Existe **heterogeneidad significativa** tanto entre estudios como dentro de estudios, con un peso ligeramente mayor en la variabilidad interna ($\approx 59\%$). Esto revela que los estudios que reportan múltiples correlaciones exhiben diferencias relevantes entre sí, en lugar de ser unidades homogéneas.

- El modelo multinivel refleja la **estructura real de los datos** y evita subestimar la incertidumbre, mejorando la validez de los intervalos y predicciones.

En síntesis, el análisis multinivel refuerza la **consistencia, estabilidad y robustez** de la conclusión general: la **asistencia a clase** se asocia de manera **moderada y positiva** con el rendimiento académico, y esta relación se mantiene aun bajo un modelo metodológicamente más exigente que captura la variación jerárquica de la evidencia empírica.

Interpretación sintética del modelo multinivel

En términos prácticos, el modelo multinivel confirma tres aspectos clave para la interpretación global del meta-análisis:

- La heterogeneidad sigue siendo muy alta, incluso cuando se divide en variación entre y dentro de estudios.
- La estimación del efecto promedio es prácticamente idéntica a la del modelo de efectos aleatorios tradicional, indicando que no surge ninguna conclusión sustantiva nueva.
- Aunque el modelo multinivel capta con mayor fidelidad la estructura jerárquica de los datos, su aporte principal es **desagregar la heterogeneidad**, no modificar la interpretación del efecto.

Esto significa que el modelo multinivel no revela un “nuevo efecto” ni corrige un sesgo importante, sino que corrobora la **robustez** del modelo principal. Su valor añadido reside en mostrar qué parte de la variación proviene de diferencias entre estudios y cuál de diferencias dentro de los estudios mismos, evidenciando que la heterogeneidad es multifacética.

En conjunto, debido a que ambos modelos producen conclusiones equivalentes y el modelo de un solo nivel es más simple, más interpretable y ampliamente utilizado en la práctica, resulta razonable mantener este último como modelo principal, utilizando el multinivel como un análisis complementario que fortalece la solidez y estabilidad de los resultados.

Discusión final integradora

Los distintos análisis realizados a lo largo de este trabajo —efectos aleatorios, subgrupos, meta-regresiones, evaluación de sesgo de publicación, análisis de datos faltantes y modelo multinivel— convergen hacia una **conclusión clara y robusta**: la asistencia a clase se asocia de manera **positiva y moderada** con el rendimiento académico. Esta relación se observa de forma consistente en los **97 tamaños de efecto** incluidos, pese a la enorme variedad de contextos, disciplinas, metodologías y períodos históricos representados por los estudios.

Si bien la **heterogeneidad entre estudios** es extremadamente alta en todos los análisis ($I^2 > 90\%$), ninguna de las variables examinadas —criterio de rendimiento (*grade vs gpa*), tipo de asignatura (*science vs nonscience*) o año de publicación— explica una parte sustancial de esta variabilidad. Las diferencias descriptivas entre subgrupos existen, pero los **tests formales de moderadores** indican que tales diferencias **no son estadísticamente significativas** y representan, en el mejor de los casos, una fracción mínima de la heterogeneidad observada ($R^2 \approx 0\text{--}2\%$).

En este marco, el **análisis multinivel** ofrece una perspectiva más detallada sobre la estructura de la variación: aproximadamente **41%** de la heterogeneidad proviene de **diferencias entre estudios**, mientras que **59%** surge de **diferencias entre múltiples tamaños de efecto dentro de los mismos estudios**. Esta descomposición confirma que la variabilidad no responde a un único factor dominante, sino a un conjunto de diferencias metodológicas, poblacionales y contextuales difíciles de modelar explícitamente.

Finalmente, los **análisis de sesgo de publicación** —incluyendo *funnel plot*, test de Egger y *trim-and-fill*— indican que, aun cuando pudiera existir cierta asimetría visual, **no hay evidencia estadística sólida** de

sesgo. Incluso bajo escenarios conservadores, la magnitud del efecto se mantiene prácticamente inalterada. Del mismo modo, los **datos faltantes** están concentrados en un único moderador y no afectan las conclusiones principales.

En conjunto, este panorama confirma que la relación **asistencia–rendimiento es consistente, robusta y relativamente estable** a través de disciplinas, contextos y décadas.

Conclusión general del meta-análisis

El meta-análisis realizado muestra que:

- La asistencia a clase se asocia de manera **positiva y moderada** con el rendimiento académico.
- La estimación combinada en la escala de correlación es aproximadamente:

$$r \approx 0.40 - 0.45,$$

según el modelo utilizado (efectos aleatorios o multinivel).

- Este efecto es **estadísticamente significativo** en todos los modelos y se mantiene estable bajo múltiples especificaciones.

Los análisis por subgrupos indican que la relación es similar:

- en calificaciones de curso (*grade*, $r \approx 0.43$),
- en promedio general (*gpa*, $r \approx 0.36$),
- en asignaturas de ciencias ($r \approx 0.47 - 0.56$),
- y en áreas no científicas ($r \approx 0.42 - 0.46$).

No existe evidencia de que la relación haya cambiado a lo largo del tiempo (1927–2009), indicando una **estabilidad histórica notable**.

El **modelo multinivel**, aunque metodológicamente más sofisticado, produce conclusiones equivalentes a las del modelo principal, reforzando su **solidez**.

En síntesis:

Los estudiantes que **asisten más a clase** tienden a **rendir mejor académicamente**, y esta relación se mantiene en una amplia variedad de contextos, disciplinas y períodos históricos.

Esta consistencia, combinada con la **robustez frente a sesgo de publicación** y a **datos faltantes**, sugiere que la asociación asistencia–rendimiento es un **fenómeno educativo real, persistente y empíricamente bien respaldado**.

Limitaciones y recomendaciones futuras

Aunque las conclusiones del meta-análisis son sólidas, el estudio presenta algunas **limitaciones** inherentes a la evidencia disponible:

1. Heterogeneidad extremadamente alta

La variabilidad entre estudios no puede explicarse por los moderadores examinados. Esto sugiere la influencia de factores no reportados, tales como:

- características institucionales,
- dinámicas de evaluación,
- perfiles sociodemográficos,
- modalidades de cursado,
- calidad y estilo docente,
- políticas de asistencia.

Futuros estudios deberían reportar estos elementos de forma más sistemática.

2. Variables de moderación incompletas

Aunque yi y vi están completos para todos los estudios, alrededor del **31%** de las observaciones carece de información sobre el **tipo de asignatura**, lo cual reduce el poder para evaluar este moderador. Sería deseable que las publicaciones futuras incluyan descripciones más detalladas y estandarizadas de las asignaturas y contextos educativos.

3. Limitación conceptual de los estudios observacionales

Casi todos los estudios incluidos son **correlacionales**. Por lo tanto:

- no permiten establecer **causalidad**,
- no controlan rigurosamente **factores de confusión**,
- y no pueden aislar el efecto de la asistencia *per se* frente a variables como motivación, responsabilidad o características individuales.

En investigaciones futuras, serían valiosos **diseños experimentales o cuasi-experimentales**, así como **análisis longitudinales** con control de covariables.

4. Escasez de información moderna

La mayoría de los estudios preceden la era digital y no reflejan:

- cursos híbridos,
- plataformas virtuales,
- videos asincrónicos,
- nuevas formas de evaluación,
- ni cambios en los hábitos de estudio.

Bibliografía de las librerías utilizadas

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2025).
meta: General Package for Meta-Analysis.
R package version 7.0-0.
Disponible en: <https://CRAN.R-project.org/package=meta>

- Viechtbauer, W. (2025).
metafor: Meta-Analysis Package for R.
R package version 4.6-0.
Disponible en: <https://CRAN.R-project.org/package=metafor>
- Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2025).
dmetar: Companion R Package for the Guide “Doing Meta-Analysis in R”.
R package version 0.0.9000.
Disponible en: <https://CRAN.R-project.org/package=dmetar>
(Nota: algunas versiones se distribuyen vía GitHub).
- Wickham, H., Lüdecke, D., Patil, I., Ben-Shachar, M., & Makowski, D. (2025).
metadat: Meta-Analysis Datasets.
R package version 1.2-0.
Disponible en: <https://CRAN.R-project.org/package=metadat>
- Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2025).
tidyverse: Easily Install and Load the Tidyverse.
R package version 2.0.0.
Disponible en: <https://CRAN.R-project.org/package=tidyverse>
- Credé, M., Roch, S. G., & Kiesczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, 80(2), 272–295. <https://doi.org/10.3102/0034654310362998>
- Viechtbauer, W. (2023). metadat: Meta-Analysis Datasets (Version X.X-X) [R package]. <https://wviechtb.github.io/metadat/>
- Viechtbauer, W. (2023). Dataset: dat.crede2010 – Studies on the relationship between class attendance and grades. In metadat: Meta-Analysis Datasets. Retrieved from <https://wviechtb.github.io/metadat/reference/dat.crede2010.html>