VISUAL REVIEW | Vol. 18, No. 4, 2026 | ISSN 2695-9631
International Visual Culture Review / Revista Internacional de Cultura Visual
https://doi.org/10.62161/revvisual.v18.6276

ARTIFICIAL INTELLIGENCE IN VISUAL LITERACY: Design and Evaluation of a Pedagogical Artefact in Higher Education

JACOB BAÑUELOS-CAPISTRÁN[1]
jcapis@tec.mx
[1] Tecnológico de Monterrey, México

KEYWORDS

Visual literacy
Artificial intelligence
Social semiotics
Image analysis
Hermeneutics
Pedagogical artefact
Visual culture
Higher education

ABSTRACT

This article presents the development and exploratory evaluation of MAVIA (Visual Analysis Model with Artificial Intelligence), a digital artefact based on generative artificial intelligence that operationalises a hermeneutic socio-semiotic model for photographic criticism (Bañuelos, 2023). The study adopts an exploratory mixed-methods approach and was conducted with 125 university students enrolled in visual culture and photography courses. The most frequently selected perspective was the historical–techno-aesthetic (78.4%). Overall, 72% of participants evaluated the tool positively, while 36.8% identified limitations related to overanalysis. The combined mode of use (39.2%) demonstrated greater critical depth than the automatic mode (60.8%). The most frequently observed emergent visual literacy indicators were multiperspectival reading (67.2%) and awareness of technological mediation (52%). The results suggest that MAVIA functions as a form of cognitive scaffolding that expands the analytical repertoire without replacing critical judgement, and that the tension between automation and interpretation constitutes a productive pedagogical space.

Received: 04/ 03 / 2026
Accepted: 13/ 04 / 2026

1. Introduction

The production, circulation and consumption of images in the contemporary digital ecosystem have reached an unprecedented scale. It is estimated that billions of photographs are shared daily on digital platforms, a figure that does not account for the total number of images generated by artificial intelligence systems. However, critical image-reading competences—what the academic tradition refers to as visual literacy (Debes, 1969; Mirzoeff, 2003)—have not progressed at the same pace as this proliferation. Despite being hyper-consumers of images, university students rarely have access to theoretical frameworks or methodological tools that enable them to move beyond the aesthetic surface or the immediate emotional impact of what they see. In disciplines such as communication, design and the visual arts, this gap between visual overproduction and the critical competence required to interpret it has become a major pedagogical challenge.

Generative artificial intelligence tools—particularly large language models (LLMs)—in principle open unprecedented avenues for bridging this gap. Several authors have emphasised the urgency of developing critical competences for the use of these technologies in the classroom (Bozkurt, 2024; Long and Magerko, 2020). However, much of the existing research has focused on the use of AI in writing tasks or information retrieval (Barrett and Pack, 2023); its role as a mediator in the critical reading of images remains largely unexplored. The question that motivates this study is therefore as follows: can an AI system, by articulating complex theoretical frameworks, function as a form of cognitive scaffolding that guides students in the construction of well-grounded visual analyses?

This question is addressed at the intersection of visual studies, social semiotics and artificial intelligence applied to education. The conceptual ambition of this research lies in an interdisciplinary integration that brings together social semiotics—with its tripartite distinction between the semantic, syntactic and pragmatic domains (Morris, 1985; Verón, 1993)—Gadamerian hermeneutics—with its concepts of finitude, prejudice and the fusion of horizons (Gadamer, 1999)—and generative artificial intelligence technologies, with the aim of enriching visual interpretation without replacing human judgement. MAVIA (Visual Analysis Model with Artificial Intelligence) (Bañuelos, 2025) is presented as a digital artefact that translates into an interactive AI system the interdisciplinary hermeneutic socio-semiotic model for photographic criticism proposed by Bañuelos (2023). This model integrates five analytical perspectives—historical–techno-aesthetic, philosophical–sociocultural, media–technological, anthropological and semiotic—articulated within the three domains of social semiotics.

1.1. General Objective and Research Questions

The general objective of this research is to describe the design, implementation and impact of MAVIA on the visual literacy practices of university students, through an exploratory qualitative study supported by descriptive statistics, involving 125 participants enrolled in photography and visual culture courses at a private university in Mexico.

This objective is articulated through two complementary research questions:

RQ1: In what ways does an AI artefact, based on an interdisciplinary hermeneutic socio-semiotic model, contribute to visual literacy practices among university students?

RQ2: What are the possibilities and limitations of this type of tool in the development of competences for the critical reading of images?

The article is structured as follows: first, the theoretical framework underpinning the study is presented, organised around four key axes; next, the methodology of the exploratory study is described; this is followed by the presentation of results organised into five thematic categories; finally, the discussion relates the findings to the theoretical framework, and the conclusions synthesise the contributions, limitations and directions for future research.

2. Theoretical framework

The theoretical framework of this research is structured around four interrelated axes that underpin both the design of the MAVIA artefact and the analysis of its effects on students' visual literacy.

2.1. Visual Literacy in the Digital Age

It was John Debes (1969) who coined the term *visual literacy* to refer to the set of competences that an individual develops by integrating sensory experience with the reading of the visible. Since then, the concept has undergone continuous reformulation. Susan Sontag (1977) showed how photography alters our relationship with reality by imposing a kind of visual "grammar" that conditions perception; W. J. T. Mitchell (1992), for his part, argued that the image in the digital age requires new forms of critical reading. With Nicholas Mirzoeff (2003), the field expanded towards what is now known as visual culture: a domain that extends beyond the aesthetic to interrogate the power relations embedded in contemporary regimes of visuality.

More recently, Karayianni and Leftheriotis (2025) have mapped the convergence between artificial intelligence and visual literacy—a field which, according to these authors, is rapidly expanding and calls for renewed theoretical frameworks—while Brumberger (2025) has argued for the need to redefine the very notion of visual literacy in light of artificial intelligence. Today, this concept encompasses the ability to interpret, produce, evaluate and communicate through images, attending both to their aesthetic and technical dimensions and to their cultural and political implications. A recent study by Sánchez Márquez et al. (2024) reinforces this perspective: in analysing the digital and audiovisual literacy of 320 university students, the authors identified favourable attitudes towards technology alongside notable deficiencies in the critical reading of visual media.

This complexity has been further intensified in recent years by the widespread emergence of images generated by artificial intelligence. Joanna Zylinska (2017) had already anticipated the advent of what she termed *non-human photography*, a phenomenon that challenges conventional notions of authorship and representation. More recently, Cooper and Tang (2024) have documented how images produced by DALL-E 3 reproduce stereotypes in scientific education contexts, while Heaton et al. (2024) explore the still unresolved tensions between the "artificial" and the "intelligent" when AI is integrated into art education. What these studies make clear is that it is no longer sufficient to know how to read images: it is essential to understand how algorithmic systems produce, classify and interpret them (Bozkurt, 2024; Long and Magerko, 2020).

The convergence between visual literacy and AI literacy constitutes an emerging field with limited literature, particularly in the Ibero-American context. Mansoor et al. (2024) report that AI literacy among university students varies significantly across cultural and disciplinary contexts, highlighting the need for situated studies that examine how specific AI tools can contribute to the development of competences in the critical reading of images within concrete university settings.

2.2. The Socio-Semiotic Hermeneutic Model for Photographic Criticism

The central theoretical foundation of MAVIA is the socio-semiotic hermeneutic interdisciplinary model for photographic criticism proposed by Bañuelos (2023). This model constitutes a systematic approach to the analysis of photographic images, integrating multiple theoretical traditions into a coherent operational framework.

In its semiological dimension, the model is grounded in the social semiotics of Charles Morris (1985) and his tripartite distinction between the semantic, syntactic and pragmatic dimensions of the sign, complemented by the social semiosis of Eliseo Verón (1993) and Peircean semiotics (Charles Sanders Peirce, 1974). In its hermeneutic dimension, the model draws on Gadamerian concepts of finitude, opacity, prejudice, tradition and the fusion of horizons (Hans-Georg Gadamer, 1999), assuming that any interpretation of an image is conditioned by the interpreter's historical situation and that no analysis can exhaust the meanings of an image.

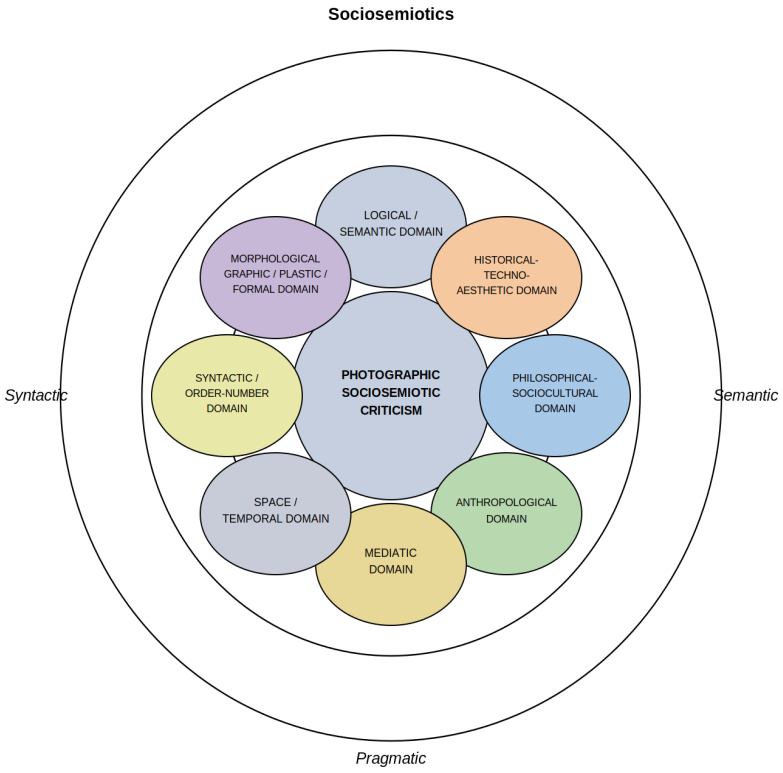
The model articulates five interdisciplinary analytical perspectives, each of which operates simultaneously across the three semiotic domains (see Table 1 and Figure 1). The historical–techno-aesthetic perspective situates the image within its technological context and within the history of photography as an expressive medium (Ritchin, 2009; Wells, 2015). The philosophical–sociocultural perspective examines social imaginaries and regimes of visuality (Mitchell, 1992; Rancière, 2007). The media–technological perspective analyses conditions of circulation and media ecology (Joan Fontcuberta, 2016). The anthropological perspective approaches the image as a vehicle of identity, memory and representation (Hans Belting, 2007). Finally, the semiotic perspective focuses on the iconic, indexical and symbolic signs that compose the image.

Table 1. Structure of the socio-semiotic model: Analytical perspectives and domains

Perspective	D. Semantic	Syntactic domain	Pragmatic domain
Historical-technological-aesthetic	Historical and aesthetic meanings of the photographic medium	Compositional conventions and techniques of the period	Social functions of photography in its historical context
Philosophical-sociocultural	Social imaginaries and image-reality relationships	Regimes of visuality and ideological structures	Effects of truth and the power of the image
Media and technology	Meanings derived from the platform and the device	Formats, resolution and conditions of circulation	Algorithm-mediated reception and media ecology
Anthropological	Representation of the body, identity and memory	Cultural codes of representation	Ritual, identity and memory functions
Semiotics	Iconic, indexical and symbolic signs	Compositional codes and intertextual relationships	Effects of meaning on the recipient

Note. Adapted from Bañuelos (2023). Each cell summarises the analytical focus of the perspective-domain intersection. Source: Author's own elaboration, 2026.

Figure 1. The socio-semiotic hermeneutic model for photographic criticism



Source: Bañuelos (2023, p. 320). Author's own elaboration.

The articulation of five perspectives across three domains generates an analytical matrix that enables the photographic image to be approached from multiple angles without reducing it to a single interpretative dimension. The originality of the model lies precisely in its integrative capacity, which combines semiotic rigour with hermeneutic openness. As shown in Figure 1, the original model proposed by Bañuelos (2023) is structured as a circular diagram in which eight analytical domains are organised around the core of socio-semiotic photographic criticism, framed by the three dimensions of semiosis represented as concentric rings.

For its implementation in MAVIA, several operational adaptations were introduced: the perspectives were grouped into five selectable modules that integrate related domains, and the three semiotic axes were structured as transversal dimensions within the system instructions. This decision entails a partial linearisation of a model that, in its theoretical conception, is fundamentally relational and interdependent. This adaptation responds to the affordances and limitations of the AI's conversational interface and should be understood as an operational interpretation of the theoretical framework rather than an exact replication.

2.3. Artificial Intelligence as a Pedagogical Tool for Visual Culture

Since ChatGPT became publicly accessible in late 2022, academic production on artificial intelligence in education has grown at a rapid pace. Cross-national surveys indicate that more than 90% of university students have already used some form of generative AI tool, although patterns of use vary considerably across disciplines and cultural contexts (Mansoor et al., 2024). For Bozkurt (2024), literacy in generative artificial intelligence is an urgent task that must integrate knowledge, practice and critical reflection. Allen and Kendeou (2024), from an interdisciplinary perspective, have developed the ED-AI Lit framework, which brings together technical, cognitive and ethical competences within a single educational proposal. In the Ibero-American context, Segovia-García (2023) reported that online postgraduate students value chatbots as a learning support, while at the same time expressing doubts about the reliability of the outputs generated by these systems and the risk of becoming dependent on them.

When the focus is narrowed to art and visual education, the research landscape remains relatively underdeveloped. Heaton et al. (2024) document the experience of three art educators in Singapore who incorporated AI into their teaching: the overall assessment was ambivalent, with gains in interdisciplinarity and student engagement, but also with a loss of manual skills and a degree of dependence on the system. Cooper and Tang (2024), for their part, warn of the reproduction of cultural stereotypes in images generated by DALL-E 3 within scientific education contexts—a finding that underscores the need to subject AI-generated visual outputs to critical scrutiny. In the field of academic writing, Barrett and Pack (2023) identified a revealing divergence: while students tend to value generative AI as a support, educators express legitimate concerns as to whether such support may ultimately undermine the autonomous development of critical thinking.

From these studies emerges a tension that might be described as constitutive: artificial intelligence can function as cognitive scaffolding—in the Vygotskian sense (Lev Vygotsky, 1978)—enhancing students' capabilities by providing conceptual frameworks that reorganise their understanding, but it may also replace the very processes of thinking that it is intended to stimulate. In the domain of image interpretation, this tension becomes more acute: critically reading an image requires a subjective positioning, grounded in experience and personal history, which no algorithmic system is capable of replicating.

2.4. From Theoretical Model to Computational Artefact: Translating Social Semiotics

Why translate a socio-semiotic model into an AI artefact? The underlying hypothesis is pedagogical and may be formulated as follows: if the model proposed by Bañuelos (2023) offers a systematic structure for analysing images, and if large language models are capable of articulating complex conceptual frameworks, then an artefact that encodes this model could mediate between social semiotics and students' concrete analytical practice.

Implementing this required a process of computational operationalisation: the perspectives, domains and categories of the model were encoded into system instructions (*system prompts*) that condition the behaviour of the language model. What emerges from this operation—MAVIA (Bañuelos, 2025)—is not a conventional chatbot that simply answers questions, but a system that reproduces the analytical logic of the theoretical framework and generates analyses that mobilise specialised vocabulary from social semiotics, visual studies and hermeneutics.

However, this translation is not without tensions. The model proposed by Bañuelos (2023) is grounded in a hermeneutic tradition according to which all understanding is conditioned by the finitude and historicity of the interpreter (Gadamer, 1999): interpreting an image requires a situated subject, endowed with prejudices—in the Gadamerian sense—traditions and a horizon of understanding of their own. An AI system, by definition, lacks such existential situatedness. This raises the question of whether

this absence undermines its use as a hermeneutic mediator or whether, on the contrary, the friction between the systematicity of the artefact and the student's interpretative openness may become pedagogically productive. It is precisely this tension that the present study seeks to examine.

3. Methodology

3.1. Type of Study and Methodological Approach

The study was designed as exploratory, with a predominantly qualitative orientation supported by descriptive statistics (Creswell and Creswell, 2022). In its qualitative dimension—which underpins the overall design—both thematic analysis of students' reflections and content analysis of the documents generated by MAVIA were employed. Descriptive statistics, in turn, were used to quantify frequencies and distributions of emergent categories, selected perspectives, modes of use and indicators of visual literacy identified in the corpus.

Two reasons justify the exploratory nature of the study: the novelty of the artefact and the near absence of prior research combining socio-semiotic models with AI tools in visual education. This approach makes it possible to identify unforeseen categories, unexpected patterns of use and tensions that emerge only through direct engagement with the data, thereby laying the groundwork for future research of greater scope. As for the decision to combine qualitative and descriptive approaches, it responds to a dual need: to gain an in-depth understanding of students' experiences and, at the same time, to assess the extent to which these patterns were present among the 125 participants.

3.2. Description of the MAVIA Artefact

MAVIA was developed on Claude AI, the platform developed by Anthropic (Sonnet 4.5) (Bañuelos, 2025), using the Projects functionality, which enables the configuration of specialised system instructions (*system prompts*). The artefact encodes the entirety of Bañuelos's (2023) **socio-semiotic,** hermeneutic model—its five analytical perspectives, three domains and subcategories of analysis—into a set of instructions that guide the language model in generating image analyses.

The interaction flow follows a defined sequence: (a) the student uploads a photographic image; (b) selects the analytical perspectives to be applied (one or more of the five available); (c) the artefact generates a structured analysis in accordance with the theoretical model; (d) the student may edit, supplement or challenge the generated analysis; (e) the system produces a final document that integrates the selected analyses. The link to the artefact is provided in Bañuelos (2025).

MAVIA offers three modes of use: automatic (the system generates the complete analysis without student intervention), manual (the student produces the analysis using the model's structure as a guide) and combined (the system generates an initial analysis that the student modifies, supplements or corrects). The latter mode was promoted as the recommended practice during the course.

3.3. Corpus and Participants

The corpus consists of the work produced by 125 university students enrolled in Communication, Design and Arts programmes at a private university in Mexico, who were taking courses in photography, documentary production and visual culture during the August–December 2025 semester. The materials analysed include: (a) social semiotic analysis documents generated using MAVIA based on photographs either produced by the students or sourced by them; and (b) written comments and reflections by students on their experience of using the artefact.

Students used MAVIA as part of their regular course activities, applying the model to their own photographs or to images not produced by them—portraits, urban landscapes, documentary, product and personal photography—with the freedom to select both the analytical perspectives and the mode of use. To preserve anonymity, participants are identified by alphanumeric codes (E-XX), where E indicates student and the letters and numbers correspond to a random assignment unrelated to personal data.

3.4. Analysis Techniques

The analysis was conducted on two complementary levels. At the qualitative level, thematic analysis (Braun and Clarke, 2006) was employed to examine students' reflections, complemented by content analysis of the documents generated by MAVIA. The coding process followed two cycles: an initial cycle

of descriptive coding, aimed at identifying recurring themes in the reflections and patterns in the generated analyses, and a second cycle of pattern coding, aimed at constructing analytical categories that articulated the findings with the theoretical framework. Both a priori categories (derived from the theoretical model: perspectives used, referential density, specialised vocabulary) and emergent categories (derived from the corpus: perceptions of overanalysis, contextual errors, metacognition) were developed.

At the descriptive quantitative level, once the thematic categories had been established through the two coding cycles, absolute frequency counts were carried out by identifying the presence or absence of each category in each participant's reflections. Coding was conducted by the principal investigator, with cross-checking of a 20% sample of the corpus by a second coder, achieving an agreement rate of 85%. Frequencies were calculated based on the total number of participants (n = 125), and since each student could refer to more than one category in their reflections and select multiple perspectives, the frequencies are not mutually exclusive. The descriptive data are not intended to establish causal relationships or statistical generalisations, but rather to quantify qualitative patterns and facilitate the visualisation of trends observed in the corpus (Creswell and Creswell, 2022). The findings were triangulated across the generated analyses, students' reflections, descriptive data and the theoretical framework.

3.5. Ethical Considerations and Researcher Positionality

The work analysed was produced as part of regular academic activities within the courses. The data were anonymised for analysis by removing names and identifying information. No sensitive personal data were collected beyond the academic work produced. The institutional guidelines for ethics in educational research were followed.

It is necessary to make explicit that the researcher responsible for this study occupied a dual position throughout the research process: as the designer of the artefact under analysis and as the instructor of the courses in which it was implemented. This dual role introduces a potential bias that must be acknowledged. To mitigate this, the following strategies were adopted: (a) students' reflections were produced as part of regular course activities, without prior knowledge that they would later be analysed for research purposes; (b) the coding process was subjected to cross-checking by a second independent coder; (c) emergent categories—particularly those critical of the artefact—were included with the same prominence as positive evaluations; and (d) the discussion explicitly acknowledges the artefact's limitations and tensions. Readers should take this positionality into account when evaluating the findings.

4. Results

The results of the analysis of the corpus comprising 125 student responses are organised into five thematic categories that emerged from the coding process. The tables and figures accompanying the qualitative analysis present the frequencies and distributions that quantify the identified patterns.

4.1. Patterns in the Analyses Generated by MAVIA

The documents produced using MAVIA exhibit a significantly higher level of theoretical density than that typically achieved by students in image analysis tasks conducted without mediation by the artefact. The generated analyses consistently mobilise theoretical references drawn from visual studies (Mitchell, Mirzoeff, Rancière), philosophy of the image (Fontcuberta, Sontag), semiotics (Peirce, Verón) and critical theory (Benjamin), constructing readings of images that articulate multiple layers of meaning. This referential density constitutes a distinctive feature of the artefact: by encoding Bañuelos's (2023) model within its system instructions, MAVIA operates as a mediator that translates specialised theoretical frameworks into analyses applied to specific images.

As shown in Table 2 and Figure 2, the semiotic perspective was the most frequently selected by students (f = 98; 78.4%), followed by the historical–techno-aesthetic (f = 76; 60.8%) and the philosophical–sociocultural (f = 71; 56.8%). The anthropological (f = 45; 36.0%) and media–technological (f = 38; 30.4%) perspectives were less frequently used. This distribution suggests that students tend to favour perspectives that are more intuitively connected to image reading—namely, the

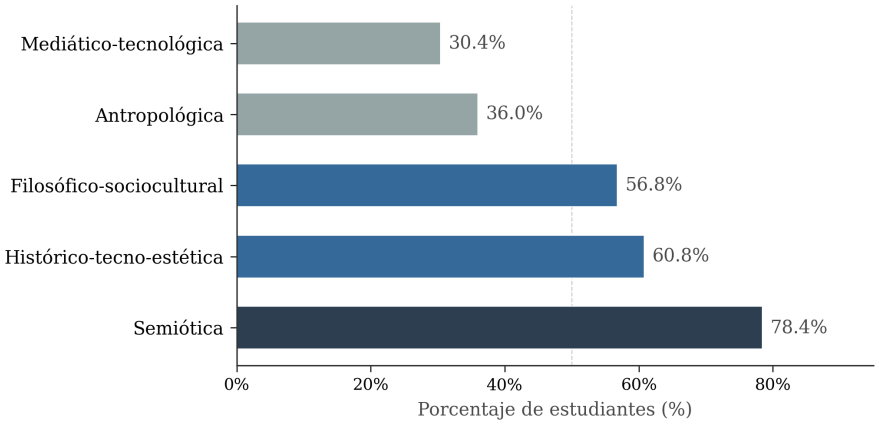
identification of signs and historical contextualisation—whereas perspectives requiring less familiar conceptual frameworks—such as media ecology or the anthropology of the image—demand additional pedagogical mediation. As one student observed: "I liked that it offered different analytical perspectives, depending on the focus of what one aims to identify" (E-M23). Analyses that combined three or more perspectives produced texts of greater interpretative complexity, although they were also those most frequently identified as prone to overanalysis.

Table 2. Frequency of selection of analytical perspectives (n = 125)

Analytical perspective	f	%
Semiotic	98	78.4
Historical-techno-aesthetic	76	60.8
Philosophical-sociocultural	71	56.8
Anthropological	45	36.0
Media-technological	38	30.4

Note. Students could select more than one perspective per analysis. Percentages are calculated based on the total number of participants (n = 125). f = absolute frequency. Source: Author's own elaboration, 2026.

Figure 2. Percentage distribution of analytical perspective selection in the socio-semiotic model.



Source: Author's own elaboration, 2026.

4.2. Perceptions of Usefulness and Expansion of the Analytical Repertoire

A significant proportion of students (f = 89; 71.2%) reported that MAVIA enabled them to access layers of meaning they had not previously perceived in their own photographs. Expressions such as "it allowed me to see the different meanings contained within the image" or "it is a very useful tool for noticing aspects within the image that you had not seen before" reflect a recurring pattern: the artefact functions as an amplifier of the analytical repertoire, expanding the vocabulary and categories through which the student approaches the image. This finding is consistent with the Vygotskian notion of the zone of proximal development: the artefact enables students to operate with conceptual frameworks that exceed their current individual competence.

Students particularly valued three contributions of the artefact (see Table 3): the introduction of specialised theoretical vocabulary (f = 74; 59.2%), including terms such as "regime of visuality", "visual intertextuality" and "social semiosis"; the capacity to establish intertextual connections (f = 62; 49.6%) between the analysed photograph and broader artistic or theoretical traditions; and the opening up of multiple reading perspectives that challenge one-dimensional interpretations. One student noted: "it provides a great deal of information from different fields; I find it interesting how an AI can analyse an image and bring out so many themes" (E-D15). Another wrote: "it allowed me to notice [meanings] that I had overlooked. Overall, I found the report quite accurate, logical and coherent, although I do not think it is advisable to rely blindly on artificial intelligence" (E-M23). This dual perception—recognition of the artefact's contribution alongside reservations about dependence—demonstrates how the artefact both defamiliarises perception and fosters an evaluative stance (Selwyn, 2024).

4.3. Critical Perceptions: Overanalysis and Contextual Errors

Alongside positive evaluations, a considerable number of students identified significant limitations in the analyses generated by MAVIA, as summarised in Table 3 and Figure 3. The most frequent criticism was the identification of overanalysis (f = 67; 53.6%): analyses perceived as excessively elaborate and disproportionate to the actual content of the image. One student described the analysis as "too extensive, becoming illogical and turning into unnecessary overanalysis. It is useful for identifying certain aspects, but it is not entirely accurate; therefore, it should be used responsibly and with awareness" (E-P08). Another student observed: "I found it very interesting, but also somewhat exaggerated for such an everyday photograph. At times, it seems that the AI exaggerates the results, yet at the same time it opens your eyes to things you would not normally question" (E-L42). This coexistence of recognition and critique is particularly revealing.

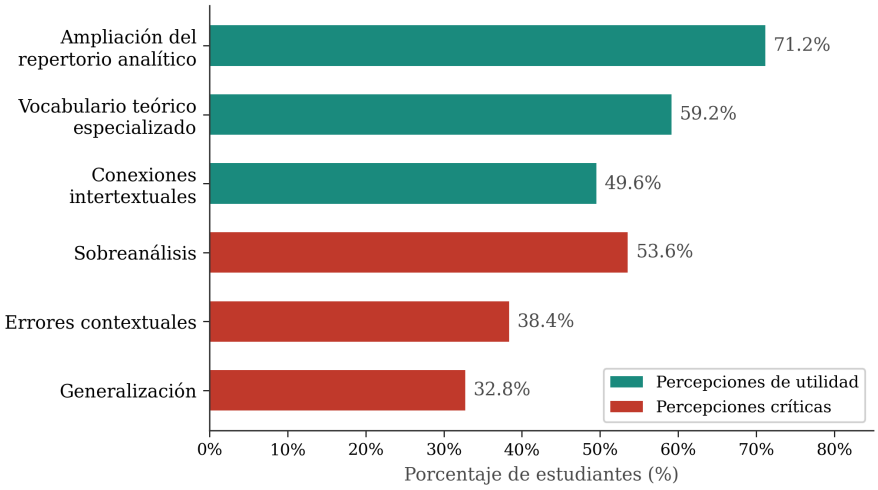
The second most frequent criticism was the identification of contextual errors (f = 48; 38.4%): instances in which the generated analysis did not correspond to the actual context in which the image was produced. One student noted that "in specific situations, it is more difficult for the AI to produce an accurate analysis, as it begins to distort or introduce elements that are not present" (E-V31), particularly in photographs involving local cultural references or personal contexts that the system could not access. A tendency towards generalisation was also identified (f = 41; 32.8%): analyses that could be applied to almost any similar image without attending to the specific characteristics of the photograph in question. These limitations are consistent with what Selwyn (2024) refers to as the "fundamental limits" of AI in education and align with the structural absence of a situated hermeneutic horizon that Gadamer (1999) identifies as a condition of all genuine interpretation.

Table 3. Emerging thematic categories: Students´ perceptions of MAVIA (n = 125)

Category	f	%	*Representative statement*
Perceptions of usefulness			
Expansion of the analytical repertoire	89	71.2	*'It allowed me to see the different meanings contained within the image'*
Specialised theoretical vocabulary	74	59.2	*"I learned terms such as regime of visuality and social semiosis"*
Intertextual connections	62	49.6	*"It connected my photograph to artistic traditions I was not familiar with"*
Critical perceptions			
Overanalysis	67	53.6	*"Too extensive, becoming illogical"*
Contextual errors	48	38.4	*"The AI did not seem to understand the image"*
Generalisation	41	32.8	*'The analysis could be applied to any similar photograph'*

Note: f = Number of students who mentioned the category in their reflections. A single student could refer to more than one category. Representative expressions are anonymised verbatim quotations.
Source: Author's own elaboration, 2026.

Figure 3. Distribution of perceptions of usefulness and critical perceptions regarding MAVIA



Source: Author's own elaboration, 2026.

4.4. Modes of Use: Automatic Versus Combined

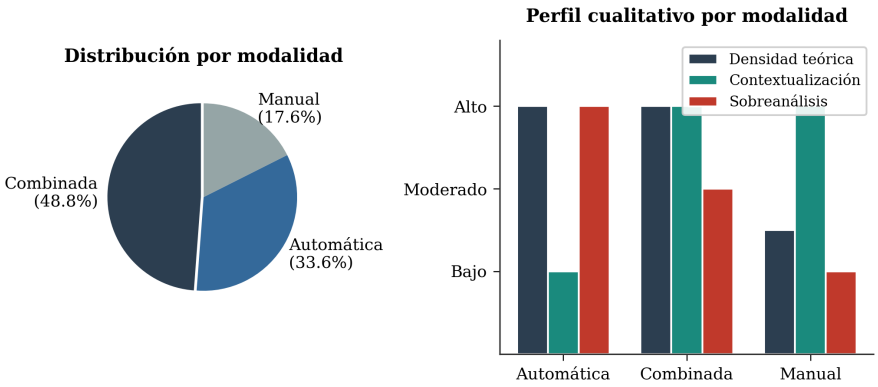
The comparison between the analyses produced across the three modes of use reveals significant qualitative differences (see Table 4 and Figure 4). The combined mode was the most frequently adopted by students (n = 61; 48.8%), followed by the automatic mode (n = 42; 33.6%) and the manual mode (n = 22; 17.6%). Combined analyses tend to be more situated: they incorporate contextual information known only to the student-photographer, correct erroneous interpretations produced by the system, and add nuances derived from authorial intent and the circumstances of the image's production. In contrast, purely automatic analyses, although more theoretically dense, more frequently exhibit the problems of overanalysis and decontextualisation identified above.

Table 4. Comparison of MAVIA modes of use (n = 125)

Modality	n	%	Theoretical density	Contextualisation	Reported over-analysis
Automatic	42	33.6	High	Low	Frequent
Combined	61	48.8	High	High	Moderate
Manual	22	17.6	Low-Medium	High	Low

Note: n = number of students who predominantly used each mode. Theoretical density, contextualisation and reported overanalysis are qualitative assessments derived from the content analysis of the corpus.
Source: Author's own elaboration, 2026.

Figure 4. Distribution and qualitative profile of MAVIA modes of use



Source: Author's own elaboration, 2026.

This distribution confirms that the combined mode constitutes the optimal use of the artefact: the system provides the theoretical framework and specialised vocabulary, while the student contributes situated knowledge, authorial intent and critical judgement regarding the relevance of the analysis. It is noteworthy that the combined mode simultaneously exhibits the highest level of theoretical density and the greatest degree of contextualisation, indicating that student intervention does not dilute the complexity of the analysis but rather anchors it in the specificity of the image. Student intervention is therefore not merely corrective; it is constitutive of the interpretative process, in line with the hermeneutic premise that the meaning of an image emerges in the encounter between the visual text and the interpreter's horizon.

4.5. Indicators of Emergent Visual Literacy

The analysis of the corpus made it possible to identify four indicators of visual literacy that emerged through the use of MAVIA (see Table 5 and Figure 5). The most frequent indicator was the incorporation of specialised terminology (f = 81; 64.8%): students began to use in their own reflections terms they had encountered through the artefact, such as "scopic regime", "punctum", "indexical sign" or "post-photography". This indicator demonstrates a lexical transfer from the artefact to the student's discourse, a necessary—though not sufficient—condition for conceptual appropriation.

The second most frequent indicator was critical positioning in relation to AI (f = 67; 53.6%): students who critically evaluated the quality and relevance of the generated analysis, identifying errors, overanalysis or disconnections from the specific image. The articulation of multiple perspectives (f = 53;

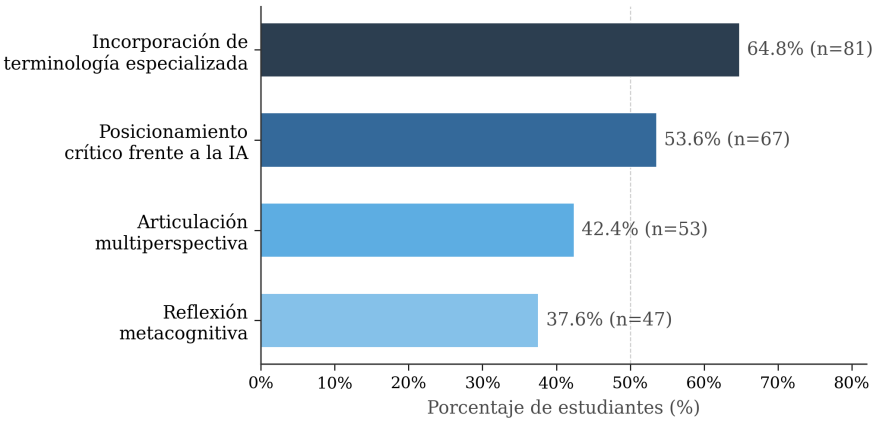
42.4%)—students who developed the ability to approach the image from complementary angles without reducing it to a unidimensional reading—and metacognitive reflection (f = 47; 37.6%) complete the set of indicators. The latter is exemplified in reflections such as: "I usually look at it without thinking too much, but when I observed it more carefully, I realised that every pose, gesture and object communicates something. It changed the way I perceived it because I noticed that even the simplest elements can have a deeper meaning" (E-A67), or "it is very important to learn how to use AI in a way that genuinely benefits society, and this is a great way to learn how to use AI tools for our benefit and to learn more, rather than making us more passive and dependent" (E-R12). It is noteworthy that more than half of the participants demonstrated critical positioning in relation to AI, suggesting that the artefact, paradoxically, stimulates students' evaluative capacity rather than diminishing it—a finding consistent with the notion of critical AI literacy that Karayianni and Leftheriotis (2025) identify as a central emerging competence in the field.

Table 5. Indicators of emergent visual literacy identified in the corpus (n = 125)

Indicator	f	%	Example
Incorporation of specialist terminology	81	64.8	*Use of terms such as punctum, scopic regime, and indexical sign in students' own reflections*
Multiperspectival articulation	53	42.4	*Analyses that integrate ≥2 perspectives in an articulated manner*
Metacognitive reflection	47	37.6	*"It changed my perception of my own photograph"*
Critical positioning in relation to AI	67	53.6	*Identification of errors, overanalysis or generalisations*

Note: f = number of students whose reflections or combined analyses evidenced the indicator. A single student could exhibit more than one indicator.
Source: Author's own elaboration, 2026.

Figure 5. Frequency of emergent visual literacy indicators



Source: Author's own elaboration, 2026.

5. Discussion

The results presented make it possible to articulate three lines of discussion that connect the empirical findings with the theoretical framework of the study and with the literature on AI in education.

5.1. MAVIA as Cognitive Scaffolding for Visual Literacy

The qualitative and descriptive data suggest that MAVIA effectively operates as cognitive scaffolding in the Vygotskian sense (Vygotsky, 1978): it provides a theoretical structure that enables students to carry out analyses of a level of complexity they would not achieve autonomously at this stage of their education. The fact that 71.2% of participants reported an expansion of their analytical repertoire and that 64.8% incorporated specialised terminology into their own reflections supports this interpretation.

The artefact does not teach seeing in a strict sense; rather, it provides conceptual frameworks and specialised vocabulary that reorganise perception. Crucially, students do not merely reproduce the generated analyses; they evaluate them, question them and, in the combined mode, actively transform them.

However, it is necessary to distinguish between two functions of scaffolding that the data make it possible to differentiate. On the one hand, MAVIA functions as informational scaffolding: it introduces references, concepts and theoretical traditions that expand the student's knowledge base. This function is evidenced by the high frequency of incorporation of specialised vocabulary and by the intertextual connections fostered by the artefact. On the other hand, it functions as procedural scaffolding: it models a multiperspectival analytical process that students may eventually internalise and transfer to other contexts of image interpretation. The data suggest that the first function is fulfilled consistently, whereas the second—visible in the 42.4% of students who articulated multiple perspectives autonomously—requires additional pedagogical mediation and, most likely, sustained use of the artefact beyond the scope of this study.

5.2. The Paradox of Overanalysis: Critique as Evidence of Literacy

One of the most significant findings of the study is the paradoxical nature of students' criticisms of the artefact. When a student detects that the AI has produced overanalysis—that it has identified meanings where none exist, that it has generated an analysis disproportionate to or disconnected from the specific image—they are exercising precisely the critical reading capacity that the artefact seeks to promote. The fact that 53.6% of participants demonstrated critical positioning in relation to AI should not be interpreted as an indicator of the artefact's failure, but rather as evidence of its effectiveness as a catalyst for evaluative thinking. In other words, a student who recognises that "the AI did not understand the image" is demonstrating that they did understand it, or at least that they possess well-founded criteria for assessing the relevance of a visual analysis.

This paradox can be productively interpreted through a Gadamerian hermeneutic lens. Gadamer (1999) argued that prejudice—understood not in a pejorative sense but as pre-understanding, as a prior horizon of meaning—is a condition of possibility for all interpretation. A student who photographs a scene understands it from a particular horizon that includes intention, context and lived experience. AI, lacking this situated horizon, produces analyses that apply theoretical frameworks systematically but at times indiscriminately. The student's recognition of this discrepancy constitutes an instance of the fusion of horizons: the student confronts their situated understanding with the artefact's systematised reading, and in that confrontation a critical judgement emerges that did not exist prior to the interaction. Perceptions of overanalysis and contextual errors, read from this perspective, are not system failures but pedagogically productive moments that activate hermeneutic reflection.

This interpretation is reinforced by the data on modes of use: the combined mode, which requires students to intervene actively in the generated analysis, was both the most widely adopted (48.8%) and the one that best balances theoretical density and contextualisation. The pedagogical value of MAVIA lies not only in the quality of the analyses it generates, but also—and perhaps more fundamentally—in the critiques it provokes and the interventions it demands.

5.3. Implications for the Pedagogy of Visual Culture

The findings of this study have direct implications for the design of pedagogical experiences that integrate AI tools into university-level visual education. First, the results suggest that the use of artefacts such as MAVIA is more productive when embedded within a teaching sequence that includes moments of critical engagement with the generated analyses, not as a final product but as a starting point for reflection. The qualitative superiority of the combined mode reinforces this recommendation: the pedagogical architecture should be designed in such a way that the student is never a passive recipient of automated analysis, but rather an active interlocutor who evaluates, corrects and supplements it.

Second, the study highlights the importance of making explicit to students the fundamental limitations of AI systems in the interpretation of images: the absence of lived experience, the inability to access the context of production, and the tendency towards generalisation. This explicitness does not weaken the tool; it strengthens it, as it transforms the student from a passive user into a critical evaluator. As Källström (2025) and Selwyn (2024) argue, AI literacy requires not only knowing how to

use tools but also understanding how generative systems mediate knowledge, shape perception and influence the production of meaning. The data from this study support this position: the detection of errors and overanalysis constitutes, in itself, a practice of dual literacy—visual and digital—of considerable educational value, aligned with what Allen and Kendeou (2024) describe as the critical–cognitive dimension of AI literacy.

Third, the experience with MAVIA suggests that AI can play a valuable role in democratising access to theoretical frameworks that would otherwise require extensive disciplinary training. Students who had not read Rancière or Fontcuberta were able, through the artefact, to experience how these frameworks illuminate the reading of an image. However, it is necessary to distinguish between initial exposure and genuine appropriation. If engagement with theoretical frameworks through MAVIA leads to sustained interest in deeper theoretical study, the artefact will have fulfilled a valuable introductory function. If, on the contrary, the student becomes satisfied with automated analysis without developing their own competences, the artefact will have replaced learning. The difference between these two scenarios ultimately depends on the pedagogical design accompanying the use of the tool and the teaching sequence in which it is embedded.

6. Conclusions

This study has examined the design, implementation and exploratory evaluation of MAVIA, an artificial intelligence artefact that translates a socio-semiotic, hermeneutic and interdisciplinary model (Bañuelos, 2023) into an interactive tool for the visual literacy of university students. The findings of the study, grounded in qualitative analysis supported by descriptive statistics on a corpus of 125 participants, allow the following conclusions to be drawn.

In response to the first research question (RQ1), the results show that MAVIA contributes to visual literacy practices at three interconnected levels. First, it is both feasible and productive to translate a complex theoretical model such as that of Bañuelos (2023) into a functional AI artefact. By encoding the five perspectives and the three domains into system instructions, MAVIA generates analyses that effectively mobilise the vocabulary and categories of the socio-semiotic framework; this indicates that large language models can serve as vehicles for specialised conceptual frameworks within visual studies. Second, the artefact functions as cognitive scaffolding: it expands students' analytical repertoire by introducing theoretical vocabulary (64.8%), intertextual connections (49.6%) and multiple perspectives (42.4%) that tangibly enrich their reading of images. This effect is consistent with both Vygotskian scaffolding theory and previous findings on the role of AI in learning (Barrett and Pack, 2023; Heaton et al., 2024). Third, the findings show that the critical use of the artefact—detecting errors, questioning overanalysis and identifying disconnections between the analysis and the specific image—constitutes in itself a practice of visual literacy. A total of 53.6% of participants demonstrated critical positioning in relation to AI, which, when interpreted through a Gadamerian hermeneutic lens, indicates that the artefact activates processes of fusion of horizons between systematised reading and the student's situated understanding.

In response to the second research question (RQ2), the study reveals that MAVIA's potential lies in its capacity to democratise access to complex theoretical frameworks and in the pedagogically productive tension it generates between automation and interpretation. Its main limitations—overanalysis, contextual errors and the tendency towards generalisation—derive from the absence of a situated hermeneutic horizon, making active student intervention (combined mode) a necessary condition for effective educational use.

Beyond these specific findings, this study offers implications that extend beyond the case examined and contribute to the broader field of visual literacy in global digital contexts. First, MAVIA demonstrates that artificial intelligence can function as a vector for the democratisation of theoretical frameworks that have historically remained confined to specialised academic circles: social semiotics, Gadamerian hermeneutics and critical visual studies are no longer the exclusive domain of those with extensive disciplinary training, but become accessible—at least as an initial operational encounter—to students from diverse educational backgrounds. This democratisation does not replace direct engagement with primary sources or rigorous theoretical training, but it can serve as an entry point that stimulates intellectual curiosity and motivates deeper exploration.

Second, the study contributes to addressing one of the most pressing gaps in the contemporary visual ecosystem: the growing distance between the overproduction of images—now further amplified by

algorithmic generation—and the interpretative competences required to read them critically. Artefacts such as MAVIA, by mediating between theory and analytical practice, offer a scalable pedagogical model that can be adapted to different cultural, linguistic and disciplinary contexts, thereby helping to mitigate this gap within higher education. Third, the interdisciplinary integration underpinning the artefact—social semiotics, hermeneutics and generative technology—positions this type of initiative as a reference for interdisciplinary pedagogies in visual culture, where AI does not function as a substitute for critical thinking but as a catalyst that both enhances students' interpretative capacities and tests their judgement. Fourth, this study contributes to the emerging field of AI in visual culture education by providing an interdisciplinary framework that integrates social semiotics, hermeneutics and technology, offering a replicable model for translating complex theoretical frameworks into AI-based pedagogical artefacts.

The study presents several limitations that must be acknowledged: its exploratory nature precludes generalisation; the absence of a control group does not allow causal attribution of the observed effects to the use of the artefact; descriptive data complement but do not replace a rigorous quantitative design; reliance on a commercial platform (Claude AI) raises issues of sustainability and replicability; and the potential for self-perception bias in students' reflections limits the depth of the analysis.

Future lines of research include: comparative pre-/post-use studies of the artefact employing validated visual literacy instruments and quasi-experimental designs; the development of versions of the artefact on other AI platforms to assess the transferability of the model; research on the transfer of competences acquired through MAVIA to other contexts of image interpretation; longitudinal studies examining the evolution of competences over time; and comparative intercultural studies to evaluate how cultural context shapes interaction with the artefact.

Declaration of the Use of Artificial Intelligence

During the preparation of this manuscript, Claude AI (Anthropic, Opus 4.6) was used as an assistance tool for linguistic revision and text editing. The author assumes full responsibility for the content, analysis, interpretations and conclusions presented in this work.

References

Allen, L. K. y Kendeou, P. (2024). ED-AI Lit: An interdisciplinary framework for AI literacy in education. Policy Insights from the Behavioral and Brain Sciences, 11(1), 3–10. https://doi.org/10.1177/23727322231220339

Bañuelos, J. (2023). Hermeneutical and interdisciplinary socio-semiotic model for a photographic critique. Case analysis: Max Siedentopf, How to survive a deadly global virus, 2020. Photographies, 16(3), 316–335. https://doi.org/10.1080/17540763.2023.2185661

Bañuelos, J. (2025). MAVIA: Modelo de Análisis Visual con Inteligencia Artificial [Artefacto digital en Claude AI]. Anthropic. https://claude.ai/public/artifacts/75b784e5-d486-4261-8cde-5f1682d2154d

Barrett, A. y Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. International Journal of Educational Technology in Higher Education, 20, 59. https://doi.org/10.1186/s41239-023-00427-0

Belting, H. (2007). Antropología de la imagen. Katz.

Bozkurt, A. (2024). Why generative AI literacy, why now and why it matters in the educational landscape? Kings, queens and GenAI dragons. Open Praxis, 16(3), 283–290. https://doi.org/10.55982/openpraxis.16.3.739

Braun, V. y Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brumberger, E. (2025). (Re)Defining visual literacy for the age of artificial intelligence. Journal of Visual Literacy, 44(4), 1–17. https://doi.org/10.1080/1051144X.2025.2566593

Cooper, G. y Tang, K.-S. (2024). Pixels and pedagogy: Examining science education imagery by generative artificial intelligence. Journal of Science Education and Technology, 33, 556–568. https://doi.org/10.1007/s10956-024-10104-0

Creswell, J. W. y Creswell, J. D. (2022). Research design: Qualitative, quantitative, and mixed methods approaches (6.ª ed.). Sage.

Debes, J. (1969). The loom of visual literacy: An overview. Audiovisual Instruction, 14(8), 25–27.

Fontcuberta, J. (2016). La furia de las imágenes. Galaxia Gutenberg.

Gadamer, H.-G. (1999). Verdad y método I. Sígueme.

Heaton, R., Low, J. H. y Chen, V. (2024). AI art education—artificial or intelligent? Transformative pedagogic reflections from three art educators in Singapore. Pedagogies: An International Journal, 19(4), 647–659. https://doi.org/10.1080/1554480X.2024.2395260

Källström, K. (2025). GIFT-AI: Damn!.jpg—visual literacy through image-making with generative AI. Frontiers in Education, 10, 1621207. https://doi.org/10.3389/feduc.2025.1621207

Karayianni, I. y Leftheriotis, K. (2025). Exploring the intersection of AI and visual literacy: Current trends. Journal of Visual Literacy, 44(4). https://doi.org/10.1080/1051144X.2025.2587419

Long, D. y Magerko, B. (2020). What is AI literacy? Competencies and design considerations. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–16. https://doi.org/10.1145/3313831.3376727

Mansoor, H. M. H., Bawazir, A., Alsabri, M. A., Alharbi, A. y Okela, A. H. (2024). Artificial intelligence literacy among university students—A comparative transnational survey. Frontiers in Communication, 9, 1478476. https://doi.org/10.3389/fcomm.2024.1478476

Mirzoeff, N. (2003). Una introducción a la cultura visual. Paidós.

Mitchell, W. J. (1992). The reconfigured eye: Visual truth in the post-photographic era. MIT Press.

Morris, C. W. (1985). Fundamentos de la teoría de los signos. Paidós.

Peirce, C. S. (1974). La ciencia de la semiótica. Nueva Visión.

Rancière, J. (2007). The future of the image. Verso.

Ritchin, F. (2009). After photography. W. W. Norton.

Sánchez Márquez, W., Peña Maldonado, A. A., Cervantes López, M. J. y Cruz Casados, J. (2024). Culture of digital and audiovisual literacy among university students. VISUAL REVIEW. International Visual Culture Review, 16(4), 245–253. https://doi.org/10.62161/revvisual.v16.5315

Segovia-García, N. (2023). Percepción y uso de los chatbots entre estudiantes de posgrado online: Un estudio exploratorio. Revista de Investigación en Educación, 21(3), 335–349. https://revistas.uvigo.es/index.php/reined/article/view/4974

Selwyn, N. (2024). On the limits of artificial intelligence (AI) in education. Nordisk Tidsskrift for Pedagogikk og Kritikk, 10, 3–14. https://doi.org/10.23865/ntpk.v10.6062
Sontag, S. (1977). On photography. Farrar, Straus & Giroux.
Verón, E. (1993). La semiosis social. Fragmentos de una teoría de la discursividad. Gedisa.
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.
Wells, L. (2015). Photography: A critical introduction (5.ª ed.). Routledge.
Zylinska, J. (2017). Nonhuman photography. MIT Press.

The complete list of references (APA 7ª citation style) should appear at the end of the article in Cambria 11, single-spaced, without blank spaces between authors and French indentation. When possible, include the DOI for each article in the bibliography and indicate the URL if you cite an open access paper. It is recommended to shorten URLs in case they take up more than one line. Examples:

Asiri, A., Panday-Shukla, P., Rajeh, H. S., & Yu, Y. (2021). Broadening perspectives on CALL teacher education: From technocentrism to integration. *TESL-EJ, 24*(4).

Cayton, C., Sherman, M., Walkington, C., & Funsch, A. (2018). Technology integration in secondary mathematics textbooks. *Conference Papers -- Psychology of Mathematics & Education of North America*, 1235–1238.

Delamont, S. (Ed.). (2013). *Handbook of qualitative research in education.* Cardiff University.

Deliveli, K. (2021). Application of sentence-based sound teaching method for first reading and writing in education. *Participatory Educational Research, 8*(2), 330–356.

Films Media Group. (2013). *Germany's education system: Dan Rather reports on global education systems*. *Films On Demand*. https://fod.infobase.com/PortalPlaylists.aspx?wID=97665&xtid=114473.

Gillen, J. (2021). Learning to connect: relationships, race, and teacher education. *Radical Teacher, 119*, 85–87. https://doi.org/10.5195/rt.2021.902

Hanna, J. M. (2015). *The quality of education leadership doctoral dissertations in the United States: An empirical review* (Order No. 3742189). Education Collection; Education Database. (1752116829). https://ezproxy.roosevelt.edu:2048/login?url=https://www.proquest.com/dissertations-theses/quality-education-leadership-doctoral/docview/1752116829/se-2?accountid=28518

Hobbs, T. D. (2021, May 13). Cheating at school is easier than ever—and It's rampant. *Wall Street Journal (Online)*. https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=150290106&authtype=sso&custid=cls58&site=ehost-live&scope=site.

Jefferson, B. (2020). Education and school reform. In A. C. Washington & B. D. Brady (Eds.), *The comprehensive history of educational practice* (pp. 258-325). Highcourt Publishing.

Ladson-Billings, G. (2018). The social funding of race: The role of schooling. *Peabody Journal of Education, 93*(1), 90–105.

Ladson-Billings, G. (2021). Three decades of culturally relevant, responsive, & sustaining pedagogy: What lies ahead? *Educational Forum, 85*(4), 351–354. https://doi.org/10.1080/00131725.2021.1957632

Sarroub, L. K., & Nicholas, C. (2021). *Doing fieldwork at home: The ethnography of education in familiar contexts*. Rowman & Littlefield Publishers.

Smith, C. (2013). *Leadership lessons from the Cherokee nation: Learn from all I observe (Audio Book)*. McGraw-Hill.

Tharp, D. S. (2020). *Doing social justice education: A practitioner's guide for workshops and structured conversations*. Stylus Publishing LLC.

Tidwell, S., Young, D. J., Wahed, J., Doron, T., Bauer, A., McGirr, M., Ludsin, S. A., Moore, C., Miro, M., & Rodgers, G. (2012, May 8). Letters to the editor. *Wall Street Journal - Eastern Edition, 259*(107), A12.

Warren, C. A. (2021). About centering possibility in Black education. school: Questions. In *Teachers College Press*. Teachers College Press.

Wegner, G. (2019, January 4). First Lego league wrap 2018. *Graham Wegner-Open educator*. https://gwegner.edublogs.org