Pérez García-Madrid, R., Castells, N. and Minguela, M. (2025). Assessing contradictory information: effectiveness of a program for the development of critical thinking in adolescents. *Revista de Investigación Educativa, 43*. DOI: https://doi.org/10.6018/rie.597921

Assessing contradictory information: effectiveness of a programme for the development of critical thinking in adolescents.

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

Rocío Pérez Garcia-Madrid*[1] , Núria Castells Gómez* and Marta Minguela Brunat*.

Department of Cognition, Development and Educational Psychology
University of Barcelona (Spain)

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.

[1] **Correspondence:** Rocío Pérez García-Madrid. rperezga15@doct.ub.edu, Paseo de la Vall d'Hebron, 171; Faculty of Psychology, Developmental and Educational Psychology Section; 08035 Barcelona.

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.

Introduction

We are currently exposed to a multitude of contradictory, misleading or false information on controversial topics. This can hinder the construction of knowledge, especially in the adolescent population. In fact, the results of recent research show that this population shows difficulties both in managing misinformation (Herrero-Curiel and La-Rosa, 2022; Herrero-Diz et al., 2021) and in the ability to argue critically about it (Castells et al., 2022), which compromises their ability to critically assess and interpret conflicting information. Considering that the training they receive on these aspects is scarce (Pérez et al., 2018), it is essential to design interventions to teach students to critically deal with texts, fostering their critical awareness or epistemic vigilance in relation to the information they read online on a daily basis (Gierth and Bromme, 2020a; Tseng et al., 2021).

Critical questions and analysis of evidence from socio-scientific texts with contradictory information

As List and Alexander (2019) point out, reading multiple sources on the same topic potentially leads to better comprehension and learning than reading a single document. However, reading sources with conflicting information requires skills that adolescent learners have not always developed (List and Alexander, 2019) and that are rarely taught (Solé, 2019). Succinctly, understanding multiple texts involves two related processes: on the one hand, understanding the relationship between the texts read (of complementarity or discrepancy) and their characteristics (source, authorship,

etc.), which Rouet and Britt (2014) call building an intertextual model ("Intertext Model"); on the other hand, it involves building an integrated representation of the content of the texts ("Integrated Model"). While integration is an essential component of achieving a deep understanding of multiple texts (contrasting and critically questioning information, answering questions, writing a text, etc.) (Barzilai et al., 2018; Miras, 2019), attending to intertextual features allows one to evaluate sources, and to determine the potential veracity of the information they convey. Thus, both processes are necessary for a critical reading of texts with contradictory information. However, given the difficulties students have in identifying conflicting information and evaluating it appropriately, it is essential to consider how to foster their critical thinking.

Critical thinking, according to Ennis (2018), is about seeking and providing alternative arguments; supporting a particular position when it is justified to do so; being well-informed; considering different points of view; and being aware of one's own beliefs. Achieving this requires reading and/or reviewing diverse texts to discern between true and false information, evidence and opinion. Indeed, Tseng et al. (2021) argue that critical awareness or epistemic vigilance in the face of information requires attention to both the evaluation of the characteristics of the sources (considering the authorship and reliability of the source) and the evaluation of the content of the text.

When insufficient information is available to evaluate sources, only the validity of the claims they contain can be assessed (Gierth and Bromme, 2020b). However, assessing the plausibility of information is complicated when there is also insufficient prior knowledge about a topic (Lombardi et al., 2018), as is often the case for adolescents on certain topics. In these cases, the only way to discern between conflicting information lies in the ability to identify and challenge arguments, which requires explicit teaching, as recently highlighted by several studies (Hruschka and Appel, 2023; Kiili et al., 2022; Tseng et al., 2021).

There are still few studies that test the effectiveness of educational programmes in this regard. Among them, Tseng et al. (2021) recently evaluated the impact of the use of a guide with questions aimed at critically assessing information on the Internet on secondary school students. Although a positive impact on the participants' ability to analyse diverse information was identified, improvements were limited, which the authors attribute to the lack of explicit instruction on the content of the guide. In another intervention, also based on the use of a guide, pairs of university students were taught to solve critical questions to evaluate different aspects of the arguments in the texts they were asked to read (in particular, clarity, accuracy, logic and multiperspectivism) (Mayweg-Paus et al., 2016). This intervention enabled students to critically evaluate the arguments in the texts. In line with these studies, in the field of teaching reading comprehension, the use of questioning is considered fundamental for the critical reading of texts, both to assess the authors' intentions and to evaluate evidence or identify biases (Cottrell, 2011). Furthermore, teaching students to generate questions has been shown to be effective (Stevens et al., 2020). Research on the approach to improving reading comprehension called Questioning the Author (Beck et al., 2020), for example, shows that learners can learn to formulate questions to help them understand and evaluate narrative and expository texts. However, we are not

aware of any studies aimed at encouraging the formulation of critical questions to evaluate the arguments presented in socio-scientific texts.

In addition to identifying arguments, it is essential to be able to critically assess the quality of the accompanying evidence. Evidence is information that is presented to support a claim (Du and List, 2020); it may be personal opinion or expert input, and it may come from an independent institution or body or from an entity with persuasive (political or financial) interests (Duncan et al., 2022). There are few studies assessing the quality of evidence (Cartiff et al., 2021), and those that have focused on how students evaluate this aspect have shown that they have serious difficulties in doing so (Kiili et al., 2022; Tseng et al., 2021).

Objectives

Based on the limitations identified in previous studies, the purpose of this study is to evaluate the effectiveness of an educational programme to promote critical thinking, focused on teaching secondary school students to develop critical questions and, through them, to recognise and assess the quality of both the arguments in texts and the evidence that accompanies those arguments. Within this framework, the specific objectives of this study focus on:

a) assess the impact of the programme on the students' ability to develop critical questions, in terms of quantity and variety.

b) to evaluate the impact of the programme on the ability to recognise and assess the quality of textual evidence.

In relation to these objectives, the following research questions (RQ) are posed:

PI1: Will students in the intervention group (IG) develop more and more varied critical questions to assess the quality of contradictory information provided by scientific texts than students in the control group (CG) after the intervention?

RQ2: Will IG students, after the intervention, be better able to identify evidence and assess it appropriately than their CG peers?

Method

Participants

Ninety 3rd year ESO students (M = 14.2 years old, SD = .402) from a 4-line secondary school in Barcelona (Group 1 n = 22; Group 2 n = 19; Group 3 n = 25; Group 4 n = 24), selected by convenience, took part in the study. Of the sample, 53.3% identified themselves as female and 46.9% identified themselves as male. The first two groups of this course formed the control group (CG, n = 41) and the other two, the intervention group (IG, n = 49). The distribution of the groups into IG and CG was randomised. The students were mostly native (only 5.5% came from outside the country), with a high

socio-economic level, and did not have specific educational support needs. The participants gave their informed consent to participate in this study. In order to respect their privacy, the first author gave them an alphanumeric coding at the time of creating the database, thus preventing them from being identified.

Instruments linked to data collection

Prior knowledge test

A test of prior knowledge of the topics in the area of natural sciences on which the texts to be used in this study would be based was developed, consisting of 13 multiple-choice questions with four response options. The reliability of this test was α = .65, which according to Hinton et al. (2004) represents moderate reliability. Other methodologists (e.g., Kerlinger and Lee, 2000) indicate that this level of reliability is acceptable depending on the use of the measure and the type of decisions made on the basis of the measure. In our study this measure is used as a control.

Reading comprehension test

Students' reading comprehension level was assessed using the Test of Comprehension Strategies (TEC) (Vidal-Abarca et al., 2007). This standardised test consists of two texts, each followed by ten multiple-choice questions that assess four reading comprehension processes: forming ideas, connecting those ideas, activating prior knowledge and forming macro-ideas. The authors of this test report a reliability of α = .80.

Pairs of texts for pre-test, post-test, and maintenance tests.

Three pairs of texts with contradictory arguments on controversial topics in the area of natural sciences were produced (one pair for each of the pre-test, post-test and maintenance phases; see example in Appendix A). The topics were selected according to the contents of the didactic sequences that were being worked on in the classroom. These six texts shared the following characteristics: number of words (approximately 350), format (4-5 paragraph news item), number of arguments and counter-arguments (see Table 1). Considering the arguments provided, one of the texts in the pair was always more rigorous than the other (e.g., more reliable sources than others; information from qualified professionals and/or experts without personal or economic interests vs. opinions of people with potential interests; scientific evidence from independent and rigorous entities vs. data of unknown origin).

Table 1

Characteristics of the texts used in the pre-test, post-test and maintenance.

Phases	**Texts (T): title, authorship, source and date of publication**	**No. of words**	**No. of arguments**			**No. of counter-arguments**	
			Without evident.	**With evidence.**	**Rel. Ideas**	**Without evident.**	**With evidence.**
Pre-test	T1: *Keto diet: Does eating more fat to lose weight work?* Clara Sánchez - TELVA 03/05/2018	336	5	4	2	1	0
	T2*: *Experts say the keto diet is not sustainable. So why is it so popular?* Sandee LaMotte- CNN 06/01/2020	362	5	5	2	1	0
Post-test	T3*: *Pollution of the air we breathe causes millions of* deaths Susan Golberg- National Geografic 02/04/2020	358	5	5	2	1	0
	T4: *US science adviser argues that polluted air helps children* Sean Reilly- EyE News 06/11/2019	350	4	3	2	0	1

Maintenance	T5: *The radical diet that can cure type 2 diabetes, says* groundbreaking *study* Miguel Sola- El confidencial 05/12/2017	342	4	4	2	1	0
	T6*: *Type 2 diabetes and nutrition* Mercedes Galindo Rubio- Spanish Diabetes Foundation 04/04/2019	346	4	4	2	0	1

Note: No evidence =no evidence, With evidence =with evidence, Rel. ideas=relate ideas
*The asterisk marks the texts of each pair that contained the most rigorous information.

Critical questions pre-test, post-test, and maintenance task

Each of the pairs of texts produced was accompanied by a task consisting of: 1) developing critical questions to assess the arguments in the texts; 2) identifying the main ideas (arguments) in each text; 3) identifying the evidence provided in the texts for each idea and assessing in each case who provided it (whether it was provided by experts or not), whether it came from studies (or not), and whether these studies were carried out by recognised bodies (or not). The task of identifying arguments and assessing evidence was organised in a table (see Appendix B).

Instruments linked to the intervention programme

Texts

For the intervention sessions, five argumentative texts were produced with similar characteristics to the texts of the pre-test, post-test and maintenance phases (see Table 2).

Table 2

Characteristics of the texts used during the intervention

Sessions (S)	**Texts (T): title, authorship, source and date of publication**	**No. of words**	**No. of arguments**			**No. of counter-arguments**	
			No evidence.	**With evidence.**	**Relationship ideas**	**No evidence.**	**With evidence.**
S2	T7: *Spain, third EU country with most salmonella outbreaks in 2018* Heraldo de Aragón 12/12/2019	436	5	4	2	0	1
S3	T8: *The end of antibiotics* Jordi Regàs and Sara Segarra - TV3 07/01/2018	346	4	4	2	1	0
		352	4	3	2	0	1
	T9: *Responsible use of antibiotics has benefits for everyone* Gencat 15/11/2017						
S4	T10*: *Even if the tobacco law is enforced, smoking is not so bad* Author unknown 02/01/2007	439	3	4	2	1	0
		396	4	5	2	1	1

T11: *Is it bad to smoke a cigarette a day?* Ángeles Gómez López 28/03/2018	

Note: The first session of the intervention programme consisted of the revision of the task carried out in the pre-test, so we worked on the texts used in that phase (T1 and T2, in Table 1).

*The asterisk marks text that contained information that was not rigorous.

Model video

Taking as a reference the use of model videos in other research (Mateos et al., 2018), a short video was recorded in which two students of the participating students' age exemplified how to use critical questions during the reading of an argumentative text to analyse the thoroughness of its content (specifically, T7, see Table 2). An extract from the video script can be found in Appendix C.

Guide to critical questions

Guide with examples of critical questions (see Appendix D) to assess different aspects of the arguments in the texts (clarity, truthfulness, logic, accuracy, multiperspectivism, identification of experts cited in the texts and characteristics of the sources) (Mayweg-Paus et al., 2016).

Explanatory document for identifying and assessing the quality of evidence

Guide with a collection of examples of evidence from the texts read in the pre-test and an explanation of their analysis, to help students assess the quality of the evidence. Based on the contributions of other authors (Duncan et al., 2022; Gough, 2021), the dimensions used for the analysis were: (1) provenance of the evidence (personal opinion vs. provided by experts); and (2) provenance of the data associated with the evidence (e.g., from studies carried out by recognised, independent bodies or institutions vs. from companies with economic interests) (see Appendix E).

Procedure

Pre-test phase

The pre-test was carried out in two sessions of 60 minutes each. In the first, students completed the test of prior knowledge and the TEC (Vidal-Abarca et al., 2007). In the second session, students were asked to read individually two argumentative texts (T1 and T2) and to complete the pre-test task using the Qualtrics programme[XM] . The same procedure was followed in all groups.

Intervention phase

The intervention was implemented in the subject of natural sciences, a subject in which it is common to read multiple argumentative sources with contradictory information on the same topic. The design of the programme was based on a socio-constructivist perspective on teaching and learning. This is reflected in aspects such as the use of students' prior knowledge as a basis for their learning, the diversification of aids and supports during the different sessions, and the progressive transfer of control and responsibility to the students (Fisher and Frey, 2013). In addition, the sessions were based on collective argumentation, inducing participants to read and reread the different texts and to question their answers (Wilkinson et al., 2017).

The intervention programme consisted of four 40-60 minute sessions. The first author implemented the intervention summarised in Table 3 with the IGs, while the CG teachers worked with the same texts in their classrooms. Semi-structured interviews were conducted with the CG teachers to ensure the use of these texts and to learn about the work that was done with them, thus ensuring that it basically referred to the comprehension of the text and not to the critical analysis of the arguments.

Table 3

Summary of intervention programme sessions, objectives and materials

Session (S) and duration	**Session activities**	**Materials used**	**Duties**
S1 (60 mins.)	Joint review of the pre-test to help understand the critical questions in the pre-test, and to review and assess the arguments, counter-arguments and evidence in the texts. -Explicit instruction in the development of critical questions.	-T1 and T2 -Students' responses to the pre-test	-Read T7

S2 (45 mins.)	- Viewing of the model video on the use of critical questions and use of the *Critical Questions Guide*. - Teaching the differences between the types of evidence from the *Explanatory Document for the selection and assessment of the quality of evidence.*	-T7 - Model video - Guide to critical questions - Explanatory document for the selection and assessment of the quality of evidence	-Read T8 and T9 - Develop critical questions and assess evidence from T8 and T9 (using the guide and the explanatory document).
S3 (40 mins.)	-Joint correction of homework to assess what has been learnt and what needs to be improved. -Answer critical questions of intertextual comparison (individual).	-Collection of students' responses to the homework assignment	-Read T10 and T11
S4 (60 mins.)	- Practice of critical questioning and evaluation of evidence (without supporting materials) from T10 and T11.	- T10 and T11 - Activity to formulate critical questions and identify and evaluate evidence (without the guide and explanatory document, but with the help of the researcher).	

Post-test and maintenance phases

The post-test and maintenance tasks were carried out in a session of one hour each. Students were asked to read a couple of texts (T3 and T4 for the post-test; T5 and T6 for the maintenance) and to solve the corresponding tasks individually (the same as in the pre-test using the Qualtrics programme[XM]). CGs and IGs solved these tasks under the same conditions, with the post-test taking place around two weeks after the end of the intervention and the maintenance one month after the post-test.

Data analysis

To answer the first research question, related to the comparison of the number and variety of critical questions posed by the IGs and CGs, we first proceeded to identify the critical questions posed by the students in the pre-test, post-test and maintenance. For this purpose, critical questions were considered to be those that lead to assessing the information and reflecting critically on it.

Secondly, the identified critical questions were classified according to the aspects they could contribute to assessing (specifically, clarity, accuracy, truthfulness, logic, multiperspectivism, identification of experts, and analysis of the source and channel of dissemination). The criteria used in each case, drawing on the contributions of authors such as Duncan et al. (2022) and Mayweg-Paus et al. (2016), were those presented in the Critical Questions Guide (see Appendix D).

To calculate the inter-rater agreement for this classification, 3 researchers coded the questions posed by 20 students from both groups, and for each of the data collection phases, independently, reaching an agreement of 90.7%. Discrepancies were resolved by discussion and one of the researchers finished coding the questions of the other participants.

Based on this coding, a count was made of the number of different types of critical questions developed in the pre-test, post-test and maintenance to identify the diversity of critical questions developed by the students.

To answer the second research question, related to the comparison of the ability to identify and evaluate evidence by CG and IG students, firstly, a count was made of the number of pieces of evidence that each student had correctly identified in the texts used in the pre-test, post-test and maintenance, and the percentage of correctly identified pieces of evidence was calculated from the total possible pieces of evidence in each case. Subsequently, in relation to the assessment of the evidence, the answers given by the students to the critical questions about the evidence were scored (0= incorrect, 1= correct) and the total number of correct answers given by each student for the pre-test, post-test and maintenance was counted.

Based on the coding of the critical questions and the students' evaluations of the evidence, two types of statistical analysis were carried out. On the one hand, to identify the degree to which students in each group (CG and GI) had improved throughout the different phases, the Friedman statistic was calculated and, when it was significant, the Wilcoxon test was used. On the other hand, to compare the performance of IG and CG students for the pre-test, post-test and maintenance, Mann-Whitney U tests were performed. Non-parametric tests were chosen because the data did not meet the assumption of normality.

Results

Prior knowledge and reading comprehension

Measures of prior knowledge and general reading comprehension (measured by TEC) (Vidal-Abarca et al., 2007) were taken as control variables. Before addressing the results linked to the research questions, mean comparisons between CG and IG for these measures are provided. The results of the Mann-Whitney U tests show that there were no significant differences between CG and IG students in either their level of prior knowledge of the subject matter of the texts used in this study (CG: M = 3.66, SD = 1.98, SD = 1.98).66, SD = 1.98; IG: M = 3.55, SD = 1.72; U = 983.5, Z = -.173, p = .863), nor in their level of general reading comprehension (CG: M = 13.54, SD = 3.66; IG: M = 13.39, SD = 3.79; U = 994.5, Z = -.081, p = .935).

Development of critical questions

In relation to the number of critical questions developed, the results of the non-parametric Friedman and Wilcoxon tests show that both GC ($\chi^2 = 16.147$; $p < .001$) and IG ($\chi^2 = 70.560$; $p < .001$) students significantly improved over the different phases of the study (pre-test and post-test Wilcoxon: GC: z = -3.333, $p = .001$; IG: z = -983.983, p = .001).001) improved significantly throughout the different phases of the study (Wilcoxon pre-test and post-test: CG: $z = -3.333$, $p = .001$; IG: $z = -5.983$, $p < .001$; pre-test and maintenance: CG: $z = -3.273$, $p = .001$; IG: $z = -5.889$, $p < .001$). Thus, both groups were able to elaborate more critical questions in the post-test and maintenance than in the pre-test.

However, the results of the Mann-Whitney U-test at pre-test, post-test and maintenance show that, although before the intervention there were no significant differences between CG and IG participants, after the intervention programme, IG students produced significantly more critical questions than CG students, both at post-test and maintenance ($p < .001$ in both cases and with a large effect size, see Table 4).

Table 4

Means, standard deviations, mean comparisons and effect sizes between IG and CG for the number of critical questions developed.

Phases	GC *M* (DT)	GI *M* (DT)	Comparison of averages	Size of the effect
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-β = 1; d = 1.64
Maintenance	1.02 (1.63)	5.41 (3.45)	$U = 214$; $Z = -6.54$; $p<.001$	1-β = 1; d = 1.62

In relation to the variety of critical questions asked, measured in terms of the number of types of critical questions asked, Friedman's test (CG: $\chi^2 = 12.603$; $p = .002$; IG: $\chi^2 = 71.326$; $p <.001$) showed that both GC and IG students evolved favourably throughout the different phases of the study (Wilcoxon pre-test and post-test: GC: $z = -3.093$, $p = .002$; IG: $z = -6.001$, $p <.001$; pre-test and maintenance: GC: $z = -2.914$, $p = .004$; IG: $z = -5.934$, $p <.001$), in both cases posing increasingly varied critical questions.

However, Mann-Whitney U analyses show that, although CG and IG did not differ in this respect at pre-test ($p = .354$), IG students asked significantly more varied critical questions after the intervention than their CG peers ($p < .001$ at both post-test and maintenance, and with a large effect size, see Table 5).

Table 5

Means, standard deviations and results of the comparison of means and effect sizes between IG and CG for the number of critical question types elaborated.

Phases	GC *M* (DT)	GI *M* (DT)	Comparison of averages	Size of the effect
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-β = 1; d = 2.01
Maintenance	.68 (1.08)	3.49 (1.89)	U = 205.5; Z = -6.64; p<.001	1-β = 1; d = 1.82

Additionally, the number of critical questions of each type produced by CG and IG was compared. As Table 6 shows, for all critical question types, students in both groups produced a similar number of questions before the intervention; however, IG students produced significantly more critical questions of the different types at both post-test and maintenance. In addition, between post-test and maintenance, IG students had an increase in the number of certain types of critical questions asked (questions on clarity, accuracy, and multiperspectivism).

Table 6

Means, standard deviations and mean comparison results of the different types of questions asked by IGs and CGs

Types of critical questions	Phases	GC *M* (DT)	GI *M* (DT)	Comparison of averages
Clarity	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
	Maintenance	.12(.458)	1.02(.829)	U = 356.5; Z = -5.89; p<.001
Accuracy	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
	Maintenance	.05(.218)	.61(.996)	U = 696.5; Z = -3.50; p<.001

Truthfulness	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
	Maintenance	.10(.300)	.41(.674)	U = 785; Z = -2.50; p = .012
Logic	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
	Maintenance	.02(.156)	.59(.814)	U = 595.5; Z = -4.41; p<.001
Multiperspectivism	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
	Maintenance	.00(.000)	.67(.944)	U = 594.5; Z = -4.57; p<.001
Identification of experts	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
	Maintenance	.20(.511)	.84(.943)	U = 575.5; Z = -4.03; p<.001
Analysis of the source and channel of dissemination	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
	Maintenance	.54(1.002)	1.27(1.238)	U = 607.5; Z = -3.46; p<.001

Identification and assessment of evidence

The results of the Friedman test indicate significant improvements in both groups in evidence identification (CG: χ^2 = 6.685, p = .035; IG: χ^2 = 70.578, p < .001). Specifically, the Wilcoxon test revealed significant differences for the IG (pre-test and post-test: z = -6.095, p <.001; pre- and maintenance: z = -6.118, p <.001), while in the CG, the significant improvement is evident between pre-test and post-test (z = -2.678, p =.007), but disappears at maintenance (z = -1.209, p = .227). Both groups also improve significantly in the assessment of evidence (Friedman: GI χ^2 = 74.351, p < .001; GC: χ^2 = 10.164, p = .006). For this variable the differences are significant both between pre-test and post-test results and between pre-test and maintenance in both groups (Wilcoxon: pre-test and post-test: GC: z = -2.987, p =.003; GI: z = -6.095, p <.001; pre-test and maintenance: GC: z = -2.276, p =.023; GI: z = -6.100, p <.001).

When comparing CG and IG in the different phases, Mann Whitney U tests show that these groups did not differ significantly in the pre-test either in the percentage of

correctly identified evidence (p = .069, see Table 7) or in the correct assessment of that evidence (p = .175, see Table 8), as measured by the total number of critical responses answered correctly. However, the results show that IG participants were significantly better in both aspects in the post-test (p < .001 in both cases and with a large effect size, see Tables 7 and 8). This significant difference was maintained at maintenance (p <.001 in both cases and with a large effect size, see Tables 7 and 8).

Table 7

Means, standard deviations and results of the comparison of means and effect sizes between IG and CG for the percentage of evidence correctly identified.

Phases	**GC** ***M* (DT)**	**GI** ***M* (DT)**	**Comparison of averages**	**Size of the effect**
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-β =1; d = 1.98
Maintenance	1.59 (2.02)	6.37 (2.24)	U = 150; Z = -7.08; p<.001	1-β = 1 ; d = 2.24

Table 8

Means, standard deviations and results of the comparison of means and effect sizes between IG and CG for the correct assessment of evidence from the total number of correct answers.

Phases	**GC** ***M* (DT)**	**GI** ***M* (DT)**	**Comparison of averages**	**Size of the effect**
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-β = 1; d = 2.08
Maintenance	4.98 (6.60)	23.02 (8.70)	U = 127.5; Z = -7.17; p<.001	1-β = 1; d = 2.33

Discussion

The aim of this research was to assess the effectiveness of an educational programme aimed at improving the critical thinking skills of 3rd year ESO students by teaching them to develop critical questions and to recognise and evaluate the evidence provided in texts based on these questions.

Although the results of the evolution of each of the groups over time show that both CG and IG students improve significantly in all the variables studied, IG participants obtained significantly better results than CG participants in all cases.

Thus, IG students asked significantly more and more diverse critical questions than CG students. The diversity of questions posed by IG students would allow for a more in-depth analysis of the arguments of the texts. Moreover, for some types of critical questions, the number of elaborated questions increases between post-test and maintenance, indicating that the participants can continue to generate relevant questions for the analysis of the texts after the intervention. This shows that students can incorporate and generate questions that facilitate the analysis of information, appropriating these resources, as studies related to reading comprehension have also shown (Beck et al., 2020).

IG students were also able to correctly identify more evidence and evaluate it more adequately than CG students. In addition, for both critical questions and evidence, the significantly higher results of the IG persisted in maintenance, which underlines the potential of the educational programme to promote lasting learning and lends validity to the intervention (Graham and Hebert, 2011).

The learning achieved by the IG students during the course of the educational programme can be attributed to several factors. Firstly, unlike the proposal by Tseng et al. (2021), in which a guide with critical questions was used, but students were not taught how to use it, in our study we opted for explicit instruction on the content of the intervention and the use of guides based on the progressive transfer of control to the students (Fisher and Frey, 2013). In addition, the resources used to teach how to develop critical questions and to identify and evaluate evidence, and the way they were used during the intervention, may possibly have favoured the effectiveness of the intervention, as the results of the effect size calculations suggest. On the one hand, the model video played by students of the same age as the participating pupils may have facilitated identification with the actresses, empathising with the difficulties, reflections and solutions negotiated between them. On the other hand, the guide of critical questions and the explanatory document for the evaluation of evidence were initially used within the framework of peer interaction; later, individually; and finally, once the contents of these resources had been assimilated and internalised, they were dispensed with. Thus, despite being a relatively short and intense educational programme, we understand that the fact of trying to progressively hand over responsibility to the students (Fisher and Frey, 2013), resorting to the use of external tools to facilitate the revision of information while they were internalising the use of these tools, contributed to providing them with the necessary resources to manage conflicting information. In this sense, it is important to highlight the importance of teaching how to contrast different texts on the same topic. Despite the difficulties involved in integrating information from texts that often have different structures, contents and intentions (Miras, 2019), it is precisely the possibility of contrasting evidence from different texts that allows students to perceive the interest of learning to formulate critical questions and to adequately assess the validity of the contents, when they do not have a sufficiently broad background of knowledge that allows for other types of analysis (Lombardi et al., 2018). Finally, although critical questions were used

in this case in the area of natural sciences, other research has used this resource in other areas and with different argumentative texts, obtaining equally satisfactory results with university students (Nussbaum and Edwards, 2011).

On the other hand, it would be interesting to investigate the reasons for the intra-group progress shown by CG pupils. In this respect, some hypotheses can be put forward. On the one hand, we know that 14-15 years of age is a turning point in students' argumentative capacity (Kuhn and Udell, 2007), which could explain the improvement in their critical abilities. On the other hand, the fact that they worked with texts with contradictory information, despite the fact that the CG teachers did not have guidelines for working with them critically, nor did they indicate that they had done so, may have encouraged processes of comparison and contrast that fostered a certain development of these critical abilities. To test these hypotheses would require a micro-genetic study of the teaching-learning processes carried out in CG classrooms. Although our study does not allow for such an analysis, the results obtained show that explicit teaching of the formulation and use of critical questions favours a better development of students' critical abilities.

Beyond the importance of the results obtained, which underline the potential of the programme, it would be interesting to be able to determine its impact in different contexts (e.g. schools with different characteristics). Moreover, the programme implemented has addressed certain dimensions of the evidence, but not many others, such as the way in which it is generated, validated, etc. (Duncan et al., 2022; Gough, 2021), which could be the focus of future lines of research.

References

Castells, N. García-Milà, M., Miralda, A., Luna, J., and Pérez, E. (2022), El razonamiento de los adolescentes para gestionar las noticias falsas. *Educación XX1: Revista de la Facultad de Educación*, *25*(2), 291-313. https://doi.org/10.5944/educxx1.31693

Barzilai, S., Zohar, A. R., & 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., & Sandora, C. A. (2020). *Robust comprehension instruction with Questioning the Author*. Guilford Publications.

Cartiff, B. M., Duke, R. F., & 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., and 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., and Chinn, C. A. (2022). The role of evidence evaluation in critical thinking: Fostering epistemic vigilance. In B. Puig, and M. P. Jiménez-Aleixandre (Eds.), *Critical thinking in Biology and Environmental Education, Contributions from biology education research* (pp. 75-92). Springer. https:/doi.org/10.1007/978-3-030-92006-7_5

Ennis, R. (2018). Critical thinking across the curriculum. *Topoi, 37*(1), 165-184. https://doi.org/10.1007/s11245-016-9401-4

Fisher, D., and Frey, N. (2013). *Better learning through structured teaching: A framework for the gradual release of responsibility* (2nd ed.). ASCD.

Gierth, L., and 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., and 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., & 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., and La-Rosa, L. (2022). Secondary school students and media literacy in the age of misinformation. *Comunicar: Scientific Journal of Communication and Education, 30*(73), 95-106. https://doi.org/10.3916/C73-2022-08

Herrero-Diz, P., Conde-Jiménez, J., and 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., and Brownlow, C. (2004). *SPSS Explained*. Routledge. https://doi.org/10.4324/9780203642597

Hruschka, T. M. J., and 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., and 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., & 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., and 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., & 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., & 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., & 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., & 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). Reading to learn by integrating multiple sources. In I. Solé (Coord.), *Leer, comprender y aprender. Proposals to promote learning from texts* (pp. 123-146). Horsori.

Nussbaum, E. M., and 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., & 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., and 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). Texts and reading tasks. In I. Solé (Coord.), *Leer, comprender y aprender. Proposals to favour learning from texts* (pp. 41-71). Stevens, E. A., Murray, C. S., Fishstrom, S., & 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., & 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., and Ferrer, C. (2007). *Manual Test de Estrategias de Comprensión. TEC*. Calasanz Institute of Educational Sciences.

Wilkinson, A., Reznitskaya, A., Bourdage, K., Oyler, J., Glina, M., Drewry, R., Kim, M., & 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

Date of receipt: 20 December, 2023.
Review date: 11 January, 2024.
Date of acceptance: 22 May, 2024.

Appendix A

Example of one of the texts produced for the study.

Experts say the keto diet is not sustainable. So why is it so popular?

By Sandee LaMotte
06:30 6 January 2020

Losing weight in a healthy way is not easy. According to data from the 2017 National Health Survey, 54.19% of people aged between 15 and 24 are following some kind of diet to lose weight. Between the ages of 25 and 34, the figure reaches 66.26%. The figures speak for themselves and we can all get an idea of how desperate some people are to shed excess weight.

Nowadays there is a wide variety of "miracle diets" that are advertised as safe for health. However, Professor Gregorio Varela, Professor of Nutrition and Bromatology at the Faculty of Pharmacy of the CEU Sant Pau University and President of the Spanish Nutrition Foundation (FEN), reminds us that these diets usually limit the foods and quantities that can be consumed, reducing the possibility of getting all the necessary nutrients.

One of the diets that has currently become most popular through the networks is the keto diet, which restricts carbohydrate intake to less than 10%, so a person following a diet of 1800 kilocalories should not eat more than 45 grams of carbohydrates a day, explains Álex Pérez, a nutritionist at the Vallcarca-Sant Gervasi Primary Care Centre (Barcelona). The rest corresponds to 15% protein and 75% fat.

Both the professor and the nutritionist state that this type of diet can produce negative psychological effects and can trigger eating disorders, as well as generating imbalances in people's metabolism. They also highlight the well-known "rebound effect", which can cause us to regain the weight we have lost plus an additional 6% when we start eating normally again.

These experts are clear: the keto diet is unhealthy. Although it eliminates the consumption of unhealthy products such as sweets, pastries, chocolates and unhealthy flours, it also restricts the intake of a number of foods that are beneficial to health and promotes the consumption of excessive fats and/or proteins. Although it allows short-term weight loss, this diet does not favour the adoption of adequate dietary habits, prolonged over time and allowing a healthy weight.

Appendix B

Task of identifying and assessing evidence in pre-test, post-test and maintenance.

To analyse the content of this text, we suggest you fill in the following table. To do so, you must follow the steps indicated below:
a) First, identify the ideas that appear in the text and write them in the first row of the table in blue.
b) Answer the questions in the first column for each of the ideas you have identified.

Write down the ideas that appear in the text	
Does this idea come from a personal opinion or are they provided by an expert?	
How would you rate the level of experience and knowledge of the expert presenting this idea?	
Is the data or evidence provided from quality studies? Indicate Yes or No	
Is this idea supported by clear and well-explained data or evidence? Indicate Yes or No	
Is it indicated which body or institution has carried out these studies? Indicates Yes or No	
Can this idea be related to other ideas in the text? Indicates Yes or No	
Does this idea point to problems with the main argument made in the text? Yes or No	

Appendix C

Excerpt from the script of the video that served as a model for the formulation and response to critical questions.

Participants in the video*:*
Carla
Lucia

Dialogue:
Carla (C): *Hi Lucia, what are you doing?*
Lucía (L): ***Hi Carla! I am doing my homework. I have to look for information on how salmonellosis works and I found an article on the internet that I think can help me.***
C: *And how will you do that?*
L: ***I will first read the whole text and then try to identify the most important points.***
C: *Well, I will help you to evaluate the information you have found, because at the Institute we have been told that it is important to learn to evaluate the accuracy of the information, so that we don't make mistakes or allow ourselves to be deceived.*
L [read entire text]:

Salmonellosis is one of the most common foodborne diseases. The report by the European Food Safety Authority and the European Centre for Disease Prevention and Control indicates that Spain was the third country in the European Union (EU) with the most salmonella outbreaks detected in 2018. The study states that in 2018 EU states reported 5,146 outbreaks of salmonella (622 in Spain), which affected 48,365 people (6,803 of them were Spanish). Even so, (...)

C: *That's very interesting. Tell me, what is the source? Where did you get this document?*
L: ***Heraldo de Aragón Editora (Digital news portal of Zaragoza, Huesca and Teruel)***
C: *Is it a reliable and trustworthy medium?*
L: ***I think so, since it is a digital newspaper.***
C: *What makes you believe that the source is or is not trustworthy?*
L: ***It gives me confidence that it is a daily newspaper, even if it is digital, because the news is usually verified.***
C: *And it's quite current, isn't it?*
L: ***Yes from 2019***
C: *OK, read the article again, this time I will ask you questions to help you assess the information it provides.*

[Lucía reads again].

C: *What exactly was investigated in the study?*
L: ***The study talks about the number of salmonella outbreaks in Europe, reporting which countries have the most cases of the disease, and in 2018 Spain was among the most affected.***

C: *Is this evidence or data that supports the idea that salmonellosis is a common disease?*
L: *Yes, because it provides data that shows that many people have suffered from this disease, so these figures exemplify it.*
C: *Do you think these data are true, are they true, how do you know?*
L: *Yes, because these are data provided by the European Food Safety Authority and the European Centre for Disease Prevention and Control, which has to be a body of experts investigating salmonellosis, and the claims and evidence they provide I think are reliable.*

Appendix D

Guide to critical questions

Criteria	Examples of critical questions
Clarity: The information in the texts is explained in a complete and comprehensible way.	Is it clear what the main idea of the text is? Is the relationship between the various ideas in the text clear? Are ideas expressed in a comprehensible way? Are there examples to help understand the text?
Accuracy: Facts are presented accurately and concretely.	Are the facts presented in the text accurate or do you doubt their accuracy? Is there evidence in the text to show that the facts presented lead to the results shown? Are the claims entirely accurate or do you doubt their accuracy? What exactly was investigated in the study?
Truthfulness: Arguments need to be verifiable. It is important that they are supported by empirical evidence.	What kind of evidence (studies, scientific data, expert input, etc.) is there in the text? Do you think that the evidence provided in the text is true? How can we know? Is all evidence equally important for the argumentation of the text? Why? Do you know of any scientific evidence that contradicts any of the information in the text?
Logic: The statements in the texts must be formally correct and logically comprehensible. In addition, the arguments must be consistent.	Is there a logic to what you are explaining, and why? Are the arguments consistent and do the ideas presented make sense? Is there a logical conclusion?
Multiperspectivism: Different perspectives on an issue are considered (i.e. arguments are made for and against). There is or is not a consensus on a particular issue.	What is the perspective from which the issue is approached? Does the text present different points of view, and between the two texts? What similarities, differences or contradictions are there between the texts? What other views and arguments can there be on this issue than those presented in the text? Could there be consensus in the scientific community on this issue?

Identification of experts: Who provides the evidence and what is their experience and level of expertise.	Is the evidence in the text provided by experts or is it just a personal opinion? Can we rely on the examples of people who have experienced what they explain in the text as if it were evidence provided by an expert? Why? How would you rate the degree of experience and knowledge of the professionals who appear in the text?
Analysis of the source and dissemination channel: Assess whether the source from which the texts originate is reliable and trustworthy.	What is the source of the text? Is the source reliable and trustworthy? What makes you believe that the source is or is not trustworthy? Can we believe everything this source explains? If yes, how can you assess its credibility? If not, what kind of sources do you consider trustworthy and why?

Appendix E

Explanatory document for the identification and assessment of the quality of evidence

When we read scientific texts on a particular subject, the ideas presented are usually accompanied by evidence to support the arguments. But is all evidence equally important, and are there different degrees or types of evidence?

We are going to compile some of the examples we have seen in the texts we have read before...

Kim Kardashian was the celebrity who unveiled this diet on social media, showing a loss of 25 kg after becoming a mother in a short period of time.

This is evidence that shows a personal example of a celebrity who found that following the keto diet worked for her to lose weight quickly. But does the fact that it worked for her mean that it can work for everyone? Can we rely 100% on this evidence when deciding whether or not to follow this type of diet? Do we know all the information when a personal example is explained to us? When a person's case is presented to us, we do not always have all the information and it is very likely that there are factors involved that we are not aware of. We need to be careful to analyse the information we are given, but also to assess what we are not told. Also, is a personal experience as reliable as evidence provided by an organisation, study or expert in the field? Is the evidence of the same quality? When information is provided by an expert, by a scientific institution or when there are data contrasted by a study, we can consider that the information we are given is contrasted, whereas when it is a personal experience or opinion, it is not.

Professor Gregorio Varela, Professor of Nutrition and Bromatology at the Faculty of Pharmacy of the CEU Sant Pau University and President of the Spanish Nutrition Foundation (FEN), reminds us that these diets usually limit the foods and quantities that can be consumed, reducing the possibility of getting all the necessary nutrients.

This is evidence provided by an expert who argues against miracle diets such as the keto diet, which are based on restricting foods and/or quantities. Unlike the evidence in the previous example, there is an expert behind this evidence. So which evidence is more reliable, Kim Kardashian's experience or the input of an expert in the field? As we have already mentioned, information provided by an expert is more credible and trustworthy than personal case examples, since the former is more objective.

Nutrition expert Susie Burrell says there is no evidence that the keto diet seriously affects health.

This is another piece of evidence from an expert who, this time, argues in favour of the keto diet, explaining that there is no evidence to suggest that following this type of diet affects health. Here we have another expert, but do all experts have the same degree of professionalism? In this case, we are told that Susie Burrell is an expert in nutrition, but

we have no other information. In contrast, in the previous evidence we are given more information about the degree of professionalism of the expert cited by indicating that Professor Gregorio Varela is Professor of Nutrition and Bromatology at the Faculty of Pharmacy of the CEU Sant Pau University and President of the Spanish Nutrition Foundation (FEN). Therefore, when we analyse evidence provided by experts, we have to assess their degree of professionalism/knowledge. Usually, an expert who is at a university or who works in a state body/institution will provide more truthful evidence than an expert who is not continuously training or informed of the latest scientific advances.

According to data from the 2017 National Health Survey, 54.19% of people aged between 15 and 24 are following some kind of diet to lose weight. Between the ages of 25 and 34, the figure reaches 66.26%.

This is evidence that provides data demonstrating the high percentage of people who follow some kind of diet to lose weight. If we compare it again with an evidence based on personal experiences like the first one, we can become aware that an evidence supported by a scientific body or institution that has carried out a survey and that provides contrasted data has more value than the experience of a person, since behind a personal experience there are many factors that we do not know (personal, economic interests...).

Therefore, in summary: evidence provided by recognised bodies or institutions or by independent experts (linked to universities or research centres) is better. It is necessary to scrutinise or question evidence provided by private professionals (who may be linked to companies or businesses) and to relativise evidence provided by non-experts in the field who may also be driven by economic and/or ideological interests.