

Evaluando información contradictoria: efectividad de un programa para el desarrollo del pensamiento crítico en adolescentes

Assessing Contradictory Information: Effectiveness of a Program for the Development of Critical Thinking in Adolescents

Rocío Pérez García-Madrid*¹, Núria Castells Gómez* y Marta Minguela Brunat*
Departamento de Cognición, Desarrollo y Psicología de la Educación. Universidad de Barcelona (España)

Resumen

Las tecnologías de la información y la comunicación ponen a nuestro alcance múltiples fuentes e informaciones diversas sobre un mismo tema, lo que dificulta la tarea de valorarlas y extraer conclusiones al respecto. En este artículo se presentan los resultados de la aplicación de un programa de intervención dirigido a promover el pensamiento crítico, centrado en enseñar al alumnado a elaborar preguntas críticas, capacitándoles para seleccionar y valorar correctamente las evidencias presentes en los textos. El estudio se realizó con cuatro grupos de 3º de ESO (2 grupos control y 2 grupos de intervención) en el área de ciencias naturales. Los resultados muestran una mayor capacidad para formular preguntas críticas sobre los textos y una mayor habilidad para seleccionar y valorar correctamente las evidencias de los textos por parte del alumnado de los grupos de intervención frente al de los grupos control tras la aplicación del programa. Estos resultados avalan la posibilidad de enseñar a los y las estudiantes a valorar la credibilidad de las informaciones que leen, lo que resulta prioritario en el contexto actual para ayudarles a desarrollar el pensamiento crítico ante las múltiples informaciones contradictorias a su alcance.

Palabras clave: pensamiento crítico; interrogación; evidencia; comprensión; lectura.

Correspondencia: Rocío Pérez García-Madrid. rperezga15@doct.ub.edu, Paseo de la Vall d'Hebron, 171; Facultad de Psicología, Sección Psicología Evolutiva y de la Educación; 08035 Barcelona.

Abstract

Information and communication technologies provide us with multiple sources and diverse information on the same subject, which makes it difficult to assess this information and draw conclusions from it. This article presents the results of an intervention program aimed at promoting critical thinking, focusing on teaching students to elaborate critical questions and enabling them to correctly select and evaluate the evidence presented in the texts. The study was carried out with four groups of secondary education students (2 control groups and 2 intervention groups) in a natural sciences course. The results show a greater ability to formulate critical questions, and a greater ability to correctly select and evaluate the evidence of the texts by participants in the intervention groups, compared to the students of the control groups. These results support the possibility of teaching students to assess the reliability of the information they read, which is a priority nowadays to help them develop critical thinking to face the multiple contradictory pieces of information at their disposal.

Keywords: critical thinking; questioning; evidence; comprehension; reading.

Introducción

Actualmente estamos expuestos a múltiples informaciones contradictorias, engañosas o falsas sobre temas controvertidos. Ello puede entorpecer la construcción del conocimiento, especialmente en la población adolescente. De hecho, los resultados de investigaciones recientes evidencian que esta población muestra dificultades tanto en la gestión de la desinformación (Herrero-Curiel y La-Rosa, 2022; Herrero-Diz et al., 2021) como en la capacidad para argumentar críticamente ante la misma (Castells et al., 2022), lo que compromete su capacidad para valorar e interpretar críticamente informaciones conflictivas. Considerando que la formación que reciben en torno a estos aspectos es escasa (Pérez et al., 2018), resulta fundamental diseñar intervenciones para enseñar al alumnado a enfrentarse de forma crítica a los textos, fomentando su conciencia crítica o vigilancia epistémica ante las informaciones que leen a diario en la red (Gierth y Bromme, 2020a; Tseng et al., 2021).

Preguntas críticas y análisis de evidencias de textos socio-científicos con informaciones contradictorias

Como señalan List y Alexander (2019), leer múltiples fuentes sobre un mismo tema permite, potencialmente, alcanzar una mejor comprensión y aprendizaje que la lectura de un solo documento. Sin embargo, la lectura de fuentes con información contradictoria requiere de habilidades que el alumnado adolescente no siempre ha desarrollado (List y Alexander, 2019) y que rara vez se enseñan (Solé, 2019). De forma sucinta, comprender múltiples textos implica dos procesos relacionados: por una parte, comprender la relación entre los textos leídos (de complementariedad o de discrepancia) y sus características (fuente, autoría, etc.), lo que Rouet y Britt (2014) denominan construir un modelo intertextual (“Intertext Model”); por otra parte, implica construir una representación integrada del contenido de los textos (“Integrated Model”). Si bien

la integración es un componente esencial para lograr una comprensión profunda de múltiples textos (contrastar y cuestionar críticamente las informaciones, responder preguntas, escribir un texto, etc.) (Barzilai et al., 2018; Miras, 2019), atender a las características intertextuales permite evaluar las fuentes, y determinar la potencial veracidad de la información que transmiten. Así, ambos procesos son necesarios para una lectura crítica de textos con información contradictoria. No obstante, dadas las dificultades que el alumnado presenta para identificar tanto la información contrapuesta como para valorarla adecuadamente, resulta imprescindible plantearse cómo fomentar su pensamiento crítico.

El pensamiento crítico, según Ennis (2018), consiste en buscar y proporcionar argumentos alternativos; respaldar una posición concreta cuando sea justificado hacerlo; estar bien informado; considerar distintos puntos de vista y ser consciente de las propias creencias. Lograrlo precisa recurrir a la lectura y/o revisión de textos diversos para discernir entre información cierta y falsa, entre las evidencias y las opiniones. Justamente, Tseng et al. (2021) plantean que la conciencia crítica o vigilancia epistémica ante la información requiere atender tanto a la evaluación de las características de las fuentes (considerando la autoría y la fiabilidad de la fuente), como a la evaluación del contenido del texto.

Cuando no se dispone de suficiente información para evaluar las fuentes, solo se puede valorar la validez de las aseveraciones que contienen (Gierth y Bromme, 2020b). Ahora bien, evaluar la plausibilidad de la información resulta complicado cuando tampoco se dispone de conocimiento previo suficiente sobre una temática (Lombardi et al., 2018), como les suele ocurrir a los y las adolescentes ante ciertos temas. En estos casos, la única forma de discernir entre informaciones contrapuestas recae en la posibilidad de identificar y cuestionar los argumentos, lo que requiere de una enseñanza explícita, como recientemente han puesto de manifiesto distintos estudios (Hruschka y Appel, 2023; Kiili et al., 2022; Tseng et al., 2021).

Son todavía escasos los estudios que comprueban la eficacia de programas educativos con relación a este aspecto. Entre ellos, recientemente Tseng et al. (2021) evaluaron el impacto del uso de una guía con preguntas orientadas a valorar críticamente informaciones de la red en alumnado de secundaria. A pesar de que se identificó un impacto positivo en la capacidad de los y las participantes para analizar informaciones diversas, las mejoras fueron limitadas, hecho que las autoras atribuyen a la falta de instrucción explícita sobre el contenido de la guía. En otra intervención, también basada en el uso de una guía, se enseñó a parejas de estudiantes universitarios a resolver preguntas críticas para evaluar distintos aspectos de los argumentos de los textos que debían leer (en concreto, claridad, exactitud, lógica y multiperspectivismo) (Mayweg-Paus et al., 2016). Esta intervención capacitó al alumnado para valorar críticamente los argumentos de los textos. En consonancia con estos estudios, en el ámbito de la enseñanza de la comprensión lectora el uso de la interrogación se considera fundamental para la lectura crítica de textos, tanto para evaluar las intenciones de los autores, como para evaluar las evidencias o identificar sesgos (Cottrell, 2011). Asimismo, se ha demostrado la efectividad de enseñar al alumnado a generar preguntas (Stevens et al., 2020). Las investigaciones sobre la propuesta para mejorar la comprensión lectora denominada Questioning the Author (Beck et al., 2020), por ejemplo, muestran que los y las apren-

dices pueden aprender a formularse preguntas que les ayuden a comprender y a evaluar textos narrativos y expositivos. Sin embargo, no tenemos conocimiento de que se hayan realizado estudios orientados a fomentar la formulación de preguntas críticas para evaluar los argumentos que se presentan en textos de temáticas socio-científicas.

Además de identificar argumentos, es imprescindible poder valorar críticamente la calidad de las evidencias que los acompañan. La evidencia es la información que se presenta para sustentar una afirmación (Du y List, 2020); puede consistir en una opinión personal o en una aportación de un experto o experta, y puede proceder de una institución u organismo independiente o bien de una entidad con fines persuasivos (intereses políticos o financieros) (Duncan et al., 2022). Existen escasos estudios en los que se evalúe la calidad de la evidencia (Cartiff et al., 2021), y, aquellos que se han centrado en indagar cómo el alumnado evalúa este aspecto han mostrado que estos presentan serias dificultades para ello (Kiili et al., 2022; Tseng et al., 2021).

Objetivos

Partiendo de las limitaciones identificadas en los estudios precedentes, la finalidad de este estudio es evaluar la eficacia de un programa educativo para fomentar el pensamiento crítico, centrado en enseñar al alumnado de secundaria a elaborar preguntas críticas y, a través de ellas, reconocer y valorar la calidad tanto de los argumentos de los textos, como de las evidencias que acompañan a esos argumentos. En este marco, los objetivos específicos de este estudio se centran en:

- a) evaluar el impacto del programa sobre la capacidad de los y las estudiantes para elaborar preguntas críticas, en términos de cantidad y variedad.
- b) evaluar el impacto del programa sobre la habilidad para reconocer y valorar la calidad de las evidencias de los textos.

En relación con estos objetivos, se plantean las siguientes preguntas de investigación (PI):

PI1: ¿El alumnado del grupo de intervención (GI) elaborará más preguntas críticas para valorar la calidad de la información contradictoria proporcionada por textos científicos, y más variadas, que el alumnado del grupo control (GC) tras la intervención?

PI2: ¿El alumnado del GI, tras la intervención, será más capaz de identificar evidencias y de valorarlas adecuadamente que sus compañeros/as del GC?

Método

Participantes

Participaron 90 estudiantes de 3º de la ESO ($M = 14.2$ años, $DT = .402$) de un instituto concertado de Barcelona de 4 líneas (Grupo 1 $n = 22$; Grupo 2 $n = 19$; Grupo 3 $n = 25$; Grupo 4 $n = 24$), seleccionados por conveniencia. El 53.3% de la muestra se identificó como mujer y el 46.9% lo hizo como hombre. Los dos primeros grupos de este curso

conformaron el grupo control (GC, $n = 41$) y los otros dos, el grupo intervención (GI, $n = 49$). La distribución de los grupos en GI y GC se hizo aleatoriamente. El alumnado era mayoritariamente autóctono (solo un 5.5% procedía de fuera del país), con un nivel socioeconómico elevado, y no presentaba necesidades específicas de apoyo educativo. Los y las participantes dieron su consentimiento informado para participar en este estudio. Con el fin de respetar su privacidad, la primera autora les otorgó una codificación alfanumérica en el momento de crear la base de datos, evitando así que pudieran ser identificados.

Instrumentos vinculados a la recogida de datos

Test de conocimientos previos

Se elaboró una prueba de conocimientos previos sobre las temáticas del área de ciencias naturales sobre las que versarían los textos que se emplearían en este estudio, constituida por 13 preguntas de opción múltiple, con cuatro opciones de respuesta. La fiabilidad de esta prueba fue $\alpha = .65$, lo que según Hinton et al. (2004) representa una fiabilidad moderada. Otros metodólogos (p.e., Kerlinger y Lee, 2000) indican que este nivel de fiabilidad resulta aceptable en función del uso que se haga de la medida y del tipo de decisiones que se tomen en base a esta. En nuestro estudio esta medida se utiliza como control.

Test de comprensión lectora

El nivel de comprensión lectora del alumnado se evaluó mediante el Test de Estrategias de Comprensión (TEC) (Vidal-Abarca et al., 2007). Esta prueba estandarizada consta de dos textos, cada uno de ellos seguido de diez preguntas de opción múltiple que evalúan cuatro procesos de comprensión lectora: formación de ideas, conexión de esas ideas, activación de conocimientos previos y formación de macroideas. Los autores de esta prueba reportan una fiabilidad de $\alpha = .80$.

Pares de textos para las pruebas pre-test, post-test, y mantenimiento

Se elaboraron tres pares de textos con argumentos contradictorios en torno a temáticas controvertidas del área de ciencias naturales (un par para cada una de las fases de pre-test, post-test y mantenimiento; ver ejemplo en Apéndice A). Las temáticas se seleccionaron en función de los contenidos de las secuencias didácticas que se estaban trabajando en el aula. Estos seis textos compartían las siguientes características: número de palabras (aproximadamente 350), formato (noticia de 4-5 párrafos), número de argumentos y contraargumentos (ver Tabla 1). Considerando los argumentos proporcionados, uno de los textos del par siempre era más riguroso que el otro (p.e., fuentes más fiables que otras; información procedente de profesionales cualificados y/o expertos/as sin intereses personales o económicos vs. opiniones de personas con potenciales intereses; evidencias científicas procedentes de entidades independientes y rigurosas vs. datos de origen desconocido).

Tabla 1

Características de los textos utilizados en el pre-test, post-test y mantenimiento

| Fases | Textos (T): título, autoría, fuente y fecha de publicación | Nº de palabras | Nº de argumentos | | | Nº de contra-argumentos | |
|---------------|--|----------------|------------------|-----------|------------|-------------------------|-----------|
| | | | Sin evid. | Con evid. | Rel. Ideas | Sin evid. | Con evid. |
| Pre-test | T1: <i>Dieta Keto: ¿comer más grasa para perder peso funciona?</i> Clara Sánchez - TELVA 03/05/2018 | 336 | 5 | 4 | 2 | 1 | 0 |
| | T2*: <i>Los expertos dicen que la dieta keto no es sostenible. Entonces, ¿por qué es tan popular?</i> Santee LaMotte-CNN 06/01/2020 | 362 | 5 | 5 | 2 | 1 | 0 |
| Post-test | T3*: <i>La contaminación del aire que respiramos provoca millones de muertes</i> Susan Golberg - National Geographic 02/04/2020 | 358 | 5 | 5 | 2 | 1 | 0 |
| | T4: <i>Un asesor científico de los EE. UU. defiende que el aire contaminado ayuda a los niños</i> Sean Reilly- EyE News 06/11/2019 | 350 | 4 | 3 | 2 | 0 | 1 |
| Mantenimiento | T5: <i>La dieta radical que puede curar la diabetes tipo 2, según un estudio revolucionario</i> Miguel Sola- El confidencial 05/12/2017 | 342 | 4 | 4 | 2 | 1 | 0 |
| | T6*: <i>Diabetes tipo 2 y alimentación</i> Mercedes Galindo Rubio- Fundación Española de Diabetes 04/04/2019 | 346 | 4 | 4 | 2 | 0 | 1 |

Nota: Sin evid.=sin evidencias, Con evid.=con evidencias, Rel. ideas=relacionan ideas

*El asterisco marca los textos de cada par que contenían información más rigurosa

Tarea de preguntas críticas pre-test, post-test, y mantenimiento

Cada uno de los pares de textos elaborados iba acompañado de una tarea consistente en: 1) elaborar preguntas críticas para valorar los argumentos de los textos; 2) identificar las ideas principales (argumentos) de cada texto; 3) identificar las evidencias aportadas en los textos para cada idea y valorar en cada caso quién las proporcionaba (si se trataba de expertos/as o no), si procedían de estudios (o no), y si estos estudios eran llevados a cabo por organismos reconocidos (o no). La tarea de identificación de argumentos y la valoración de evidencias se organizó en una tabla (ver Apéndice B).

Instrumentos vinculados al programa de intervención

Textos

Para las sesiones de intervención, se elaboraron cinco textos argumentativos con características similares a los textos de las fases pre-test, post-test y mantenimiento (ver Tabla 2).

Tabla 2

Características de los textos utilizados durante la intervención

| Sesiones (S) | Textos (T): título, autoría, fuente y fecha de publicación | Nº de palabras | Nº de argumentos | | | Nº de contra-argumentos | |
|--------------|--|----------------|------------------|-----------|--------------|-------------------------|-----------|
| | | | Sin evid. | Con evid. | Relac. ideas | Sin evid. | Con evid. |
| S2 | T7: <i>España, tercer país de la Unión Europea con más brotes de salmonela el 2018</i> Heraldo de Aragón 12/12/2019 | 436 | 5 | 4 | 2 | 0 | 1 |
| S3 | T8: <i>El fin de los antibióticos</i> Jordi Regàs y Sara Segarra - TV3 07/01/2018 | 346 | 4 | 4 | 2 | 1 | 0 |
| | T9: <i>El uso responsable de los antibióticos tiene beneficios para todos</i> Gencat 15/11/2017 | 352 | 4 | 3 | 2 | 0 | 1 |

| | | | | | | | |
|----|---|-----|---|---|---|---|---|
| S4 | T10*: <i>Aunque se aplique la ley seca del tabaco, fumar no es tan malo</i> Autor desconocido 02/01/2007 | 439 | 3 | 4 | 2 | 1 | 0 |
| | T11: <i>¿Es malo fumar un cigarro al día?</i> Ángeles Gómez López 28/03/2018 | 396 | 4 | 5 | 2 | 1 | 1 |

Nota: La primera sesión del programa de intervención consistió en la revisión de la tarea realizada en el pre-test, por lo que se trabajó sobre los textos empleados en esa fase (T1 y T2, en Tabla 1).

*El asterisco marca el texto que contenía información escasamente rigurosa.

Vídeo modelo

Tomando como referencia el uso de vídeos modelo en otras investigaciones (Mateos et al., 2018), se grabó un vídeo breve en el que dos estudiantes de la edad del alumnado participante ejemplificaban cómo usar preguntas críticas durante la lectura de un texto argumentativo para analizar la rigurosidad de su contenido (concretamente, el T7, ver Tabla 2). Puede consultarse un extracto del guión del vídeo en el Apéndice C.

Guía de preguntas críticas

Guía con ejemplos de preguntas críticas (ver Apéndice D) para valorar distintos aspectos de los argumentos de los textos (claridad, veracidad, lógica, exactitud, multiperspectivismo, identificación de expertos citados en los textos y características de las fuentes) (Mayweg-Paus et al., 2016).

Documento explicativo para la identificación y valoración de la calidad de las evidencias

Guía con una recopilación de ejemplos de las evidencias de los textos leídos en el pre-test y una explicación de su análisis, para ayudar al alumnado a valorar la calidad de las evidencias. Partiendo de las aportaciones de otros autores (Duncan et al., 2022; Gough, 2021), las dimensiones empleadas para el análisis fueron: (1) procedencia de la evidencia (opinión personal vs. proporcionada por expertos/as); y (2) procedencia de los datos asociados a la evidencia (p.e., de estudios realizados por organismos o instituciones reconocidas e independientes vs. de empresas con intereses económicos) (ver Apéndice E).

Procedimiento

Fase pre-test

El pre-test se realizó en dos sesiones de 60 minutos cada una. En la primera, el alumnado resolvió la prueba de conocimientos previos y el TEC (Vidal-Abarca et al., 2007). En la segunda, se les pidió que leyeran individualmente dos textos argumentativos (T1 y T2) y que completaran la tarea de pre-test mediante el programa Qualtrics^{XM}. Se siguió el mismo procedimiento en todos los grupos.

Fase de intervención

La intervención se implementó en la asignatura de ciencias naturales, materia en la que es habitual la lectura de múltiples fuentes argumentativas con informaciones contradictorias sobre un mismo tema. El diseño del programa partió de una perspectiva socio-constructivista sobre la enseñanza y el aprendizaje. Esto se refleja en aspectos como el uso de los conocimientos previos del alumnado como base para su aprendizaje, la diversificación de las ayudas y apoyos durante las diferentes sesiones, y el traspaso progresivo del control y la responsabilidad al alumnado (Fisher y Frey, 2013). Además, las sesiones se basaron en la argumentación colectiva, induciendo a los y las participantes a leer y releer los diferentes textos y a cuestionarse sus respuestas (Wilkinson et al., 2017).

El programa de intervención consistió en cuatro sesiones de 40-60 minutos. La primera autora implementó la intervención que se resume en la Tabla 3 con los GI, mientras el profesorado de los GC trabajó los mismos textos en sus aulas. Se realizaron entrevistas semiestructuradas con las docentes de los GC para asegurar el uso de estos textos y conocer el trabajo que se realizó con ellos, asegurando así que este remitía básicamente a la comprensión del texto y no al análisis crítico de los argumentos.

Tabla 3

Resumen de las sesiones, objetivos y materiales del programa de intervención

| Sesión (S) y duración | Actividades de la sesión | Materiales utilizados | Deberes |
|------------------------------|---|--|----------------|
| S1 (60 mins.) | -Repaso conjunto del pre-test para ayudar a comprender las preguntas críticas del pre-test, y revisar y valorar los argumentos, contraargumentos y evidencias de los textos. -Instrucción explícita para la elaboración de preguntas críticas. | -T1 y T2 -Respuestas del alumnado al pre-test | -Leer T7 |

| | | | |
|------------------|---|--|--|
| S2 (45 mins.) | <ul style="list-style-type: none"> - Visionado del vídeo modelo del uso de preguntas críticas y utilización de la <i>Guía de preguntas críticas</i>. - Enseñanza de las diferencias entre los tipos de evidencias a partir del <i>Documento explicativo para la selección y valoración de la calidad de las evidencias</i>. | <ul style="list-style-type: none"> -T7 - Vídeo modelo - Guía de preguntas críticas - Documento explicativo para la selección y valoración de la calidad de evidencias | <ul style="list-style-type: none"> -Leer T8 y T9 - Elaborar preguntas críticas y valorar las evidencias de T8 y T9 (usando la guía y el documento explicativo) |
| S3 (40 mins.) | <ul style="list-style-type: none"> -Corrección conjunta de los deberes para evaluar los aspectos aprendidos y aquellos en los que era necesario incidir más. -Responder a preguntas críticas de comparación intertextual (individual). | <ul style="list-style-type: none"> -Recopilación de las respuestas del alumnado a la tarea de deberes | <ul style="list-style-type: none"> -Leer T10 y T11 |
| S4 (60 mins.) | <ul style="list-style-type: none"> - Práctica de la elaboración de preguntas críticas y evaluación de evidencias (sin materiales de apoyo) de los T10 y T11. | <ul style="list-style-type: none"> - T10 y T11 - Actividad para formular preguntas críticas e identificar y evaluar evidencias (sin la guía ni el documento explicativo, pero con ayuda de la investigadora) | |

Fases de post-test y mantenimiento

Las tareas del post-test y mantenimiento se realizaron en una sesión de una hora cada una. En ellas, se pidió al alumnado que leyera un par de textos (T3 y T4 para el post; T5 y T6 para el mantenimiento) y que resolviera las tareas correspondientes individualmente (las mismas que en el pre-test utilizando el programa Qualtrics^{XM}). Los GC y GI resolvieron estas tareas en las mismas condiciones, realizándose el post-test en torno a dos semanas después de finalizar la intervención y el mantenimiento un mes después del post-test.

Análisis de datos

Para responder a la primera pregunta de investigación, relacionada con la comparación de la cantidad y variedad de preguntas críticas planteadas por los GI y GC, se procedió, en primer lugar, a identificar las preguntas críticas elaboradas por el alumnado en el pre-test, post-test y mantenimiento. Para ello, se consideraron preguntas críticas aquellas que conducen a valorar la información y a reflexionar críticamente sobre la misma.

En segundo lugar, se clasificaron las preguntas críticas identificadas según los aspectos que podrían contribuir a valorar (en concreto, claridad, exactitud, veracidad, lógica, multiperspectivismo, identificación de expertos/as, y análisis de la fuente y el canal de divulgación). Los criterios empleados en cada caso, partiendo de las aportaciones de autores como Duncan et al. (2022) y Mayweg-Paus et al. (2016), fueron los que se presentan en la Guía de preguntas críticas (ver Apéndice D).

Para calcular el acuerdo interjueces para esta clasificación, 3 investigadoras codificaron las preguntas planteadas por 20 estudiantes de ambos grupos, y para cada una de las fases de recogida de datos, de forma independiente, llegando a un acuerdo del 90.7%. Las discrepancias se resolvieron mediante discusión y una de las investigadoras acabó de codificar las preguntas del resto de participantes.

Partiendo de esta codificación, se realizó un conteo del número de tipos de preguntas críticas distintas que elaboraron en el pre-test, en el post-test y en el mantenimiento para identificar la diversidad de preguntas críticas elaboradas por los y las estudiantes.

Para responder a la segunda pregunta de investigación, relacionada con la comparación de la capacidad para identificar y valorar evidencias por parte del alumnado de GC y GI, en primer lugar, se realizó un conteo del número de evidencias que cada estudiante había identificado correctamente en los textos empleados en el pre-test, post-test y mantenimiento, y se calculó el porcentaje de evidencias correctamente identificadas a partir del total de evidencias posibles en cada caso. Posteriormente, y en relación con la valoración de las evidencias, se puntuaron las respuestas que proporcionó el alumnado a las preguntas críticas sobre dichas evidencias (0= incorrecto, 1= correcto) y se hizo un conteo del total de respuestas correctas de cada estudiante para el pre-test, post-test y mantenimiento.

A partir de las codificaciones de las preguntas críticas y de las valoraciones de las evidencias realizadas por los y las estudiantes, se realizaron dos tipos de análisis estadísticos. Por un lado, para identificar el grado en el que el alumnado de cada grupo (CG y GI) había mejorado a lo largo de las distintas fases, se calculó el estadístico Friedman y, cuando este resultó significativo, la prueba de Wilcoxon. Por otro lado, para comparar la actuación del alumnado de los GI y GC para el pre-test, el post-test y el mantenimiento, se realizaron pruebas de U de Mann-Whitney. Se optó por pruebas no paramétricas debido a que los datos no cumplían el supuesto de normalidad.

Resultados

Conocimientos previos y comprensión lectora

Las medidas de conocimientos previos y comprensión lectora general (medida por medio del TEC) (Vidal-Abarca et al., 2007) se tomaron como variables de control. Antes de abordar los resultados vinculados a las preguntas de investigación, se proporcionan las comparaciones de medias entre los GC y GI para estas medidas. Los resultados de las pruebas de U de Mann-Whitney muestran que no había diferencias significativas entre el alumnado del GC y del GI ni en su nivel de conocimientos previos sobre las temáticas de los textos empleados en este estudio (GC: $M = 3.66$, $DT = 1.98$; GI: $M =$

3.55, $DT = 1.72$; $U = 983.5$, $Z = -.173$, $p = .863$), ni en su nivel de comprensión lectora general (GC: $M = 13.54$, $DT = 3.66$; GI: $M = 13.39$, $DT = 3.79$; $U = 994.5$, $Z = -.081$, $p = .935$).

Elaboración de preguntas críticas

En relación con la cantidad de preguntas críticas elaboradas, los resultados de las pruebas no paramétricas de Friedman y de Wilcoxon muestran que tanto el alumnado del GC ($\chi^2 = 16.147$; $p < .001$) como el del GI ($\chi^2 = 70.560$; $p < .001$) mejoró significativamente a lo largo de las distintas fases del estudio (Wilcoxon pre-test y post-test: GC: $z = -3.333$, $p = .001$; GI: $z = -5.983$, $p < .001$; pre-test y mantenimiento: GC: $z = -3.273$, $p = .001$; GI: $z = -5.889$, $p < .001$). Así, ambos grupos fueron capaces de elaborar más preguntas críticas en el post-test y el mantenimiento que en el pre-test.

Con todo, los resultados de la prueba U de Mann-Whitney en el pre-test, post-test y mantenimiento muestran que, aunque antes de la intervención no había diferencias significativas entre los y las participantes de los GC y GI, tras el programa de intervención el alumnado del GI elaboró un número significativamente mayor de preguntas críticas que el del GC, tanto en el post-test, como en el mantenimiento ($p < .001$ en ambos casos y con un tamaño del efecto grande, ver Tabla 4).

Tabla 4

Medias, desviaciones típicas, comparación de medias y tamaño del efecto entre GI y GC para la cantidad de preguntas críticas elaboradas

| Fases | GC M (DT) | GI M (DT) | Comparación de medias | Tamaño del efecto |
|---------------|--------------|--------------|--|----------------------------|
| Pre-test | .15 (.422) | .22 (1.08) | $U = 948$; $Z = -.928$; $p = .354$ | |
| Post-test | 1.12 (1.72) | 5.63 (3.48) | $U = 180.5$; $Z = -6.79$; $p < .001$ | $1-\beta = 1$; $d = 1.64$ |
| Mantenimiento | 1.02 (1.63) | 5.41 (3.45) | $U = 214$; $Z = -6.54$; $p < .001$ | $1-\beta = 1$; $d = 1.62$ |

En relación con la variedad de las preguntas críticas formuladas, medida en términos de la cantidad de tipos de preguntas críticas elaboradas, la prueba de Friedman (GC: $\chi^2 = 12.603$; $p = .002$; GI: $\chi^2 = 71.326$; $p < .001$) evidenció que tanto el alumnado del GC como el del GI evolucionó favorablemente a lo largo de las diferentes fases del estudio (Wilcoxon pre-test y post-test: GC: $z = -3.093$, $p = .002$; GI: $z = -6.001$, $p < .001$; pre-test y mantenimiento: GC: $z = -2.914$, $p = .004$; GI: $z = -5.934$, $p < .001$), planteando en ambos casos preguntas críticas cada vez más variadas.

Ahora bien, los análisis de U de Mann-Whitney muestran que, aunque los GC y GI no diferían en este aspecto en el pre-test ($p = .354$), el alumnado del GI realizó preguntas críticas significativamente más variadas después de la intervención que sus compañeros del GC ($p < .001$ tanto en el post-test como en el mantenimiento, y con un tamaño del efecto grande, ver Tabla 5).

Tabla 5

Medias, desviaciones típicas y resultados de la comparación de medias y tamaño del efecto entre GI y GC para la cantidad de tipos de preguntas críticas elaboradas

| Fases | GC M (DT) | GI M (DT) | Comparación de medias | Tamaño del efecto |
|---------------|--------------|--------------|----------------------------------|-------------------------|
| Pre-test | .15 (.422) | .16 (.717) | $U = 948; Z = -.928; p = .354$ | |
| Post-test | .78 (1.19) | 3.73 (1.70) | $U = 177; Z = -6.85; p < .001$ | $1-\beta = 1; d = 2.01$ |
| Mantenimiento | .68 (1.08) | 3.49 (1.89) | $U = 205.5; Z = -6.64; p < .001$ | $1-\beta = 1; d = 1.82$ |

Adicionalmente, se comparó el número de preguntas críticas de cada tipo elaboradas por el GC y el GI. Como muestra la Tabla 6, para todos los tipos de preguntas críticas el alumnado de ambos grupos elaboró un número de preguntas similar antes de la intervención; sin embargo, el alumnado del GI formuló un número significativamente superior de preguntas críticas de los distintos tipos tanto en el post-test, como en el mantenimiento. Además, entre el post y el mantenimiento, en el GI hubo un incremento en el número de ciertos tipos de preguntas críticas formuladas (preguntas sobre la claridad, la exactitud, y el multiperspectivismo).

Tabla 6

Medias, desviaciones típicas y resultados de la comparación de medias de los diferentes tipos de preguntas formuladas por los GI y GC

| Tipos de preguntas críticas | Fases | GC M (DT) | GI M (DT) | Comparación de medias |
|-----------------------------|---------------|--------------|--------------|----------------------------------|
| Claridad | Pre-test | .00(.000) | .08(.571) | $U = 984; Z = -.915; p = .360$ |
| | Post-test | .10(.490) | .65(.723) | $U = 530.5; Z = -4.73; p < .001$ |
| | Mantenimiento | .12(.458) | 1.02(.829) | $U = 356.5; Z = -5.89; p < .001$ |
| Exactitud | Pre-test | .00(.000) | .00(.000) | $U = 1004.5; Z = .000; p = 1$ |
| | Post-test | .05(.218) | .37(.727) | $U = 802.5; Z = -2.59; p = .009$ |
| | Mantenimiento | .05(.218) | .61(.996) | $U = 696.5; Z = -3.50; p < .001$ |
| Veracidad | Pre-test | .05(.218) | .02(.143) | $U = 976; Z = -.743; p = .458$ |
| | Post-test | .15(.358) | .92(.812) | $U = 445; Z = -5.12; p < .001$ |
| | Mantenimiento | .10(.300) | .41(.674) | $U = 785; Z = -2.50; p = .012$ |

| | | | | |
|---|---------------|------------|-------------|----------------------------------|
| Lógica | Pre-test | .00(.000) | .00(.000) | $U = 1004.5; Z = .000; p = 1$ |
| | Post-test | .05(.218) | .53(.739) | $U = 638.5; Z = -3.95; p < .001$ |
| | Mantenimiento | .02(.156) | .59(.814) | $U = 595.5; Z = -4.41; p < .001$ |
| Multipers- pectivismo | Pre-test | .02(.156) | .00(.000) | $U = 980; Z = -1.09; p = .274$ |
| | Post-test | .02(.156) | .45(.765) | $U = 699; Z = -3.63; p < .001$ |
| | Mantenimiento | .00(.000) | .67(.944) | $U = 594.5; Z = -4.57; p < .001$ |
| Identificación de expertos | Pre-test | .05(.218) | .06(.242) | $U = 992; Z = -.255; p = .799$ |
| | Post-test | .15(.358) | 1.08(.812) | $U = 336.5; Z = -5.97; p < .001$ |
| | Mantenimiento | .20(.511) | .84(.943) | $U = 575.5; Z = -4.03; p < .001$ |
| Análisis de la fuente y del canal de divulgación | Pre-test | .02(.156) | .10(.368) | $U = 946.5; Z = -1.18; p = .236$ |
| | Post-test | .61(1.093) | 1.63(1.439) | $U = 519; Z = -4.15; p < .001$ |
| | Mantenimiento | .54(1.002) | 1.27(1.238) | $U = 607.5; Z = -3.46; p < .001$ |

Identificación y valoración de evidencias

Los resultados de la prueba de Friedman indican mejoras significativas en ambos grupos en la identificación de evidencias (GC: $\chi^2 = 6.685, p = .035$; GI: $\chi^2 = 70.578, p < .001$). En concreto, la prueba de Wilcoxon puso de manifiesto diferencias significativas para el GI (pre-test y post-test: $z = -6.095, p < .001$; pre y mantenimiento: $z = -6.118, p < .001$), mientras que en el GC, la mejora significativa se manifiesta entre el pre-test y el post-test ($z = -2.678, p = .007$), pero desaparece en el mantenimiento ($z = -1.209, p = .227$). Ambos grupos mejoran también significativamente en la valoración de las evidencias (Friedman: GI $\chi^2 = 74.351, p < .001$; GC: $\chi^2 = 10.164, p = .006$). Para esta variable las diferencias son significativas tanto entre los resultados del pre-test y el post-test, como entre los del pre-test y el mantenimiento en ambos grupos (Wilcoxon: pre-test y post-test: GC: $z = -2.987, p = .003$; GI: $z = -6.095, p < .001$; pre-test y mantenimiento: GC: $z = -2.276, p = .023$; GI: $z = -6.100, p < .001$).

Al comparar los GC y GI en las distintas fases, las pruebas U de Mann Whitney muestran que estos grupos no diferían significativamente en el pre-test ni en el porcentaje de evidencias correctamente identificadas ($p = .069$, ver Tabla 7) ni en la valoración correcta de dichas evidencias ($p = .175$, ver Tabla 8), medida a partir del número total de respuestas críticas respondidas correctamente. Sin embargo, los resultados muestran que los y las participantes del GI fueron significativamente mejores en ambos aspectos en el post-test ($p < .001$ en los dos casos y con un tamaño del efecto grande, ver Tablas 7 y 8). Esta diferencia significativa se mantuvo en el mantenimiento ($p < .001$ en los dos casos y con un tamaño del efecto grande, ver Tablas 7 y 8).

Tabla 7

Medias, desviaciones típicas y resultados de la comparación de medias y tamaño del efecto entre GI y GC para el porcentaje de evidencias correctamente identificadas

| Fases | GC M (DT) | GI M (DT) | Comparación de medias | Tamaño del efecto |
|---------------|--------------|--------------|--------------------------------|-------------------------|
| Pre-test | 1.29 (1.82) | .71 (1.42) | $U = 811; Z = -1.81; p = .069$ | |
| Post-test | 2.51 (2.50) | 7.76 (2.79) | $U = 189; Z = -6.69; p < .001$ | $1-\beta = 1; d = 1.98$ |
| Mantenimiento | 1.59 (2.02) | 6.37 (2.24) | $U = 150; Z = -7.08; p < .001$ | $1-\beta = 1; d = 2.24$ |

Tabla 8

Medias, desviaciones típicas y resultados de la comparación de medias y tamaño del efecto entre los GI y GC para la valoración correcta de las evidencias a partir del número total de respuestas correctas

| Fases | GC M (DT) | GI M (DT) | Comparación de medias | Tamaño del efecto |
|---------------|--------------|--------------|----------------------------------|-------------------------|
| Pre-test | 2.73 (4.26) | 1.82 (3.80) | $U = 860; Z = -1.36; p = .175$ | |
| Post-test | 5.61 (5.78) | 21.90 (9.40) | $U = 173.5; Z = -6.75; p < .001$ | $1-\beta = 1; d = 2.08$ |
| Mantenimiento | 4.98 (6.60) | 23.02 (8.70) | $U = 127.5; Z = -7.17; p < .001$ | $1-\beta = 1; d = 2.33$ |

Discusión

El objetivo de esta investigación era valorar la eficacia de un programa educativo orientado a mejorar el pensamiento crítico del alumnado de 3º de ESO, enseñándoles a elaborar preguntas críticas y a reconocer y valorar las evidencias proporcionadas en los textos a partir de esas preguntas.

A pesar de que los resultados de la evolución de cada uno de los grupos a lo largo del tiempo muestran que tanto el alumnado del GC como el del GI mejora significativamente en todas las variables estudiadas, los y las participantes del GI obtuvieron resultados significativamente mejores que los y las del GC en todos los casos.

Así, el alumnado del GI formuló un número significativamente mayor de preguntas críticas, y más diversas, que el del GC. La diversidad de preguntas planteadas por el alumnado del GI permitiría un análisis más profundo de los argumentos de los textos. Además, para algunos tipos de preguntas críticas la cantidad de preguntas elaboradas aumenta entre el post-test y el mantenimiento, indicando que los y las participantes pueden seguir generando preguntas pertinentes para el análisis de los textos tras la intervención. Ello demuestra que el alumnado puede incorporar y generar cuestiones que le faciliten el análisis de la información, apropiándose de estos recursos, como han mostrado también estudios relacionados con la comprensión lectora (Beck et al., 2020).

El alumnado del GI también pudo identificar correctamente más evidencias y valorarlas de forma más adecuada que el del GC. Además, tanto para las preguntas críticas como para las evidencias, los resultados significativamente superiores del GI

persistieron en el mantenimiento, lo que subraya el potencial del programa educativo para promover aprendizajes duraderos y otorga validez a la intervención realizada (Graham y Hebert, 2011).

Los aprendizajes realizados por el alumnado del GI en el transcurso del programa educativo pueden atribuirse a distintos factores. En primer lugar, a diferencia de la propuesta de Tseng et al. (2021), en la que se utilizó una guía con preguntas críticas, pero no se enseñó al alumnado a utilizarla, en nuestro estudio se optó por la instrucción explícita de los contenidos objeto de la intervención y por el uso de guías basado en el traspaso progresivo del control al alumnado (Fisher y Frey, 2013). Además, posiblemente los recursos utilizados para enseñar a elaborar preguntas críticas e identificar y valorar evidencias, y la forma de utilizarlos durante la intervención, pueden haber favorecido la eficacia de la misma, como apuntan los resultados de los cálculos del tamaño del efecto. Por una parte, el vídeo modelo interpretado por estudiantes de la misma edad que el alumnado participante pudo facilitar la identificación con las actrices, empatizando con las dificultades, reflexiones y soluciones negociadas entre ellas. Por otra parte, la guía de preguntas críticas y el documento explicativo para la valoración de evidencias se utilizaron inicialmente en el marco de la interacción entre iguales; posteriormente, de forma individual; y finalmente, una vez los contenidos de esos recursos se habían asimilado e interiorizado, se prescindió de ellos. Así, a pesar de tratarse de un programa educativo relativamente corto e intenso, entendemos que el hecho de tratar de ceder progresivamente la responsabilidad al alumnado (Fisher y Frey, 2013), recurriendo al uso de herramientas externas para facilitar la revisión de la información mientras iban interiorizando el uso de dichas herramientas, contribuyó a dotarlos de los recursos necesarios para gestionar informaciones contrapuestas. En este sentido, conviene destacar la importancia de enseñar a contrastar textos diversos sobre un mismo tema. A pesar de las dificultades que entraña integrar informaciones de textos que suelen tener estructuras, contenidos e intenciones distintas entre sí (Mirás, 2019), es justamente la posibilidad de contrastar las evidencias de textos diversos lo que permite que los y las estudiantes perciban el interés de aprender a formular preguntas críticas y que puedan valorar adecuadamente la validez de los contenidos, cuando no se cuenta con un bagaje suficientemente amplio de conocimientos que permita otro tipo de análisis (Lombardi et al., 2018). Finalmente, a pesar de que las preguntas críticas se trabajaron en este caso en el área de ciencias naturales, otras investigaciones han utilizado este recurso en otras áreas y con textos argumentativos diversos, obteniendo resultados igualmente satisfactorios con población universitaria (Nussbaum y Edwards, 2011).

Por otro lado, sería interesante indagar los motivos del progreso que, a nivel intra-grupal, muestra el alumnado del GC. Al respecto, se pueden apuntar algunas hipótesis. Por un lado, sabemos que los 14-15 años suponen un punto de inflexión en la capacidad argumentativa del alumnado (Kuhn y Udell, 2007), lo que podría explicar la mejora en sus capacidades críticas. Por otra parte, el hecho de que trabajaran con textos con información contradictoria, a pesar de que el profesorado de los GC no contaba con pautas para trabajarlos de forma crítica ni indicaron haberlo hecho, pudo incentivar procesos de comparación y contraste que fomentasen un cierto desarrollo de dichas capacidades críticas. Contrastar estas hipótesis requeriría de un estudio microgenético de

los procesos de enseñanza-aprendizaje llevados a cabo en las aulas de los GC. A pesar de que nuestro estudio no permite este análisis, los resultados obtenidos evidencian que la enseñanza explícita de la formulación y uso de preguntas críticas favorece un mejor desarrollo de las capacidades críticas del alumnado.

Más allá de la importancia de los resultados obtenidos, que subrayan la potencialidad del programa, sería interesante poder determinar su impacto en contextos diversos (p.e. centros con distintas características). Además, el programa implementado ha atendido a determinadas dimensiones de las evidencias, pero no a otras muchas como la forma en la que se generan, validan, etc. (Duncan et al., 2022; Gough, 2021), que podrían ser el foco de futuras líneas de investigación.

Referencias

- Castells, N. García-Milà, M., Miralda, A., Luna, J., y Pérez, E. (2022), El razonamiento de los adolescentes para gestionar las noticias falsas. *Educación XXI: Revista de la Facultad de Educación*, 25(2), 291-313. <https://doi.org/10.5944/educxx1.31693>
- Barzilai, S., Zohar, A. R., y Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review*, 30(3), 973–999. <https://doi.org/10.1007/s10648-018-9436-8>
- Beck, I. L., McKeown, M. G., y Sandora, C. A. (2020). *Robust comprehension instruction with Questioning the Author*. Guilford Publications.
- Cartiff, B. M., Duke, R. F., y Greene, J. A. (2021). The effect of epistemic cognition interventions on academic achievement: A meta-analysis. *Journal of Educational Psychology*, 113(3), 477-498. <https://doi.org/10.1037/edu0000490>
- Cottrell, S. (2011). *Critical thinking skills: Developing effective analysis and argument* (2nd ed.). MacMillan.
- Du, H., y List, A. (2020). Researching and writing based on multiple texts. *Learning and Instruction*, 66, 101297. <https://doi.org/10.1016/j.learninstruc.2019.101297>
- Duncan, R. G., Cavera, V. L., y Chinn, C. A. (2022). The role of evidence evaluation in critical thinking: Fostering episCritical thinking in Biology and Environmental Education, *Contributions from biology education research*
- Ennis, R. (2018). Critical thinking across the curriculum. *Topoi*, 37(1), 165–184. <https://doi.org/10.1007/s11245-016-9401-4>
- Fisher, D., y Frey, N. (2013). *Better learning through structured teaching: A framework for the gradual release of responsibility* (2nd ed.). ASCD.
- Gierth, L., y Bromme, R. (2020a). Beware of vested interests: Epistemic vigilance improves reasoning about scientific evidence (for some people). *PLoS ONE*, 15(4), e0231387. <https://doi.org/10.1371/journal.pone.0231387>
- Gierth, L., y Bromme, R. (2020b). Attacking science on social media: How user comments affect perceived trustworthiness and credibility. *Public Understanding of Science*, 29(2), 230–247. <https://doi.org/10.1177/0963662519889275>

- Gough, D. (2021). Appraising evidence claims. *Review of Research in Education*, 45(1), 1-26. <https://doi.org/10.3102/0091732X20985072>
- Graham, S., y Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81(4), 710-744. <https://doi.org/10.17763/haer.81.4.t2k0m13756113566>
- Herrero-Curiel, E., y La-Rosa, L. (2022). Los estudiantes de secundaria y la alfabetización mediática en la era de la desinformación. *Comunicar: Revista Científica de Comunicación y Educación*, 30(73), 95-106. <https://doi.org/10.3916/C73-2022-08>
- Herrero-Diz, P., Conde-Jiménez, J., y Reyes-de-Cózar, S. (2021). Spanish adolescents and fake news: level of awareness and credibility of information. *Culture and Education*, 33(1), 1-27. <https://doi.org/10.1080/11356405.2020.1859739>
- Hinton, P. R., McMurray, I., y Brownlow, C. (2004). *SPSS Explained*. Routledge. <https://doi.org/10.4324/9780203642597>
- Hruschka, T. M. J., y Appel, M. (2023). Learning about informal fallacies and the detection of fake news: An experimental intervention. *PLoS ONE*, 18(3), e0283238. <https://doi.org/10.1371/journal.pone.0283238>
- Kerlinger, F. N., y Lee, H. B. (2000). *Foundations of behavioral research* (4th ed.). Harcourt College Publishers.
- Kiili, C., Lakkala, M., Ilomäki, L., Toom, A., Coiro, J., Hämäläinen, E., y Sormunen, E. (2022). Designing classroom practices for teaching online inquiry: Experiences from the field. *Journal of Adolescent & Adult Literacy*, 65(4), 297-308. <https://doi.org/10.1002/jaal.1206>
- Kuhn, D., y Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking & Reasoning*, 13(2), 90-104. <http://dx.doi.org/10.1080/13546780600625447>
- List, A., y Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist*, 54(1), 20-39. <https://doi.org/10.1080/00461520.2018.1505514>
- Lombardi, D., Bailey, J. M., Bickel, E. S., y Burrell, S. (2018). Scaffolding scientific thinking: Students' evaluations and judgments during earth science knowledge construction. *Contemporary Educational Psychology*, 54, 184-198. <https://doi.org/10.1016/j.cedpsych.2018.06.008>
- Mateos, M., Martín, E., Cuevas, I., Villalón, R., Martínez, I., y González-Lamas, J. (2018). Improving written argumentative synthesis by teaching the integration of conflicting information from multiple sources. *Cognition and Instruction*, 36(2), 119-138. <https://doi.org/10.1080/07370008.2018.1425300>
- Mayweg-Paus, E., Thiebach, M., y Jucks, R. (2016). Let me critically question this! Insights from a training study on the role of questioning on argumentative discourse. *International Journal of Educational Research*, 79, 195-210. <https://doi.org/10.1016/j.ijer.2016.05.017>
- Miras, M. (2019). Leer para aprender integrando varias fuentes. En I. Solé (Coord.), *Leer, comprender y aprender. Propuestas para favorecer el aprendizaje a partir de textos* (pp. 123-146). Horsori.

- Nussbaum, E. M., y Edwards, O. V. (2011). Critical questions and argument stratagems: A framework for enhancing and analyzing students' reasoning practices. *Journal of the Learning Sciences*, 20(3), 443-488. <https://doi.org/10.1080/10508406.2011.564567>
- Pérez, A., Potocki, A., Stadtler, M., Macedo-Rouet, M., Paul, J., Salmerón, L., y Rouet, J. F. (2018). Fostering teenagers' assessment of information reliability: Effects of a classroom intervention focused on critical source dimensions. *Learning and Instruction*, 58, 53-64. <https://doi.org/10.1016/j.learninstruc.2018.04.006>
- Rouet, J. F., y Britt, A. (2014). Multimedia learning from multiple documents. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 813-841). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.039>
- Solé, I. (2019). Textos y tareas de lectura. En I. Solé (Coord.), *Leer, comprender y aprender. Propuestas para favorecer el aprendizaje a partir de textos* (pp. 41-71). Horsori.Stevens, E. A., Murray, C. S., Fishstrom, S., y Vaughn, S. (2020). Using question generation to improve reading comprehension for middle-grade students. *Journal of Adolescent & Adult Literacy*, 64(3), 311-322. <https://doi.org/10.1002/jaal.1105>
- Tseng, A. S., Bonilla, S., y MacPherson, A. (2021). Fighting "bad science" in the information age: The effects of an intervention to stimulate evaluation and critique of false scientific claims. *Journal of Research in Science Teaching*, 58(8), 1152-1178. <https://doi.org/10.1002/tea.21696>
- Vidal-Abarca, E., Gilabert, R., Martínez, T., Sellés, P., Abad, N., y Ferrer, C. (2007). *Manual Test de Estrategias de Comprensión*. TEC. Instituto Calasanz de Ciencias de la Educación.
- Wilkinson, A., Reznitskaya, A., Bourdage, K., Oyler, J., Glina, M., Drewry, R., Kim, M., y Nelson, K. (2017). Toward a more dialogic pedagogy: Changing teachers' beliefs and practices through professional development in language arts classrooms. *Language and Education*, 31(1), 65-82. <https://doi.org/10.1080/09500782.2016.1230129>

Fecha de recepción: 20 diciembre, 2023.

Fecha de revisión: 11 enero, 2024.

Fecha de aceptación: 22 mayo, 2024.

Apéndice A

Ejemplo de uno de los textos elaborados para el estudio.



Los expertos dicen que la dieta keto no es sostenible. Entonces, ¿por qué es tan popular?

Por Sandee LaMotte
06:30 6 Enero 2020

Adelgazar de forma saludable no es fácil. Según los datos de la Encuesta Nacional de Salud del año 2017, el 54,19% de las personas de entre 15 y 24 años está siguiendo alguna dieta para perder peso. Entre los 25 y los 34 años, la cifra alcanza el 66,26%. Las cifras hablan por sí solas y todos nos podemos hacer una idea de la desesperación de algunas personas para desprenderse del exceso de peso.

Hoy en día hay una gran variedad de "dietas milagro" que se anuncian como seguras para la salud. Sin embargo, el profesor Gregorio Varela, catedrático de Nutrición y Bromatología de la Facultad de Farmacia de la Universidad CEU Sant Pau y presidente de la Fundación Española de Nutrición (FEN), recuerda que estas dietas normalmente limitan los alimentos y las cantidades que se pueden consumir, disminuyendo la posibilidad de conseguir todos los nutrientes necesarios.

Una de las dietas que se ha hecho más popular actualmente a través de las redes es la dieta keto, que restringe la aportación de carbohidratos por debajo del 10%, por lo cual una persona que siga una dieta de 1800 kilocalorías no tendría que ingerir más de 45 gramos de carbohidratos en el día, explica Álex Pérez, nutricionista del Centro de Atención Primaria Vallcarca-Sant Gervasi (Barcelona). El resto se corresponde con un 15% de proteínas y un 75% de grasas.

Tanto el profesor como el nutricionista aseguran que este tipo de alimentación puede producir efectos psicológicos negativos y puede desencadenar trastornos del comportamiento alimentario, además de generar desequilibrios en el metabolismo de las personas. También destacan el conocido "efecto rebote" que nos puede hacer recuperar el peso perdido más un 6% adicional en el momento en que volvemos a comer con normalidad.

Estos expertos son claros: la dieta keto no es saludable. A pesar de que se elimina el consumo de productos nocivos para la salud, como por ejemplo golosinas, bollería, chocolates y harinas no integrales, también se restringe la ingesta de una serie de alimentos beneficiosos para la salud y se promueve el consumo de grasas y/o proteínas de manera excesiva. A pesar de que permite perder peso a corto plazo, esta dieta no favorece la adopción de unos adecuados hábitos dietéticos, prolongados en el tiempo y que permitan un peso saludable.

Apéndice B

Tarea de identificación y valoración de evidencias en el pre-test, post-test y mantenimiento.

Para analizar el contenido de este texto, te proponemos rellenar la siguiente tabla. Para hacerlo debes seguir los pasos que se indica a continuación:

- a) En primer lugar, identifica las ideas que aparecen en el texto y escríbelas en la primera fila de la tabla, de color azul.
- b) Responde a las preguntas de la primera columna para cada una de las ideas que has identificado.

| Escribe las ideas que aparecen en el texto | | |
|---|--|--|
| ¿Esta idea procede de una opinión personal o son aportadas por un experto? | | |
| ¿Cómo valorarías el nivel de experiencia y conocimiento del experto que presenta esta idea? | | |
| ¿Los datos o evidencias que se proporcionan proceden de estudios de calidad? Indica Sí o No | | |
| ¿Esta idea se apoya en datos o evidencias claras y bien explicadas? Indica Sí o No | | |
| ¿Se indica qué organismo o institución ha llevado a cabo estos estudios? Indica Sí o No | | |
| ¿Esta idea se puede relacionar con otras ideas del texto? Indica Sí o No | | |
| ¿Esta idea señala los problemas que presenta el argumento principal que se defiende en el texto? Indica Sí o No | | |

Apéndice C

Extracto del guión del vídeo que sirvió de modelo para la formulación y respuesta a pregunta críticas.

Participantes en el vídeo:

Carla

Lucía

Diálogo:

Carla (C): Hola Lucía, ¿qué haces?

Lucía (L): ¡Hola Carla! Estoy haciendo los deberes. Tengo que buscar información sobre cómo actúa la salmonelosis y he encontrado un artículo por Internet que creo que me puede ayudar.

C: ¿Y cómo lo harás?

L: Primero leeré el texto entero y después intentaré identificar lo más importante.

C: Pues yo te ayudaré a valorar la información que has encontrado, porque en el Instituto nos han dicho que es importante aprender a valorar la rigurosidad de la información, para no equivocarnos, ni dejarnos engañar.

L [lee el texto entero]:

La salmonelosis es una de las enfermedades de transmisión alimentaria más comunes. El informe de la Autoridad Europea de Seguridad Alimentaria y el Centro Europeo para la Prevención y el Control de Enfermedades indica que España fue el tercer país de la Unión Europea (UE) con más brotes detectados el 2018 de salmonella. En el estudio se precisa que el 2018 los estados de la UE notificaron 5.146 brotes de salmonella (622 en España), que afectaron 48.365 personas (6.803 de ellas eran españolas). Aun así (...)

C: Es muy interesante. Dime ¿cuál es la fuente? ¿de dónde has obtenido este documento?

L: Herald de Aragón Editora (Portal digital de noticias de Zaragoza, Huesca y Teruel)

C: ¿Es un medio fiable y de confianza?

L: Yo creo que sí, puesto que es un diario digital.

C: ¿Qué te hace creer que la fuente es o no es de confianza?

L: A mí me da confianza que sea un diario, aunque sea digital, porque las noticias acostumbran a ser contrastadas.

C: *Y es bastante actual, ¿no?*

L: *Sí del 2019*

C: *Vale, vuelve a leer el artículo, esta vez te haré preguntas para ayudarte a valorar la información que aporta.*

[Lucía vuelve a leer]

C: *¿Qué se ha investigado exactamente en el estudio?*

L: *El estudio habla de la cantidad de brotes de salmonela que hay en Europa, informando de cuáles son los países con más casos de esta enfermedad y, justamente, en 2018 España fue de los más afectados.*

C: *¿Esta es una evidencia o dato que evidencia la idea de que la salmonelosis es una enfermedad común?*

L: *Sí, porque aporta datos que demuestran que mucha gente ha sufrido esta enfermedad, por lo tanto con estas cifras se ejemplifica.*

C: *¿Crees que estos datos son veraces, son ciertos? ¿Cómo lo puedes saber?*

L: *Sí, ya que son datos aportados por la Autoridad Europea de Seguridad Alimentaria y el Centro Europeo para la Prevención y el Control de Enfermedades que tiene que ser un organismo de expertos que investigan sobre la salmonelosis, y las afirmaciones y evidencias que aportan considero que son de confianza.*

Apéndice D

Guía de preguntas críticas

| Criterios | Ejemplos de preguntas críticas |
|--|--|
| Claridad: Las informaciones que aparecen en los textos se explican de forma completa y comprensible. | ¿Queda claro cuál es la idea principal del texto? ¿Queda clara la relación entre las diversas ideas del texto? ¿Las ideas se expresan de manera comprensible? ¿Hay ejemplos que ayuden a entender el texto? |
| Exactitud: Los hechos se presentan de manera precisa y concreta. | ¿Son precisos los hechos que se presentan en el texto o dudas de su exactitud? ¿Hay pruebas en el texto que demuestren que los hechos que se presentan conducen a los resultados que se muestran? ¿Son del todo exactas las afirmaciones o dudas de su exactitud? ¿Qué se ha investigado exactamente en el estudio? |
| Veracidad: Los argumentos tienen que ser verificables. Es importante que estén apoyados por evidencia empírica. | ¿Qué tipo de evidencias (estudios, datos científicos, aportaciones de expertos, etc.) hay en el texto? ¿Crees que las evidencias aportadas por el texto son veraces, son ciertas? ¿Cómo lo podemos saber? ¿Todas las evidencias son igual de importantes para la argumentación del texto? ¿Por qué? ¿Conoces alguna evidencia científica que contradiga alguna de las informaciones del texto? |
| Lógica: Las afirmaciones de los textos tienen que ser formalmente correctas y lógicamente comprensibles. Además, los argumentos tienen que ser consistentes. | ¿Tiene lógica lo que está explicando? ¿Por qué? ¿Los argumentos son consistentes? ¿Tienen sentido las ideas que se presentan? ¿Se aporta una conclusión lógica? |

Multiperspectivismo: Se consideran diferentes perspectivas sobre un tema (es decir, se aportan argumentos a favor y en contra). Existe o no consenso acerca de una temática concreta.

¿Cuál es la perspectiva desde la cual se plantea el tema?
¿El texto presenta diferentes puntos de vista?
¿Y entre los dos textos?
¿Qué similitudes, diferencias o contradicciones hay entre los textos?
¿Qué otros puntos de vista y argumentos pueden existir sobre este tema a parte de los que se presentan en el texto?
¿Podría haber consenso en la comunidad científica sobre esta temática?

Identificación de expertos: Quien aporta las evidencias y cuál es su experiencia y nivel de conocimiento.

¿Las evidencias del texto son aportadas por expertos o son solo una opinión personal?
¿Nos podemos fiar de los ejemplos de personas que hayan vivenciado lo que explican en el texto igual que si fuera una evidencia aportada por un experto? ¿Por qué?
¿Cómo valorarías el grado de experiencia y conocimiento de los profesionales que aparecen en el texto?

Análisis de la fuente y canal de divulgación: Valorar si la fuente de la cual proceden los textos es fiable y de confianza.

¿Cuál es la fuente del texto?
¿La fuente es fiable y de confianza?
¿Qué te hace creer que la fuente es o no de confianza?
¿Podemos creer en todo lo que explica esta fuente? Si es que sí, ¿cómo puedes valorar su credibilidad? Si es que no, ¿qué tipo de fuentes consideras de confianza y por qué?

Apéndice E

Documento explicativo para la identificación y valoración de la calidad de las evidencias

Cuando leemos textos científicos sobre una temática concreta, las ideas que se exponen habitualmente se presentan acompañadas de evidencias para apoyar los argumentos. Pero, ¿son de igual importancia todas las evidencias? ¿Existen grados o tipos diferentes de evidencias?

Vamos a recopilar algunos de los ejemplos que hemos visto en los textos que hemos leído con anterioridad

Kim Kardashian fue la famosa que dio a conocer esta dieta a través de las redes sociales, mostrando una pérdida de 25 kg después de haber sido madre en un periodo corto de tiempo.

Esta es una evidencia que muestra un ejemplo personal de una famosa a la que seguir la dieta keto le funcionó para perder peso rápidamente. Pero, ¿que le haya funcionado a ella quiere decir que le puede funcionar a todo el mundo? ¿Nos podemos fiar al 100% de esta evidencia a la hora de decidir si seguir o no este tipo de dieta? ¿Conocemos toda la información cuando se nos explica un ejemplo personal? Cuando nos presentan el caso de una persona no siempre contamos con toda la información y es muy probable que haya factores implicados que desconocemos. Tenemos que tener cuidado y analizar bien la información que nos dan, pero también valorar lo que no nos dicen. Además, ¿es igualmente confiable una experiencia personal que una evidencia aportada por un organismo, estudio o experto en el ámbito? ¿Son evidencias de la misma calidad? Cuando una información es aportada por un experto, por una institución científica o cuando hay datos contrastados por un estudio podemos considerar que la información que nos dan está contrastada, en cambio, cuando es una experiencia u opinión personal, no.

El profesor Gregorio Varela, catedrático de Nutrición y Bromatología de la Facultad de Farmacia de la Universidad CEU Sant Pau y presidente de la Fundación Española de la Nutrición (FEN), recuerda que estas dietas normalmente limitan los alimentos y las cantidades que se pueden consumir, disminuyendo la posibilidad de conseguir todos los nutrientes necesarios.

Esta es una evidencia aportada por un experto que argumenta en contra de las dietas milagro como la dieta keto, las cuales se basan en restringir alimentos y/o cantidades. A diferencia de la evidencia del ejemplo anterior, detrás de esta evidencia hay un experto. Entonces, ¿qué evidencia es más confiable, la vivencia de Kim Kardashian o las aportaciones de un experto en la materia? Como ya hemos comentado, una información aportada por un experto es de más credibilidad y confianza que los ejemplos de casos personales, puesto que la primera es más objetiva.

La experta en nutrición, Susie Burrell, asegura que no existen evidencias que la dieta keto afecte seriamente la salud.

Esta es otra evidencia aportada por un experto que, esta vez, argumenta a favor de la dieta keto explicando que no hay ninguna prueba que indique que seguir este tipo de alimentación afecta la salud. Aquí tenemos a otro experto, pero ¿tienen todos los expertos igual grado de profesionalidad? En este caso, nos indican que Susie Burrell es experta en nutrición, pero no tenemos ningún otro dato. En cambio, en la evidencia anterior nos dan más información sobre el grado de profesionalidad del experto citado al indicar que el profesor Gregorio Varela es catedrático de Nutrición y Bromatología de la Facultad de Farmacia de la Universidad CEU Sant Pau y presidente de la Fundación Española de la Nutrición (FEN). Por lo tanto, cuando analizamos evidencias aportadas por expertos tenemos que valorar su grado de profesionalidad/conocimiento. Habitualmente, aportará evidencias más veraces un experto que esté en una universidad o que trabaje en un organismo/institución del estado, que un experto que no esté formándose continuamente, ni informado de los últimos avances científicos.

Según los datos de la Encuesta Nacional de Salud del año 2017, el 54,19% de las personas de entre 15 y 24 años está siguiendo alguna dieta para perder peso. Entre los 25 y los 34 años, la cifra alcanza el 66,26%.

Esta es una evidencia que aporta datos que demuestran el alto porcentaje de personas que siguen algún tipo de dieta para perder peso. Si volvemos a compararla con una evidencia basada en experiencias personales como la primera, podemos tomar conciencia de que tiene más valor una evidencia sustentada por un organismo o institución científica que ha elaborado una encuesta y que aporta datos contrastados, que la vivencia de una persona, puesto que detrás de una experiencia personal hay muchos factores que desconocemos (intereses personales, económicos...).

Por lo tanto, en síntesis: son mejores las evidencias aportadas por organismos o instituciones reconocidas o por expertos independientes (vinculados a universidades o centros de investigación). Hay que poner bajo escrutinio o cuestionar evidencias aportadas por profesionales particulares (que pueden estar vinculados a empresas o negocios) y relativizar evidencias aportadas por personas no expertas en el ámbito que también pueden actuar movidas por intereses económicos y/o ideológicos.