# Uncovering adolescents profiles of critical thinking dispositions through fuzzy clustering: Implications for developmentally-tailored learning analytics and video literacy instruction

**Abstract**

This article investigates how integrated configurations of critical thinking dispositions - combining self-efficacy (SE), epistemic beliefs, and reasoning style - influence the evaluation of educational videos and the quality of school-related usage decisions among 1,119 secondary school students (776 middle students, 343 high students). Using fuzzy c-means clustering, three distinct disposition profiles were identified within each educational level. Among high school students, two subtypes of "high dispositions" emerged: an intuitive profile (high self-efficacy, low analytical reasoning) and an analytical profile (moderate self-efficacy, high analytical reasoning). Only the latter was associated with stronger critical thinking skills and resisted confirmation bias, whereas the intuitive profile, despite high confidence, did not outperform students with low dispositions. Among middle students, performance was primarily driven by the overall intensity of dispositions, with no moderating role of reasoning style. Effective critical thinking in high school depends on the coupling of motivation and analytical reasoning, rather than self-confidence alone. The study provides concrete avenues for differentiated pedagogical design and decision-support systems in digital education.

*Keywords:* Critical thinking dispositions, Video literacy, Self-efficacy, Epistemic beliefs, Cognitive reflection, Analytic vs. intuitive reasoning, Fuzzy clustering, Learning analytics, K-12 development, Educational decision-making

## 1. Introduction

Video consumption, particularly on platforms such as YouTube, has grown exponentially among adolescents (Twenge et al., 2019) and we observe a profound transformation in media practices: students increasingly prefer videos over texts for autonomous learning (Livingstone and Blum-Ross, 2020; Vogels et al., 2022). This transition toward an audiovisual information environment introduces new educational challenges, notably the dissemination of misleading or erroneous content (Kiili et al., 2019). In this context, the ability to identify reliable and relevant content has become a central yet complex component of digital literacy (Potocki et al., 2020; **?**).

Critical thinking, defined as the ensemble of dispositions and skills that enable the evaluation of the epistemic quality of information to determine the degree of trust warranted (), plays a foundational role in knowledge construction (Binkley et al., 2011). Although research emphasises the importance of developing these competencies from an early age (Chen et al., 2024; **?**), their effective implementation depends less on procedural mastery than on the voluntary mobilisation of stable cognitive and motivational dispositions (Facione, 1990; Facione et al., 1995; Ennis, 1996; **?**). This distinction between competence and disposition underpins contemporary approaches to critical thinking: possessing evaluative skills does not ensure their activation when encountering online information.

In France, this awareness has led to the integration of the critical evaluation of digital content into school curricula. Since 2016, the common framework of knowledge, skills, and culture (*Socle commun de connaissances, de compétences et de culture*[1]) has been introduced in middle school. It sets out what every student should know and be able to do by the age of 16. It brings together all the knowledge, skills, values, and attitudes needed to succeed at school, in life, and as a future citizen. Moreover, since 2019, the PIX certification[2] has been introduced in high school to promote and certify digital skills. These pedagogical frameworks aim to develop students' capacity to analyse videos in terms of their credibility and the epistemic trustworthiness of their sources (McPhee, 2016).

Yet videos present specific challenges: their transient and multimodal

[1] https://www.education.gouv.fr/le-socle-commun-de-connaissances-de-competences-et-de-culture-12512

[2] https://pix.org/en/certify-my-digital-skills

nature heavily taxes working memory, often leading to superficial evaluation strategies that neglect argument quality or speaker expertise (Mayer, 2002; Merkt et al., 2018). Despite these specificities, few studies have explored how students evaluate videos; research has primarily focused on textual documents (Abed and Barzilai, 2023). Consequently, the links between critical thinking dispositions and video processing remain poorly understood (Carmichael et al., 2018), limiting, for instance, the design of individualised pedagogical supports (Yoon et al., 2021).

The present study is part of the [anonym] project[3], funded by the French Ministry of National Education and aimed at evaluating the effectiveness of a digital platform designed to foster critical thinking through video analysis. It seeks to deepen the understanding of the cognitive processes and individual factors that influence how students evaluate online videos. This work extends the findings of [anonym], who demonstrated that certain dispositions (intellectual humility, self-efficacy, and epistemic beliefs) predict students' ability to discriminate between videos based on source expertise.

Compared to the reference paper, our contribution differs in four essential aspects:

1. *Methodology*: we adopt a person-centred (rather than variable-centred) approach by identifying integrated profiles of critical thinking dispositions through Fuzzy C-Means clustering (FCM).
2. *Theoretical design*: we extend their approach, which encompasses credibility judgements, including trustworthiness and information credibility, assessing the quality of educational decisions to use, namely the inclination to select videos authored by experts rather than those authored by laypersons for school presentations.
3. *Developmental perspective*: we extend the investigation of adolescents' critical thinking dispositions in predicting their evaluation of informational videos, examining the differences between middle and high school students. We also focus on the moderating role of reasoning style (intuitive vs. analytical), measured using an adapted version of the Cognitive Reflection Test (CRT-D) (Young and Shtulman, 2020).
4. *Sample size*: Although based on the same experimental protocol, our study relies on a distinct and substantially larger sample (N = 1,119 vs. 363), enabling robust profiling analyses.

[3] [anonym]

The novelty of our work resides in its dual goal: *(i)* characterising integrated configurations of cognitive and motivational dispositions, and *(ii)* assessing their impact on epistemic cognition as applied to video evaluation. Using the Fuzzy C-Means (FCM) clustering method (Bezdek, 1981), which remains underutilised in this domain, we provide a more nuanced understanding of critical thinking processes. This opens concrete avenues for differentiated pedagogical design and learning analytics systems. Three research questions guide our investigation:

- RQ1: Do critical thinking dispositions differ significantly between middle and high school students?
- RQ2: What profiles of critical thinking dispositions can be identified in the analysed population?
- RQ3: To what extent do these profiles influence *(a)* judgements of information credibility and source trustworthiness, and *(b)* the quality of educational usage decisions regarding videos among middle and high school students?

The article is structured as follows: Section 2 presents the theoretical framework and prior research on the critical evaluation of digital content and the dimensions of critical thinking dispositions. Section 3 describes the methodology, including the educational context, participants, data collection procedures, and analytical methods. Section 4 reports results on developmental differences, critical thinking profiles, and their influence on students' judgements and decisions. The discussion in Section 5 situates these findings in broader theoretical and empirical contexts. Sections 6 and 7, respectively, outline pedagogical and methodological implications and propose directions for future research.

## 2. Background

### *2.1. Digital content and videos: between bounded rationality and epistemic demands*

Given the vast amount of information available, individuals often have to assess the credibility of the content they encounter (Metzger and Flanagin, 2013). This cognitive demand operates within a framework of bounded rationality (Simon, 1955), where limited attentional and mnemonic resources constrain the depth of information processing. Under these conditions, they frequently employ heuristic strategies, so-called satisficing approaches, that prioritise temporal efficiency and cognitive economy over exhaustive analysis (Pirolli, 2005; Fu & Pirolli, 2007). While adaptive in many contexts, such strategies increase vulnerability to systematic biases, especially when the information is complex, conflicting, or deliberately misleading (Tversky & Kahneman, 1974). More specifically, critical evaluation of digital content relies on two interdependent types of judgements (Abed and Barzilai, 2023; Kiili et al., 2019; Kohnen et al., 2022; Anonymous, 2025): *i)* judgement of information credibility, concerning the perceived quality of the message (accuracy, completeness, evidentiary justification) *ii)* judgement of the source's epistemic trustworthiness, focused on perceived expertise, integrity, and benevolence. This evaluative process ultimately informs concrete decisions—such as sharing a video, incorporating it into a school presentation, or discarding it; thus, reflecting what Sperber et al. (2010) describe as contextualised epistemic vigilance.

Nonetheless, online videos, representing the dominant format among adolescents (Livingstone and Blum-Ross, 2020; Ofcom, 2023), impose specific cognitive constraints: their transient nature heavily taxes working memory (Mayer, 2002), while their multimodal character (voice tone, visual appearance, narrative style) fosters the emergence of affective heuristics (Sundar et al., 2021; Aljalabneh, 2024). These heuristics often lead individuals to prioritise surface-level cues, such as fluency of speech or perceived likability, over a rigorous assessment of epistemic validity (Irwan, 2024; McGrew et al., 2018). In this context, affect itself becomes a source of information, actively shaping judgements (Jonas et al., 2024; Schwarz et al., 2021). Against this backdrop, critical thinking dispositions (CTDs) play a crucial moderating role: they determine not only the ability to identify relevant epistemic cues but, more importantly, the willingness to invest the cognitive resources necessary for their thoughtful processing. As (Anonymous, 2025) emphasise, these dispositions vary across three interrelated dimensions: motivational (self-efficacy), epistemic (schemes of knowledge justification), and cognitive (reasoning style). The present study extends this insight by adopting a person-centred, profile-based approach, moving beyond the isolated effects of individual dispositions to identify integrated configurations that differentially modulate both evaluative judgements and educational usage decisions.

### 2.2. Profiles of critical thinking dispositions in video evaluation

Critical thinking does not reduce to a set of technical skills; it requires the integrated mobilisation of motivational, epistemic, and cognitive factors (Facione, 1990; Ennis, 1996). In the context of online video evaluation, three interdependent dimensions structure dispositions toward critical thinking (Anonymous, 2025):

1. **Self-efficacy (SE)**: Defined as the belief in one's ability to successfully perform a specific task (Bandura, 1997), SE acts as a foundational motivational lever. High SE promotes engagement in cognitively demanding tasks(Hayat et al., 2020; Tang et al., 2022) and predicts more accurate video evaluations (Anonymous, 2025). However, this relationship is not linear: poorly calibrated self-efficacy, disconnected from actual competence, can lead to epistemic overconfidence, characterised by underestimating task complexity and limited recourse to verification (Dunlosky and Rawson, 2012; Anttonen et al., 2024; Pennycook et al., 2021).
2. **Epistemic beliefs - schemes of knowledge justification**: Evaluating a video also involves activating beliefs about what constitutes valid knowledge. Three justification schemes consistently emerge in the literature (Anmarkrud et al., 2019; Hämäläinen et al., 2021):
    - *Justification by authority* (**JA**): reliance on the source's status;
    - *Justification by multiple sources* (**JMS**): seeking concordance through cross-verification;
    - *Justification by personal knowledge* (**JPK**): anchoring validity in personal experience or prior beliefs.

    These orientations are not mutually exclusive, but they entail distinct epistemic practices: a strong preference for personal justification is associated with limited sourcing, whereas valuing source multiplicity promotes corroboration behaviours (Barzilai et al., 2015).
3. **Reasoning style**: this dimension is measured through the Cognitive Reflection Test (CRT) (Frederick, 2005) and its developmental version (CRT-D) (Young and Shtulman, 2020), which measures the tendency to inhibit intuitive responses in favour of more deliberate reasoning. In the video context, this distinction predicts the ability to resist surface-level biases (Pennycook and Rand, 2019; Gervais, 2015). Information processing oscillates between two modes: a fast, intuitive, and heuristic

*System 1* with low CRT, and a slow, analytical, and reflective *System 2* with high CRT (Kahneman, 2011; Evans and Stanovich, 2013).

Identifying profiles that combine SE with reasoning style is therefore essential to distinguish students who are confident based on “sound reasoning” from those whose confidence rests on “uncritical intuition”. Moreover, the originality of our study lies in examining how the three justification schemes (epistemic beliefs) interact with SE and reasoning style to form differentiated evaluator profiles. We hypothesise that reasoning style does not operate in isolation but moderates the effects of SE and epistemic beliefs: for instance, high SE should ensure rigorous evaluation only when coupled with an analytical reasoning style (*System 2*).

Our goal is to identify typical configurations within the student population that combine varying levels of self-efficacy, epistemic beliefs, and reasoning style to examine how these profiles differ in their judgements and decisions regarding videos. This perspective is particularly relevant for informing pedagogical design and personalising digital learning tools, as educators and adaptive systems interact not with isolated variables, but with individuals who embody integrated dispositional configurations.

### *2.3. Toward an original contribution: profiling, decision-making, and developmental differentiation*

The literature on digital content evaluation has primarily focused on *i)* judgement processes (credibility, trustworthiness), *ii)* associated individual factors (dispositions, competencies), and *iii)* tasks centred on textual documents, with only recent, and still limited, attention devoted to videos (Abed and Barzilai, 2023). The work of (Anonymous, 2025) represents a significant advance by demonstrating that certain critical thinking dispositions (CTDs) predict students' ability to discriminate between videos based on source expertise. Nevertheless, three structural limitations warrant the present extension:

1. Methodological: their variable-centred approach does not capture the complex interactions among CTD dimensions.
2. Operational: their analyses focus on evaluation scores alone, without examining the quality of educational usage decisions—a crucial dimension in learning contexts.

3. Developmental: their comparison between middle and high school students remains at a global level, without exploring how CTD profiles evolve with cognitive maturation.

Our study addresses these gaps by: *i)* identifying CTD profiles through Fuzzy C-Means clustering and adopting a person-centred perspective, *ii)* assessing their impact not only on evaluative judgements but also on pedagogical usage decisions, and *iii)* analysing developmental trajectories between middle and high school, with particular attention to the moderating role of reasoning style (System 1 - intuitive vs. System 2 - analytical). As we will illustrate in Section (XXX) this approach opens concrete avenues in learning analytics and instructional design: by shifting from a logic of prediction to one of profiling, it enables the design of differentiated interventions tailored to the specific needs of each learner type.

## 3. Methodology

### *3.1. Ethical Statement*

The study strictly adheres to ethical guidelines governing research involving minors in educational settings. The protocol was approved by the Research Ethics Committee of the University of [anonymous] (reference: 2023_650) and was formally validated in advance by the relevant academic authorities. All participants received age-appropriate information about the study's objectives, duration, and procedures. Informed consent was obtained from both students and their legal guardians prior to data collection. Participants were explicitly assured of their right to withdraw at any time without any negative consequences for their academic progression. All data were collected anonymously; no personally identifying information was recorded. Only random identifiers and aggregated sociodemographic variables (grade level, gender, age) were retained. Data storage complied with the General Data Protection Regulation (GDPR) and was carried out on a secure university server.

### *3.2. Educational context*

The study was conducted as part of the [Anonymous] project, funded by the French Ministry of National Education. It involved a network of 17 public secondary schools (11 middle schools, 6 high schools), recruited in collaboration with regional academic authorities. The final sample comprised 1,119 students (mean age = 14.04; SD = 2.08), ranging from Grade 6 to Grade 12 (French *6ème* to *Terminale*). Forty participants were excluded due to excessive missing data (>20% of items missing), aberrant response patterns (e.g., invariant responses or clear contradictions with task instructions), or because they were considered outliers (Sect. 3.5.1). The sample was distributed as follows: 776 middle school students (mean age = 12.88; SD = 1.27) and 343 high school students (mean age = 16.21; SD = 1.16). Table 1 provides an overview of participants by school level and gender.

Table 1: Participants by school level and gender

| Genre | Middle school | High school |
|---|---|---|
| Female | 413 | 172 |
| Male | 346 | 157 |
| Other | 9 | 6 |
| PNS | 8 | 8 |

PNS = prefer not to say

This cohort differs from that utilised by Anonymous (2025) (N = 363), despite both originating from the same research programme. In our study, the larger sample size (N = 1,119) enables robust profile-based analyses, particularly within a person-centred framework (Sect. XXX).

### 3.3. Study design and data collection

The study employed a *within-subject* experimental design, closely replicating the protocol of Anonymous (2025) but enriched by an explicit focus on educational usage decisions. Data obtained via a questionnaire were gathered in person, with teachers present, during regular school hours in a single session lasting 50 to 60 minutes, using an online platform (LimeSurvey). The questionnaire unfolded in three sequential parts:

- Part 1 (**P1**): Sociodemographic questionnaire. Students reported their age, gender, grade level as illustrated in Table 1. Gender was not employed as a discriminant indicator in this study; however, it will be implemented in subsequent investigations.
- Part 2 (**P2**): Evaluation of experimental videos. Students watched four videos (45–55 seconds each) addressing the socio-scientific issue

(Mazac et al., 2025) regarding the reduction of meat consumption to facilitate the growth of organic farming. The four videos were created for an earlier research project that examined students' assessments of conflicting videos (Lescarret et al., 2024). They were composed of fabricated interviews in which the interviewee's positions on the issue (for vs. against reducing meat consumption) and their levels of expertise (expert vs. layperson) were conflicting. A banner was displayed at the bottom of the screen at the start of each video, showing the interviewee's first name and age in the layperson condition, and their name, occupation (agronomy researcher), and professional affiliation in the expert condition. In order to underscore the disparity in proficiency, specialists were interviewed in an office setting, while laypersons were interviewed while shopping at a supermarket.

The experimental design orthogonally crossed two factors: source expertise (expert vs. non-expert) and stance on the issue (in favour of vs. against reducing meat consumption). The four resulting video conditions are detailed in Table 2. For example, $V_1$ is from a non-expert in favour, while $V_4$ is from an expert against. To mitigate order effects, the video sequence was counterbalanced across classes; specifically, the videos were distributed randomly.

Table 2: Evaluation of experimental videos

| Attitude | Non Expert | Expert |
|---|---|---|
| In favour | Video 1 ($V_1$) | Video 2 ($V_2$) |
| Against | Video 3 ($V_3$) | Video 4 ($V_4$) |

After each video, students rated on 9-point Likert scales (1=not at all/very low; 9=completely/very high): *(i)* the epistemic trustworthiness (**ET**) of the source (perceived expertise, integrity, benevolence); *(ii)* the information credibility (**IC**) of the source (accuracy, completeness, justification); *(iii)* their intention to use the video for a school presentation (**SU**).

- Part 3 (**P3**): Measurement of critical thinking dispositions (Sect. 2.2). in this part, students answered several questions assessing: *(i)* self-efficacy (SE) specific to video evaluation (4 items); *(ii)* epistemic beliefs (12 items) structured into three subscales: justification by authority (JA), multiple sources (JMS), and personal knowledge (JPK); *(iii)* reasoning style, measured via a composite version of the Cognitive Reflection Test (CRT) (Frederick, 2005) (5 items) and its developmental adaptation (CRT-D) (Young & Shtulman, 2020), tailored for adolescents (3 items). We collected one response for each item, resulting in a total of 24 responses.

For more details on the questionnaire (anonymous link), please refer to Anonymous (2025).

### *3.4. Measures and operationalisation of variables*

The operationalisation of variables and their corresponding measures were organised into two main categories: *i)* video-related judgements and *ii)* critical thinking dispositions,

#### *3.4.1. Video-related judgements*

For each of the four videos ($V_1$, $V_2$, $V_3$, $V_4$), we computed the following indicators :

- Epistemic trustworthiness of the source (ET): this statistic represents the average rating of (xxx) items evaluating perceived source expertise, integrity, and benevolence;

- Francisco, à ajouter ET_score

- Information credibility (IC): the average rating of (xxx) items evaluating the accuracy, completeness, and justification of the message;

- School usage intention (SU): a single-item score reflecting the student's stated likelihood of using the video for a school presentation.

- School usage intention quality ($SU_{score}$): a global performance score indicating the extent to which students preferentially selected expert-authored videos over non-expert videos for school usage. This index was computed by assigning a binary value to each video choice (1 = selected an expert video; 0 = selected a non-expert video) and summing these values across the four videos.

FRANCISCO: à ajouter: Please consider that we did not use the IC_score since...

### *3.4.2. Critical thinking dispositions*

The following composite indices were derived to capture students' critical thinking dispositions (CTD):

1. Self-efficacy (SE): the mean of 4 items assessing self-efficacy specific to video evaluation.
2. Epistemic beliefs: the mean of 12 items assessing epistemic beliefs specific to video evaluation, measured through three subscales (JA, JMS, JPK).
3. Reasoning style ($CRT\_score$): the total number of correct responses, combining scores from the Cognitive Reflection Test (CRT) and its developmental adaptation (CRT-D).

Table 3 summarises all constructed variables and their abbreviations.

Table 3: Variables used in the study

| Part | Measure | Description | Type | [Points] |
|---|---|---|---|---|
| Video-related judgements | ET ($V_1$, $V_2$, $V_3$, $V_4$) | Epistemic trustworthiness of the source (per video) | Mean | [0-1] |
| | $\text{ET}_{score}$ | Epistemic trustworthiness scoring quality for all videos | Sum | [0-4] |
| | IC ($V_1$, $V_2$, $V_3$, $V_4$) | Information credibility (per video) | Mean | [0-1] |
| | SU ($V_1$, $V_2$, $V_3$, $V_4$) | Intention to use the video for school presentation (per video) | Single value | [0-1] |
| | $\text{SU}_{score}$ | School usage intention quality for all videos | Sum | [0-4] |
| Critical Thinking Dispositions | SE | Self-efficacy | Mean | [0-1] |
| | JA | Epistemic beliefs: Justification by authority | Mean | [0-1] |
| | JMS | Epistemic beliefs: Justification by multiple sources | Mean | [0-1] |
| | JPK | Epistemic beliefs: Justification by personal knowledge | Mean | [0-1] |
| | $\text{CRT}_{score}$ | Ability to inhibit spontaneous responses (CRT and CRT-D) | Sum | [0-8] |

In accordance with methodological recommendations in the literature (Moeller, 2015; Little, 2024), all Likert-scale responses were standardised using the Proportion of Maximum Scaling (POMS) method, which rescales scores onto a [0, 1] interval. This transformation preserves the rank order of responses while ensuring metric comparability across items and subscales, thereby enhancing compatibility with standard statistical procedures. This enabled us to compare the scores of the answers to the questions.

### *3.5. Data Analysis*

All analyses were conducted using R (version 4.4.2), combining non-parametric comparative approaches (Kruskal–Wallis test, Mann–Whitney U tests, and Wilcoxon signed-rank tests) with exploratory fuzzy C-means clustering (via the `cmeans` function from the `e1071` package, version 1.7.16[4]). Non-parametric methods were preferred due to unequal group sizes (Table 1) and deviations from normality in the distribution of study variables (Table 3). To control the Type I error rate (i.e., false positives) associated with multiple comparisons (Noble, 2009), p-values were adjusted using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995). This correction was chosen because it offers greater statistical power and is less conservative than traditional alternatives, such as Bonferroni or Holm corrections (Glickman et al., 2014; Ryan, 2021)

### *3.5.1. Data preprocessing and quality control*

Data were first inspected for missing values, logical inconsistencies, and aberrant response patterns. Participants with more than 20% missing data on key variables or clearly non-serious response patterns (e.g., identical responses across all items, explicit disagreement with instructions) were excluded (n = 23: 20 middle schoolers, 3 high schoolers).

An initial exploratory data analysis of the entire dataset was conducted, using the methodologies outlined in the subsequent sections. Then, following established practices (Liu et al., 2019), we conducted multivariate outlier detection to minimise the disproportionate influence of atypical response profiles on cluster partitioning and enhance data structure (Nowak-Brzezińska and Gaibei, 2022). This procedure combined two complementary approaches: *(i)* robust Mahalanobis distance, which identifies global deviations from the multivariate centroid, and *(ii)* the Local Outlier Factor (LOF), which compares the local density of a point to that of its neighbours to detect local anomalies (Liu et al., 2019). Only participants flagged as outliers by both methods were excluded from the clustering analysis (n = 17: 9 middle schoolers, 8 high schoolers).

Results reported in Section 4 are based on the dataset without outliers, which yielded more stable partitions, as reflected by an improved average fuzzy silhouette score, indicating greater cluster compactness and separation.

### *3.5.2. Non-parametric analyses (RQ1, RQ3)*

To address RQ1 (*Do critical thinking dispositions differ significantly between middle and high school students?*), we compared the distributions of

[4]https://www.rdocumentation.org/packages/e1071/versions/1.7-17/topics/cmeans

critical thinking dispositions (CTD variables - Sect. 3.4.2) across school levels using Mann–Whitney U tests.

To address RQ3 (*To what extent do these profiles influence (a) judgements of information credibility and source trustworthiness, and (b) the quality of educational usage decisions regarding videos among middle and high school students?*), we examined how the identified profiles in RQ2, defined in Section 4.2, predicted outcomes from video-related judgements (ET, $\mathrm{ET}_{score}$, IC, SU, $\mathrm{SU}_{score}$ - Sect. 3.4.1). Between-profile comparisons were conducted using Kruskal–Wallis tests, followed by pairwise Mann–Whitney U tests with Benjamini–Hochberg correction to control the false discovery rate (Benjamini and Hochberg, 1995) whenever post-hoc contrasts were warranted. This non-parametric approach was selected to ensure robust inference while respecting the structure of our dataset.

### *3.5.3. Fuzzy C-means clustering analysis (RQ2, RQ3)*

To address RQ2 (*What profiles of critical thinking dispositions can be identified in the analysed population?*), we performed fuzzy C-means clustering separately for middle and high school students using the CTD variables: SE, JA, JMS, JPK, $\mathrm{CRT}_{score}$ (Table 3) since this approach was particularly suited to our data. The Hopkins statistic, a measure of cluster tendency (Hopkins and Skellam, 1954), was low (H = 0.35; computed with the `performance` R package, v0.12.4[5]), indicating a weak natural tendency toward partitioning and fuzzy boundaries between potential groups. In such cases, hard-clustering methods (e.g., K-means, hierarchical clustering) yield suboptimal performance (Rai et al., 2010), whereas fuzzy clustering accommodates the gradual, non-binary nature of critical thinking dispositions. The optimal number of clusters was determined by triangulating multiple validity indices: the Fuzzy Partition Coefficient (FPC; (Bezdek, 1974)) and the average fuzzy silhouette width (Campello and Hruschka, 2006). Solutions with k = 2 and k = 3 clusters were evaluated. We selected the three-cluster solution as it balanced statistical validity (peak silhouette score) with pedagogical interpretability, a key functional criterion in educational profiling (Aschenbruck and Szepannek, 2020). The fuzziness parameter ($m$) was optimised by varying it between 1.25 and 2.5 (Amane et al., 2023); $m = 1.75$ yielded the highest fuzzy silhouette index and was retained. The algorithm

[5] https://www.rdocumentation.org/packages/performance/versions/0.12.4

was run with a maximum of 100 iterations to ensure convergence while maintaining computational efficiency, and random initialisation was repeated 20 times to guarantee solution stability.

For subsequent comparative analyses (RQ3), each student was assigned to the cluster with the highest membership degree. This fuzzy clustering approach respects the continuous and overlapping nature of CTDs, avoids overly rigid categorisation, and provides a meaningful, interpretable foundation for analysing how CTD profiles relate to students' evaluative judgements and educational usage decisions.

## 4. Results

This section presents the findings organised by RQ. Descriptive statistics for all measures and detailed results of statistical tests (RQ1–RQ3) are provided in the supplementary materials to ensure transparency and analytical reproducibility.

### *4.1. Developmental differences in critical thinking dispositions (RQ1)*

As illustrated in Figure 1 and in the Kruskal–Wallis tests reported in Table 4, we observe significant developmental differences between middle and high school students across all CTD measures.

Table 4: Kruskal–Wallis tests. Comparison between middle and high school. $\varepsilon^2$ denotes effect sizes.

| Measure | H | *df* | $p$ | $\varepsilon^2$ |
|---|---|---|---|---|
| SE | 7.37 | 1 | $< .01$ | .01 |
| JA | 14.63 | 1 | $< .001$ | .01 |
| JPK | 13.44 | 1 | $< .001$ | .01 |
| JMS | 3.86 | 1 | $= .049$ | .003 |
| $\text{CRT}_{score}$ | 217.91 | 1 | $< .001$ | .19 |

For all metrics, high school students reported higher scores than middle schoolers (both in means and medians). The largest effect was observed for analytical reasoning ($CRT_{score}$), suggesting substantial cognitive maturation between early and late adolescence. In contrast, although statistically significant, the effect for JMS was negligible ($\varepsilon^2 = 0.003$), highlighting persistent difficulties in systematically deploying corroboration strategies—even among high school students—a finding consistent with Anmarkrud et al. (2019).

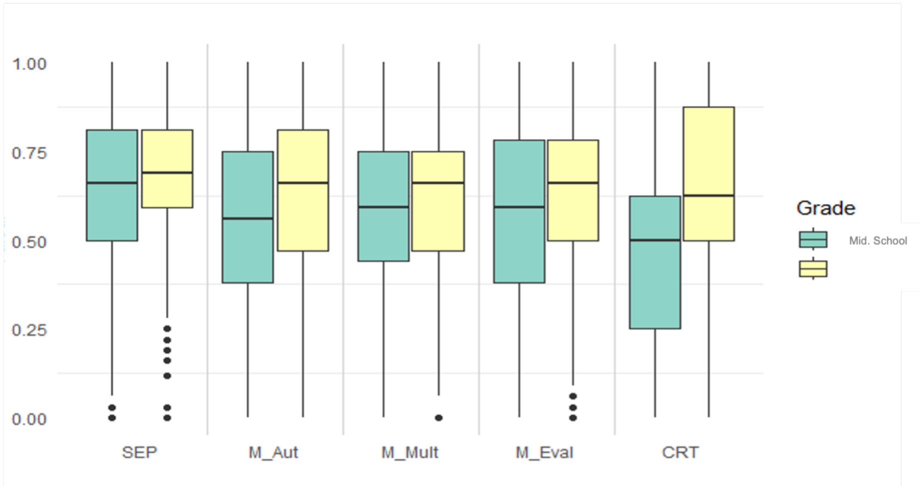


Figure 1: Distribution of the five measures selected to identify CTDs (middle school and high school students)

These developmental gaps reinforce the hypothesis of a progressive maturation of CTDs, shaped both by cognitive development (Steinberg, 2005) and by repeated exposure to institutional expectations regarding source evaluation (French national curricula, 2016; PIX certification, 2019 - Sect. 1).

### *4.2. Identification of critical thinking disposition profiles (RQ2)*

Fuzzy C-means clustering revealed three distinct profiles within each educational level (Figure 2, Table 5). These profiles are interpreted in light of the Kruskal–Wallis test results (Table 6).

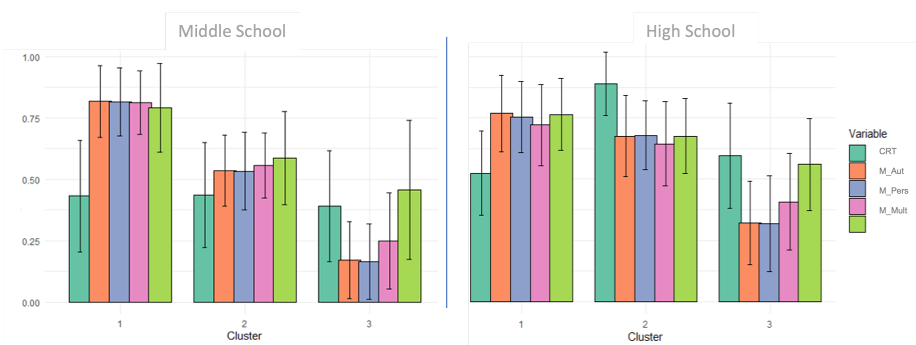


Figure 2: CTD profiles computed for middle and high school students

Table 5: Means and sample sizes for the different metrics by cluster and school type (SD in round brackets)

| School | Cluster | Size | CRT | SE | JA | JPK | JMS |
|---|---|---|---|---|---|---|---|
| Middle | 1 (High CTD) | 281 | 3.46 (1.81) | 7.33 (1.44) | 7.54 (1.16) | 7.53 (1.10) | 7.50 (1.03) |
| | 2 (Medium CTD) | 325 | 3.49 (1.71) | 5.69 (1.52) | 5.28 (1.16) | 5.27 (1.28) | 5.46 (1.06) |
| | 3 (Low CTD) | 170 | 3.14 (1.80) | 4.66 (2.25) | 3.28 (1.25) | 3.26 (1.23) | 3.00 (1.56) |
| High | 1 (High CTD - System 1) | 137 | 4.20 (1.38) | 7.11 (1.17) | 7.45 (1.23) | 7.03 (1.15) | 6.77 (1.32) |
| | 2 (High CTD - System 2) | 125 | 7.11 (1.03) | 6.41 (1.21) | 6.41 (1.31) | 6.43 (1.12) | 6.16 (1.38) |
| | 3 (Low CTD) | 81 | 4.77 (1.73) | 5.48 (1.52) | 3.57 (1.36) | 3.55 (1.56) | 4.26 (1.57) |

#### *4.2.1. Critical thinking disposition profiles among middle school students*

We identified three distinct profiles:

- Cluster 1 (High CTD, n = 281): Characterised by high levels of self-efficacy (SE = 7.33) and elevated scores across all epistemic justification dimensions (JA = 7.54; JPK = 7.53; JMS = 7.50). This profile corresponds to students who consistently report engaging in critical evaluation strategies. Notably, reasoning style (CRT_*score*) did not significantly differ from the other clusters ($\varepsilon^2 = 0.00$), indicating that analytic versus intuitive reasoning does not discriminate among middle schoolers.

- Cluster 2 (Moderate CTD, n = 325): Exhibits intermediate levels of SE (5.69) and epistemic justifications, reflecting occasional or context-dependent critical practices.

- Cluster 3 (Low CTD, n = 170): Displays low SE (4.66) and markedly low epistemic justification scores (all $<$ 3.00), indicating a pronounced disengagement from critical evaluation.

#### *4.2.2. Critical thinking disposition profiles among high-school students*

A developmental heterogeneity emerges among high school students, with two distinct subtypes of high CTD:

- Cluster 1 (High CTD – System 1, n = 137): characterised by high SE and elevated epistemic justifications, yet the lowest reasoning score ($\mathrm{CRT}_{score} = 4.20$), indicating a predominant reliance on intuitive/heuristic processing.

Table 6: Kruskal–Wallis ANOVA results on the clustering variables, by school type and group. $\varepsilon^2$ = ANOVA effect size; Post-hoc = direction of the pairwise comparisons for significant tests (see the supplementary document for calculation details).

| School | Features | $\chi^2$ | df | p | $\varepsilon^2$ | Post_hoc |
|---|---|---|---|---|---|---|
| Middle School | SE | 230.86 | 2 | $< .0001$ | .30 | c1 > c2 > c3 |
| | $\text{CRT}_{score}$ | 5.18 | 2 | ns | | |
| | JA | 560.12 | 2 | $< .0001$ | .73 | c1 > c2 > c3 |
| | JPK | 557.01 | 2 | $< .0001$ | .72 | c1 > c2 > c3 |
| | JMS | 518.04 | 2 | $< .0001$ | .67 | c1 > c2 > c3 |
| High School | SE | 63.32 | 2 | $< .0001$ | .19 | c1 > c2 > c3 |
| | $\text{CRT}_{score}$ | 179.12 | 2 | $< .0001$ | .52 | c1 > c2 > c3, c3 > c1 |
| | JA | 167.69 | 2 | $< .0001$ | .49 | c1 > c2 > c3 |
| | JPK | 162.86 | 2 | $< .0001$ | .48 | c1 > c2 > c3 |
| | JMS | 103.59 | 2 | $< .0001$ | .30 | c1 > c2 > c3 |

- Cluster 2 (High CTD – System 2, n = 125): shows slightly lower SE and epistemic justifications, but the highest reasoning score ($\text{CRT}_{score}$ = 7.11), reflecting sustained engagement in analytical thinking.

- Cluster 3 (Low CTD, n = 81): exhibits low SE (5.48), low epistemic justifications (all $< 4.26$), and an intermediate $\text{CRT}_{score}$ (4.77), suggesting the absence of a coherent evaluative strategy.

This distinction is absent among middle school students, for whom the reasoning style ($\text{CRT}_{score}$) does not discriminate between profiles. This pattern appears to reflect a late-adolescent cognitive specialisation, during which critical thinking becomes contingent upon the interaction between motivation (SE) and reasoning style. This developmental shift aligns with a well-documented trajectory: metacognitive capacities and the ability to inhibit heuristic responses (System 1) gradually strengthen throughout adolescence, driven by the maturation of the prefrontal cortex (Steinberg, 2005; Lee et al., 2013). Unlike younger students, high schoolers increasingly develop the capacity to couple self-confidence with deliberate analytical processing—a crucial condition for effective critical thinking (Kendeou et al., 2017). Notably, this specific interaction is absent in middle schoolers, a finding that converges with (Anonymous, 2025). In their variable-centred analysis, they found no significant effect of reasoning style, precisely because their sample (younger adolescents) had not yet reached the developmental stage where analytical reasoning modulates dispositional effects. This indicates that, before late adolescence, critical performance is predominantly influenced by the total intensity of inclinations, without the moderation of analytical thinking, a cognitive capacity that only becomes functional with age and neural development.

#### *4.2.3. Relationship between self-efficacy and epistemic justifications*

At both educational levels (middle and high school), self-efficacy was positively correlated with epistemic justifications (Figure 3, Table 7). In other words, the more confident students felt in their ability to evaluate videos, the more frequently they reported engaging in critical evaluation strategies. This association was consistent across both age groups. However, the form of this relationship diverged developmentally:

- In middle school, the relationship was linear: higher SE systematically corresponded to higher use of all epistemic justification strategies, following a clear trend: low SE < moderate SE < high SE (0<1<2).

- In high school, the relationship became non-linear and more nuanced. Only students with high SE (level 2) showed a significant increase in the use of justification by authority (JA) and justification by source multiplicity (JMS). In contrast, students with moderate SE did not differ significantly from those with low SE in their reported use of these strategies (Table 8).

This finding suggests that by late adolescence, self-confidence alone is no longer sufficient. A critical threshold of self-efficacy appears necessary to trigger the mobilisation of more demanding epistemic strategies, such as corroboration or deference to authority, indicating a qualitative shift in how motivation translates into critical practice.

Table 7: Correlation (Kendall's tau) between self-efficacy (SE) and the different epistemic justifications. Note: * $p < .05$, ** $p < .01$, *** $p < .001$

| School | Features | | JA | JMS | JPK |
|---|---|---|---|---|---|
| Middle School | disc.SE | Kendall Tau | 0.332 *** | 0.407 *** | 0.333 *** |
| | | p-value | < .001 | < .001 | < .001 |
| High School | disc.SE | Kendall Tau | 0.320 *** | 0.294 *** | 0.268 *** |
| | | p-value | < .001 | < .001 | < .001 |

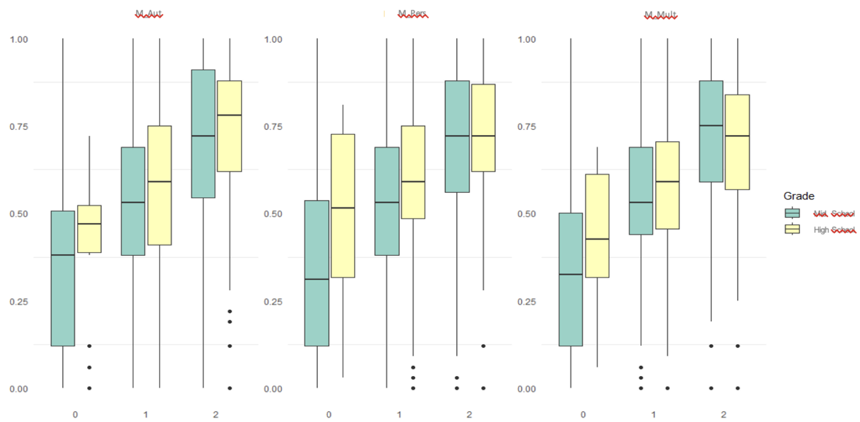


Figure 3: Distribution by school type of epistemic justification levels (JA, JPK, and JMS) by level of self-efficacy (discrete variable: $0$ = low, $1$ = medium, $2$ = high).

Table 8: Kruskal–Wallis ANOVA results testing the effect of SE (discretised variable, disc.SE) on students' epistemic justifications (monotonic values), by school type. $\varepsilon^2$ = ANOVA effect size; Post-hoc = direction of the pairwise comparisons for significant tests; three self-efficacy groups of disc. SE ($0$ = low, $1$ = medium, $2$ = high).

| School | Features | $\chi^2$ | *df* | *p* | $\varepsilon^2$ | Post_hoc |
|---|---|---|---|---|---|---|
| Middle School | JA | 126.35 | 2 | $< .0001$ | .16 | $0 < 1 < 2$ |
| | JPK | 187.80 | 2 | $< .0001$ | .15 | $0 < 1 < 2$ |
| | JMS | 120.00 | 2 | $< .0001$ | .24 | $0 < 1 < 2$ |
| High School | JA | 52.19 | 2 | $< .0001$ | .15 | $0 = 1, 2 > 0, 2 > 1$ |
| | JPK | 36.79 | 2 | $< .0001$ | .11 | $0 < 1 < 2$ |
| | JMS | 43.72 | 2 | $< .0001$ | .13 | $0 = 1, 2 > 0, 2 > 1$ |

### *4.3. Impact of critical thinking disposition profiles on judgements and decisions (RQ3)*

#### *4.3.1. judgements of epistemic trustworthiness and information credibility*

Middle school students exhibited strong inter-cluster differentiation. For all four videos, Cluster 1 (High CTD) rated both sources and information significantly more favourably than Clusters 2 and 3 (Table 10), following a consistent pattern: C1 > C2 > C3. This hierarchy was mirrored in their educational usage intentions, particularly for Videos 1, 2, and 4. In contrast, high school students showed globally homogeneous judgements, with no significant differences across clusters in most cases. Only two specific contrasts emerged: for Video 3 (layperson, against meat reduction - Table 9), Cluster 2 rated information credibility more negatively than the other clusters; for Video 4 (expert, against meat reduction), Cluster 1 judged source trustworthiness more positively. This stability suggests that, beyond a certain developmental threshold, high schoolers base their evaluations on more stable epistemic criteria, being less swayed by motivational fluctuations or dispositional intensity alone.

Table 9: Performance of video-use decisions and video sourcing processes among middle school and high school students

| **QUALITY OF DECISIONS (propensity to favor experts)** | | | | | |
|---|---|---|---|---|---|
| **School** | **H** | **df** | **p** | $\varepsilon^2$ | **Post-hoc** |
| Middle School | 14.08 | 2 | $< .001$ | .02 | $c1 > c2$, $c1 > c3$, $c2 = c3$ |
| High School | 12.18 | 2 | $< .01$ | .04 | $c2 > c3$, $c2 > c1$, $c1 = c3$ |
| **QUALITY OF SOURCING** | | | | | |
| **School** | **H** | **df** | **p** | $\varepsilon^2$ | **Post-hoc** |
| Middle School | 4.11 | 2 | *ns* | – | – |
| High School | 7.49 | 2 | .00 | .02 | *ns* |

$\varepsilon^2$ = ANOVA *effect size*. Post-hoc = direction of the *pairwise comparisons for significant tests.*

#### *4.3.2. Quality of educational usage decisions*

An analysis of decision reliability—operationalised as the tendency to select expert-authored videos—revealed a developmental reversal (Table 11). Among middle schoolers, Cluster 1 (High CTD) was the most accurate in its decisions (C1 > C2 = C3). Among high schoolers, Cluster 2 (High CTD – System 2) was the most accurate (C2 > C1 = C3). This finding is pivotal: it demonstrates that at the high school level, decision quality depends not on the overall intensity of dispositions, but on reasoning style. Students in Cluster 1 (high self-efficacy, intuitive reasoning) did not outperform those with low CTD, illustrating the risks of miscalibrated self-efficacy—a well-documented cognitive trap (Dunlosky and Rawson, 2012; Anttonen et al.,

Table 10: Kruskal–Wallis ANOVA results on the information credibility, epistemic trustworthiness, and video-use decision metrics. $\varepsilon^2$ = ANOVA effect size; Post-hoc = direction of the pairwise comparisons for significant tests (see the supplementary document for calculation details).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **EPISTEMIC TRUSTWORTHINESS** | | | | | | | |
| | Video | $H$ | $df$ | $p$ | $\eta^2$ | $\varepsilon^2$ | Post_hoc |
| Middle School | $V_1$ | 35.47 | 2 | .001 | .08 | .05 | c1 → c2 > c3 |
| | $V_2$ | 34.02 | 2 | .001 | .08 | .05 | c1 > c3, c2 > c3 |
| | $V_3$ | 27.71 | 2 | .001 | .07 | .05 | c1 > c2 > c3 |
| | $V_4$ | 24.24 | 2 | .001 | .06 | .05 | c1 → c2, c1 → c3 |
| High School | $V_1$ | 9.40 | 2 | .009 | .02 | .02 | c1 > c2, c1 > c3 |
| | $V_2$ | 7.42 | 2 | .024 | .02 | .02 | c1 > c3 |
| | $V_3$ | 6.78 | 2 | .034 | .02 | .02 | c1 > c3 |
| | $V_4$ | 9.11 | 2 | .011 | .02 | .02 | c1 → c2, c1 → c3 |
| **INFORMATION CREDIBILITY** | | | | | | | |
| | Video | $H$ | $df$ | $p$ | $\eta^2$ | $\varepsilon^2$ | Post_hoc |
| Middle School | $V_1$ | 59.46 | 2 | .001 | .11 | .08 | c1 > c2 > c3 |
| | $V_2$ | 50.27 | 2 | .001 | .10 | .07 | c1 > c2 > c3 |
| | $V_3$ | 40.19 | 2 | .001 | .09 | .07 | c1 > c2 > c3 |
| | $V_4$ | 36.60 | 2 | .001 | .08 | .07 | c1 > c2 > c3 |
| High School | $V_1$ | 7.30 | 2 | .026 | .02 | .02 | c1 > c3 |
| | $V_2$ | 7.97 | 2 | .019 | .02 | .02 | c1 > c3 |
| | $V_3$ | 8.63 | 2 | .013 | .02 | .02 | c1 > c3 |
| | $V_4$ | 9.22 | 2 | .010 | .02 | .02 | c1 → c2, c1 → c3 |
| **DECISION TO USE THE VIDEO** | | | | | | | |
| | Video | $H$ | $df$ | $p$ | $\eta^2$ | $\varepsilon^2$ | Post_hoc |
| Middle School | $V_1$ | 16.72 | 2 | .001 | .06 | .05 | c1 > c2 > c3 |
| | $V_2$ | 17.38 | 2 | .001 | .06 | .05 | c1 > c2, c1 > c3 |
| | $V_3$ | 17.22 | 2 | .001 | .06 | .05 | c1 > c2, c1 > c3 |
| | $V_4$ | 11.79 | 2 | .003 | .04 | .03 | c1 > c3 |
| High School | $V_1$ | 1.39 | 2 | n.s. | - | - | - |
| | $V_2$ | 0.76 | 2 | n.s. | - | - | - |
| | $V_3$ | 1.02 | 2 | n.s. | - | - | - |
| | $V_4$ | 1.29 | 2 | n.s. | - | - | - |

2024), wherein confidence is uncoupled from analytical vigilance. A two-tailed Wilcoxon test (for high school students, differences between Cluster 2 and Cluster 1 were further examined using a one-sided superiority test, as the results of the two-sided analysis indicated a trend, p = .058) confirmed that high school Cluster 2 (System 2) was significantly more accurate than middle school Cluster 1 (W = 20,679, p < 0.01, r = 0.22), despite comparable levels of self-efficacy—highlighting that reasoning style, not just motivation, drives mature critical performance.

#### *4.3.3. Resistance to motivational biases: The role of prior attitude*

Beyond judgements based on source expertise, the quality of critical thinking also manifests in the ability to override preexisting beliefs when evaluating counter-attitudinal information. This resistance to confirmation bias—the tendency to favour belief-congruent information—is considered an advanced epistemic competence (Kuhn et al., 2000). We therefore examined whether students' prior stance on meat consumption moderated their evaluations, and whether this sensitivity varied by CTD profile. As shown in Table 11, a clear developmental divergence emerged. Among middle schoolers, prior attitude significantly influenced judgements of epistemic trustworthiness (Wilcoxon: $W = 119{,}888$, $p = 0.02$, $r = 0.04$), with a marked preference for videos against meat reduction. This bias was especially pronounced in Clusters 2 and 3 (moderate and low CTD), where students consistently rated against-videos as more credible and trustworthy than pro-videos ($p < 0.05$, $r = 0.06$–$0.11$). This asymmetry suggests that, in the absence of robust critical dispositions, middle schoolers' judgements are guided by their prior beliefs—a pattern consistent with the cognitive superficiality hypothesis in video processing (Salmerón et al., 2020). In stark contrast, no significant effect of prior attitude was observed among high school students, either globally or within any cluster. This indicates that, by late adolescence, students develop epistemic autonomy—the capacity to decouple source evaluation from ideological preferences. This maturity coincides with the emergence of the "High CTD – System 2" profile (high CRT), which also demonstrated the highest decision accuracy. It is thus plausible that resistance to confirmation bias is tightly linked to a systematic reliance on analytical reasoning, which enables individuals to suspend judgement and consider alternative viewpoints. Far from being mere noise, prior attitude reveals a structural developmental divide: critical thinking in adolescence evolves from monitoring the source to monitoring the self. This metacognitive vigilance, absent in younger students, emerges gradually and becomes fully operative only when self-efficacy, epistemic beliefs, and analytical reasoning are harmoniously integrated—an equilibrium achieved by only a subset of high school students.

sistemare tabella 11

Table 11: Analysis of the effects of students' attitudes toward the videos' topic (in favor vs. against) on their judgements

By school type

| School | Variable | $W_s$ | $p$ | $\varepsilon$ | Description |
|---|---|---|---|---|---|
| Middle School | ET | 119,888 | .02 | .04 | against > in favor |
| | IC | 125,949 | ns. | – | – |
| High School | ET | 24,950 | ns. | – | – |
| | IC | 24,373 | ns. | – | – |

By school type and cluster

| School | Variable | Clusters | $W_s$ | $p$ | $\varepsilon$ | Description |
|---|---|---|---|---|---|---|
| Middle School | ET | 1 (High CTD →) | 70722.5 | .04 | .08 | against > in favor |
| | | 2 (Medium CTD) | 46043.5 | .01 | .11 | against > in favor |
| | | 3 (Low CTD) | 71254 | ns. | – | – |
| | IC | 1 (Intrans CTD →) | 3779 | ns. | – | – |
| | | 2 (Normac CTD) | 32727 | ns. | – | – |
| | | 3 (Intrans CTD) | 1528.5 | ns. | – | – |
| | | 4 (Intrans CTD →) | 2393 | ns. | – | – |
| | | 5 (Intrans CTD →) | 1947.5 | ns. | – | – |
| High School | ET | – | – | ns. | – | – |
| | IC | – | – | ns. | – | – |

$\varepsilon$ = *effect size* (Wilcoxon signed-rank test, paired samples; description: explanation of the effect)

# 5. Discussion

## *5.1. Development of critical thinking dispositions: progressive maturation (RQ1)*

Our findings confirm a significant developmental progression in CTD between early and late adolescence: high school students reported higher levels of self-efficacy, greater reliance on justification by authority (JA) and personal knowledge justification (JPK), and, most notably, a markedly more analytical reasoning style (CRT$_{score}$. This developmental gap aligns with well-established neurocognitive models: prefrontal cortex maturation during adolescence enhances cognitive regulation, the inhibition of heuristic processing (System 1), and engagement in deliberate, reflective reasoning (System 2) (Steinberg, 2005; Lee et al., 2013). Two results, however, need particular attention. First, the negligible effect size ($\varepsilon^2 = 0.003$) for justification by source multiplicity (JMS) suggests that—even at the high school level—the corroboration strategy, though institutionally promoted, remains poorly internalised. This finding resonates with (Anmarkrud et al., 2019), who emphasise the cognitive cost of cross-source verification, which often exceeds students' available resources. Our results indicate that this skill necessitates both motivation and automated cognitive routines, the latter likely emerging solely through regular, guided practice. Second, the absence of a relationship between self-efficacy (SE) among middle school students indicates that, at this stage, confidence in one's evaluative abilities has not yet become coupled with a specific mode of reasoning; a coupling that clearly emerges by late adolescence.

### *5.2. Critical thinking disposition profiles: Toward a developmental typology (RQ2)*

The clustering analysis reveals a cognitive specialisation emerging in late adolescence. Among middle school students, profiles are primarily distinguished by the overall intensity of CTDs: low, moderate, or high. In contrast, among high school students, two distinct subtypes of "high CTD" emerge, differentiated by reasoning style:

- High CTD – System 1: high self-efficacy (SE) and well-developed epistemic beliefs, but intuitive reasoning;
- High CTD – System 2: slightly lower SE, yet markedly analytical reasoning.

This distinction is absent among middle schoolers, for whom the Cognitive Reflection Test (CRT) does not differentiate clusters. It reflects an increasing integration between motivation and cognition: by late adolescence, critical thinking is no longer a matter of dispositional intensity alone, but of a specific configuration in which self-efficacy must be coupled with deliberate, analytical reasoning to become fully functional. This finding extends the literature on epistemic development (Kuhn et al., 2000) by demonstrating that cognitive maturation does not merely enhance the strength of dispositions, but reshapes their very structure—shifting from a unidimensional continuum to a multidimensional, functionally integrated system.

### *5.3. Impact of profiles on judgements and decisions*

#### *5.3.1. The overconfidence trap (RQ3)*

Our response to RQ3 underscores the central contribution of this study: the coupling of self-efficacy (SE) and reasoning style ($\mathrm{CRT}_{score}$) determines the quality of decisions but not the favorability of subjective judgements. Among middle school students, the High CTD profile (Cluster 1) issues more favourable judgements and makes the most accurate usage decisions. At this developmental stage, motivation alone appears sufficient to discriminate between expert and non-expert sources—likely because the task is perceived as school-related and inherently motivating, prompting engagement even in the absence of analytical reasoning. Among high school students, judgements become remarkably homogeneous, yet decisions diverge sharply: only students in Cluster 2 (High CTD – System 2) demonstrate significantly superior decisional performance. In contrast, students in Cluster 1 (High CTD – System 1), despite high confidence and self-reported use of critical strategies, do not outperform those with low CTD. This developmental reversal is critical: it illustrates the risks of miscalibrated self-efficacy (Dunlosky and Rawson, 2012; Anttonen et al., 2024). High confidence, when disconnected from analytical reasoning, can foster an illusory epistemic vigilance—a phenomenon Anonymous (2025) could not observe, as their younger sample (Mean age = 13.17) had not yet developed this dissociation, and their variable-centred approach failed to capture the SE × ($CRT_{score}$) interaction. Finally, the absence of prior attitude effects among high schoolers—in stark contrast to middle schoolers—confirms that the ability to overcome confirmation bias (Nickerson, 1998) emerges in late adolescence, likely due to the combined influence of cognitive maturation and curricular scaffolding.

### *5.4. Resistance to motivational biases: critical thinking as epistemic self-distancing*

The analysis of prior attitude, introduced as a complementary investigation, reveals a decisive developmental divide. Middle school students, particularly those with moderate or low CTD, systematically rated videos against meat reduction as more credible, revealing a clear confirmation bias (Nickerson, 1998). This pattern aligns with the cognitive superficiality hypothesis in video processing (Salmerón et al., 2020), where emotions and prior beliefs override argument analysis. High school students, by contrast, showed no sensitivity to their initial attitude, regardless of CTD profile. This epistemic autonomy reflects a mature capacity to decouple source evaluation from ideological preferences—an advanced hallmark of critical thinking (Kuhn et al., 2000). This shift co-occurs with the emergence of the System 2 profile, which also demonstrates the highest decisional accuracy, suggesting that critical maturity extends beyond monitoring the source to monitoring the self. Thus, far from being mere noise, prior attitude serves as a diagnostic marker of critical depth: its neutralisation signals the transition from heuristic judgement to authentic epistemic evaluation. This finding complements related work from the same research project (Anonymous2, 2024), where the authors identified attitude effects on video evaluation but did not examine developmental moderators. Our study demonstrates that only high school students with analytical reasoning skills successfully override this bias, thereby reinforcing our central thesis: effective critical thinking arises not from dispositional intensity alone, but from a synergistic integration of motivation and cognition.

## 6. Pedagogical and methodological implications

The findings of this study yield concrete implications at three interconnected levels: curriculum design, teacher education, and the development of digital pedagogical tools. They align with a growing body of research emphasising that critical thinking education cannot be one-size-fits-all but must be tailored to learners' cognitive-motivational profiles (Kendeou et al., 2017; Anmarkrud et al., 2019)

### *6.1. Toward profile-based differentiated instruction*

Our identification of three distinct critical thinking disposition profiles among middle schoolers and three more nuanced profiles among high schoolers, particularly the two subtypes of "high CTD", calls for a rethinking of differentiation in digital literacy instruction. Rather than issuing generic prompts such as "evaluate the reliability of this video", teachers could adapt their interventions based on the learner's diagnosed profile:

- For students with low CTD (middle or high school): Strengthen self-efficacy through scaffolded tasks (e.g., identify the source → assess its expertise → choose between two videos). These activities follow principle of "micro-successes", which gradually calibrate self-perceived competence and reduce cognitive overload.
- For students with high CTD – System 1 (intuitive high schoolers): Explicitly address the mismatch between confidence and analytical rigour. Cognitive decentration exercises (Kuhn et al., 2000), such as asking students to write a refutation of their chosen video or explain why a peer might reject it, can foster engagement of System 2 (analytical thinking).
- For all students: Systematically promote justification through source multiplicity, a skill found to be underdeveloped even at the high school

level. This requires moving beyond the analysis of a single video toward multi-video corroboration tasks, as advocated in collaborative fact-checking approaches (McGrew et al., 2018).

These recommendations resonate with the principles of explicit metacognitive instruction (), wherein epistemic strategies are not only taught but also embedded within authentic educational decision-making contexts.

### 6.2. Curricular alignment and teacher professional development

The observed developmental trajectory, particularly the maturation of analytical reasoning ($\text{CRT}_{score}$) and the emergence of distinct CTD profiles in late adolescence, highlights the need for a fine-grained articulation between middle and high school curricula in digital literacy. While the French socle commun (2016) and the PIX certification (2019) (Sect. 1) provide a useful framework, they remain insufficient if they fail to differentiate expectations by developmental stage. In middle school, instruction should prioritise the identification of surface-level cues (source expertise, speaker intentions) and the construction of self-efficacy. In high school, the focus should shift toward overcoming cognitive biases (confirmation bias, overconfidence) and cultivating analytical reasoning. Yet, current teacher education, both initial and in-service, often addresses these competencies only marginally (Wineburg and McGrew, 2019). We therefore propose integrating dedicated modules into **teacher training programmes (INSPE) on: *(i)* critical evaluation of educational videos; *(ii)* design of differentiated tasks; and *(iii)* interpretation of learner profiles using learning analytics. As  argue, effective teacher preparation must bridge the didactics of digital media and the cognitive psychology of the digital learner, a dual competence essential for fostering critical engagement in today's information-rich classrooms.

### 6.3. Design of digital tools and learning analytics

This study validates the relevance of person-centred approaches in Educational Technology. The use of fuzzy C-Means clustering not only aligns with prior research (), but further demonstrates the algorithm's capacity to uncover nuanced dispositional patterns in educational contexts, thereby opening new avenues for the development of intelligent adaptive systems:

- Digital platforms could diagnose students' critical thinking disposition profiles in real time, using either brief questionnaires or behavioural traces (e.g., interaction logs, pause frequency, source-corroboration behaviour).

- They could then deliver personalised feedback: for instance, to a "High CTD – System 1" student: "You selected this video with confidence. Take two minutes to consider whether an alternative viewpoint exists."

- Conversational agents (pedagogical chatbots) could be integrated to scaffold specific epistemic strategies, such as JA or JMS (), through guided dialogue and reflective prompts.

In this regard, our work contributes to a vision of data-informed critical thinking education (), in which learner data do not serve surveillance, but rather enable a fine-grained, ethical, and individualised pedagogy attuned to students' cognitive-motivational configurations.

## 7. Conclusion, limitations, and future directions

This study aimed to examine how integrated configurations of critical thinking dispositions, combining self-efficacy, epistemic beliefs, and reasoning style, influence not only judgements of information credibility and source trustworthiness but, more crucially, the quality of educational usage decisions (i.e., preference for expert-authored videos) among secondary school students. By adopting a person-centred approach on a large and representative sample (N = 1,119), we identified developmentally differentiated CTD profiles between middle and high school students, revealing subtle dynamics between motivation, cognition, and decisional performance.

However, despite the robustness of its methodology, this study reveals some limitations:

### 7.1. Methodological limitations

First, all indicators of critical thinking dispositions (CTD) rely on self-report measures. As (Dunlosky and Rawson, 2012; Anttonen et al., 2024) emphasised, self-efficacy is particularly prone to miscalibration relative to actual competence: students may report using critical strategies while applying them only superficially. Future studies should therefore triangulate self-reports with objective behavioural traces—such as video viewing time, pause usage, cross-source corroboration behaviours, or eye-tracking measures focused on epistemic cues (e.g., expertise indicators). Second, the assessment context, conducted in classrooms with an explicit mention of "critical thinking evaluation", exposes the study to social desirability bias (Paul et al., 2017). Students may have responded according to what they believed teachers or researchers expected. More ecologically valid protocols, for instance, asking students to autonomously search for videos on a simulated YouTube-like platform without explicit reference to "critical thinking", would mitigate this effect. Third, prior knowledge of the topic ("reducing meat consumption to promote organic farming") was not introduced in the analysis, although the literature clearly demonstrates that this factor strongly moderates source evaluation competence (). Integrating a factual knowledge pretest and analysing its interaction with CTD profiles would constitute a methodological advancement of primary importance. Finally, the ecological validity of our findings is constrained by the use of laboratory-produced videos featuring homogeneous actors (white males aged 40). While necessary for experimental control, this standardisation contrasts sharply with the sociodemographic diversity of content creators on YouTube (Li et al., 2020), limiting generalisability. Moreover, the exclusive use of male presenters prevents examination of gender effects on perceived expertise—a documented bias in source evaluation (Hendriks et al., 2015). Future replications using authentic videos, pre-screened for misinformation by expert raters, would significantly enhance external validity.

### *7.2. Conceptual limitations*

Two conceptual limitations also deserve attention. First, the person-centred approach provides a cross-sectional snapshot of CTD at a single point in time, precluding insight into their developmental dynamics. Yet, as Facione (1990) notes, these dispositions are inherently malleable and responsive to educational experiences. Future work should employ longitudinal designs or pre-post experimental interventions (e.g., cognitive decentration workshops) to model how profiles evolve and assess their plasticity. Second, our clustering model does not incorporate task-specific motivational variables: interest in the topic, perceived academic stakes, or affective responses to the videos. Yet, such factors strongly modulate cognitive engagement and epistemic processing quality (). Although our supplementary analysis revealed the moderating role of prior attitude—especially among middle schoolers—this variable was not integrated a priori into profile construction. A promising direction would be to enrich future profiling models with measures of subjective task value ()—including perceived utility, intrinsic interest, and perceived costs, to better capture the full complexity of epistemic judgement in authentic educational contexts.

### 7.3. Future research directions

Four key avenues emerge for future research. First, studies should examine whether CTD profiles predict not only final decisions but also online processing strategies (eye-fixations, pause usage, rewatching), enabling integrated modelling of epistemic cognition in real time. Second, within a design-based research framework (), profile-tailored interventions could be developed and evaluated—e.g., cognitive decentration activities specifically targeting "High CTD – System 1" students. Third, this typology should be extended to other media formats (TikTok, podcasts, infographics), accounting for their distinct modal and cognitive constraints (Sundar et al., 2021). Finally, co-design with teachers could yield practical diagnostic tools (e.g., classroom observation rubrics, behavioural indicators) to operationalise these profiles in daily practice and ensure the ecological transfer of research insights.

## Acknowledgements

[anonym]

## References

Abed, F., Barzilai, S., 2023. Can students evaluate scientific youtube videos? examining students' strategies and criteria for evaluating videos versus webpages on climate change. Journal of Computer Assisted Learning 39, 558–577. doi:10.1111/jcal.12754.

Aljalabneh, A.A., 2024. Visual media literacy: educational strategies to combat image and video disinformation on social media. Frontiers in Communication 9, 1490798. doi:10.3389/fcomm.2024.1490798.

Amane, M., Aissaoui, K., Berrada, M., 2023. Enhancing learning object analysis through fuzzy c-means clustering and web mining methods. Emerg Sci J 7, 799–807.

Anmarkrud, Ø., Andresen, A., Bräaten, I., 2019. Cognitive load and working memory in multimedia learning: Conceptual and measurement issues. Educational Psychologist 54, 61–83. doi:10.1080/00461520.2019.1583130.

Anonymous, 2025. Anonymous. Anonymous .

Anonymous2, 2024. Anonymous2. Anonymous2 .

Anttonen, R., Räikkönen, E., Kiili, K., Kiili, C., 2024. Sixth graders evaluating online texts: Self-efficacy beliefs predict confirming but not questioning the credibility. Scandinavian Journal of Educational Research 68, 1214–1230. doi:10.1080/00313831.2023.2192056.

Aschenbruck, R., Szepannek, G., 2020. Cluster validation for mixed-type data. Archives of Data Science, Series A 6, 02.

Barzilai, S., Tzadok, E., Eshet-Alkalai, Y., 2015. Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. Instructional Science 43, 737–766. doi:10.1007/s11251-015-9351-8.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) 57, 289–300.

Bezdek, J.C., 1974. Numerical taxonomy with fuzzy sets. Journal of mathematical biology 1, 57–71.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers, Boston, MA.

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., Rumble, M., 2011. Defining twenty-first century skills, in: Assessment and Teaching of 21st Century Skills. Springer, pp. 17–66. doi:10.1007/978-94-007-2324-5_2.

Campello, R.J., Hruschka, E.R., 2006. A fuzzy extension of the silhouette width criterion for cluster analysis. Fuzzy Sets and Systems 157, 2858–2875.

Carmichael, M., Reid, A., Karpicke, J.D., 2018. Assessing the impact of educational video on student engagement, critical thinking and learning. Technical Report. SAGE. URL: https://www.sagepub.com/sites/default/files/upm-binaries/33258_Chapter1.pdf.

Chen, J., Wang, X., Zheng, X., 2024. The investigation of critical thinking disposition among chinese primary and middle school students. Thinking Skills and Creativity 51, 101444. doi:10.1016/j.tsc.2024.101444.

Dunlosky, J., Rawson, K.A., 2012. Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. Learning and Instruction 22, 271–280. doi:10.1016/j.learninstruc.2011.10.003.

Ennis, R.H., 1996. Critical thinking dispositions: Their nature and assessability. Informal Logic 18, 115–135. doi:10.22329/il.v18i2.2378.

Evans, J.S.B., Stanovich, K.E., 2013. Dual-process theories of higher cognition: Advancing the debate. Perspectives on Psychological Science 8, 223–241. doi:10.1177/1745691612460685.

Facione, P., 1990. Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report). Technical Report. California Academic Press.

Facione, P.A., Sanchez, C.A., Facione, N.C., Gainen, J., 1995. The disposition toward critical thinking. The Journal of General Education 44, 1–25.

Gervais, W.M., 2015. Override the controversy: Analytic thinking predicts endorsement of evolution. Cognition 142, 312–321. doi:10.1016/j.cognition.2015.06.006.

Glickman, M.E., Rao, S.R., Schultz, M.R., 2014. False discovery rate control is a recommended alternative to bonferroni-type adjustments in health studies. Journal of clinical epidemiology 67, 850–857.

Hämäläinen, E.K., Kiili, C., Räikkönen, E., Marttunen, M., 2021. Students' abilities to evaluate the credibility of online texts: The role of internet-specific epistemic justifications. Journal of Computer Assisted Learning 37, 1409–1422. doi:10.1111/jcal.12581.

Hayat, A.A., Shateri, K., Amini, M., Shokrpour, N., 2020. Relationships between academic self-efficacy, learning-related emotions, and metacognitive learning strategies with academic performance in medical students: a structural equation model. BMC Medical Education 20, 76. doi:10.1186/s12909-020-01995-9.

Hendriks, F., Kienhues, D., Bromme, R., 2015. Measuring laypeople's trust in experts in a digital age: The muenster epistemic trustworthiness inventory (meti). PloS one 10, e0139309.

Hopkins, B., Skellam, J.G., 1954. A new method for determining the type of distribution of plant individuals. Annals of Botany 18, 213–227.

Irwan, I., 2024. Introduction to interactive video-based e-learning to improve critical thinking skills in vocational high school students. Unram Journal of Community Service 5, 176–181. doi:10.29303/jpk.v5i3.368.

Jonas, M., Kerwer, M., Chasiotis, A., Rosman, T., 2024. Indicators of trustworthiness in lay-friendly research summaries: Scientificness surpasses easiness. Public Understanding of Science 33, 37–57. doi:10.1177/09636625231185782.

Kahneman, D., 2011. Thinking, Fast and Slow. Farrar, Straus and Giroux, New York, NY.

Kendeou, P., Butterfuss, R., Van Boekel, M., O'Brien, E.J., 2017. Integrating relational reasoning and knowledge revision during reading. Educational Psychology Review 29, 27–39.

Kiili, C., Coiro, J., Räikkönen, E., 2019. Students' evaluation of information during online inquiry: Working individually or in pairs. The Australian Journal of Language and Literacy 42, 167–183.

Kohnen, A., Dawson, K., Mertens, G., 2022. Three continua of online credibility strategies used by eighth graders. TechTrends 66, 643–653. doi:10.1007/s11528-022-00722-5.

Kuhn, D., Black, J., Keselman, A., Kaplan, D., 2000. The development of cognitive skills to support inquiry learning. Cognition and instruction 18, 495–523.

Lee, K., Bull, R., Ho, R.M., 2013. Developmental changes in executive functioning. Child development 84, 1933–1953.

Lescarret, C., Magnier, J., Le Floch, V., Sakdavong, J.C., Boucheix, J.M., Amadieu, F., 2024. Do you trust this speaker? the impact of prompting on middle-school students' consideration of source when watching conflicting videos. Instructional Science 52, 41–69.

Li, H.O.Y., Bailey, A., Huynh, D., Chan, J., 2020. Youtube as a source of information on covid-19: a pandemic of misinformation? BMJ global health 5.

Little, T.D., 2024. Longitudinal structural equation modeling. Guilford Publications.

Liu, H., Li, J., Wu, Y., Fu, Y., 2019. Clustering with outlier removal. IEEE transactions on knowledge and data engineering 33, 2369–2379.

Livingstone, S.M., Blum-Ross, A., 2020. Parenting for a Digital Future: How Hopes and Fears about Technology Shape Children's Lives. Oxford University Press, Oxford, UK. doi:10.1093/oso/9780190874704.001.0001.

Mayer, R.E., 2002. Multimedia learning, in: Psychology of Learning and Motivation. Elsevier. volume 41, pp. 85–139. doi:10.1016/S0079-7421(02)80005-1.

Mazac, R., Sahlin, K.R., Hyypiä, I., Keränen, F., Niva, M., Berglund, N., Herzon, I., 2025. Does "better" mean "less"? sustainable meat consumption in the context of natural pasture-raised beef. Agriculture and Human Values 42, 1637–1651.

McGrew, S., Breakstone, J., Ortega, T., Smith, M., Wineburg, S., 2018. Can students evaluate online sources? learning from assessments of civic online reasoning. Theory & Research in Social Education 46, 165–193. doi:10.1080/00933104.2017.1413990.

McPhee, R.D., 2016. A virtue epistemic approach to critical thinking. Ph.D. thesis. Bond University. URL: https://research.bond.edu.au/en/studentTheses/a-virtue-epistemic-approach-to-critical-thinking.

Merkt, M., Ballmann, A., Felfeli, J., Schwan, S., 2018. Pauses in educational videos: Testing the transience explanation against the structuring explanation. Computers in Human Behavior 89, 399–410. doi:10.1016/j.chb.2018.08.020.

Metzger, M.J., Flanagin, A.J., 2013. Credibility and trust of information in online environments: The use of cognitive heuristics. Journal of Pragmatics 59, 210–220. doi:10.1016/j.pragma.2013.07.007.

Moeller, J., 2015. A word on standardization in longitudinal studies: don't.

Nickerson, R.S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology 2, 175–220.

Noble, W.S., 2009. How does multiple testing correction work? Nature biotechnology 27, 1135–1137.

Nowak-Brzezińska, A., Gaibei, I., 2022. How the outliers influence the quality of clustering? Entropy 24, 917.

Ofcom, 2023. Children and parents: Media use and attitudes report 2023. Technical Report. Ofcom. URL: https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-children/children-and-parents-media-use-and-attitudes-report-2023/. published 29 March 2023.

Paul, J., Macedo-Rouet, M., Rouet, J.F., Stadtler, M., 2017. Why attend to source information when reading online? the perspective of ninth grade students from two different countries. Computers & Education 113, 339–354.

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A.A., Eckles, D., Rand, D.G., 2021. Shifting attention to accuracy can reduce misinformation online. Nature 592, 590–595. doi:10.1038/s41586-021-03334-6.

Pennycook, G., Rand, D.G., 2019. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. Cognition 188, 39–50. doi:10.1016/j.cognition.2019.01.012.

Potocki, A., de Pereyra, G., Ros, C., Macedo-Rouet, M., Stadtler, M., Salmerön, L., Rouet, J.F., 2020. The development of source evaluation skills during adolescence: exploring different levels of source processing and their relationships. Journal for the Study of Education and Development 43, 19–59. doi:10.1080/02103702.2020.1733065.

Rai, P., Singh, S., et al., 2010. A survey of clustering techniques. International Journal of Computer Applications 7, 1–5.

Ryan, C., 2021. Data Science with R for Psychologists and Healthcare Professionals. CRC Press.

Salmerón, L., Sampietro, A., Delgado, P., 2020. Using internet videos to learn about controversies: Evaluation and integration of multiple and multimodal documents by primary school students. Computers & Education 148, 103796.

Schwarz, N., Jalbert, M., Noah, T., Zhang, L., 2021. Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. Consumer Psychology Review 4, 4–25. doi:10.1002/cpr3.1042.

Steinberg, L., 2005. Cognitive and affective development in adolescence. Trends in cognitive sciences 9, 69–74.

Sundar, S.S., Molina, M.D., Cho, E., 2021. Seeing is believing: Is video modality more powerful in spreading fake news via online messaging apps? Journal of Computer-Mediated Communication 26, 301–319. doi:10.1093/jcmc/zmab015.

Tang, Y., Tseng, H., Tang, X., 2022. The impact of information-seeking self-efficacy and online learning self-efficacy on students' performance proficiency. The Journal of Academic Librarianship 48, 102584. doi:10.1016/j.acalib.2022.102584.

Twenge, J.M., Martin, G.N., Spitzberg, B.H., 2019. Trends in us adolescents' media use, 1976–2016: The rise of digital media, the decline of tv, and the (near) demise of print. Psychology of Popular Media Culture 8, 329–342. doi:10.1037/ppm0000203.

Vogels, E.A., Gelles-Watnick, R., Massarat, N., 2022. Teens, Social Media and Technology 2022. Technical Report. Pew Research Center. URL: https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/.

Wineburg, S., McGrew, S., 2019. Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. Teachers College Record 121, 1–40.

Yoon, M., Lee, J., Jo, I.H., 2021. Video learning analytics: Investigating behavioral patterns and learner clusters in video-based online learning. The Internet and Higher Education 50, 100806. doi:10.1016/j.iheduc.2021.100806.

Young, A.G., Shtulman, A., 2020. Children's cognitive reflection predicts conceptual understanding in science and mathematics. Psychological Science 31, 1396–1408. doi:10.1177/0956797620945012.

## Appendix A. Example Appendix Section

Appendix text.