

Same Standards, Different Pathways: Forgotten Voices, Neurodiversity, and Critical AI Literacy in Ancient History Teaching

Isaías Arrayás Morales

Universitat Autònoma de Barcelona
isaias.arrayas@uab.cat
[0000-0003-4079-0134](https://orcid.org/0000-0003-4079-0134)

Carlos Heredia Chimeno

Universitat Autònoma de Barcelona
carlos.heredia@uab.cat
[0000-0003-2866-5883](https://orcid.org/0000-0003-2866-5883)

ABSTRACT — Growing attention to neurodivergence in higher education and the rapid expansion of generative artificial intelligence have made visible a series of problems that extend beyond either diagnosis or technology: how students organise complex academic work, how disciplinary competences are distinguished from the conventions through which they are normally demonstrated, and how plausible reconstructions can be separated from interpretations that are actually supported by evidence. This article develops a transferable framework from a teaching innovation project in Ancient History implemented during the 2025–26 academic year. The project brought together three concerns: the incorporation of historically marginalised actors into teaching, neurodiversity-aware educational design, and the critical use of generative AI. Drawing primarily on two classroom experiences, one in an *Introductory Ancient History* course and another in an advanced course on *War and Imperialism in Antiquity*, within a broader teaching innovation project involving other courses in the Ancient History area, we gradually came to distinguish between disciplinary operations that should remain common and organisational forms that can legitimately vary. From this process we develop a three-pathway model—guided and sequential, systematic and comparative, and synthetic and visual—available to all students without diagnostic assignment and assessed through identical historical criteria. We further argue that critical AI literacy need not entail compulsory personal adoption of generative AI, since students may analyse documented instructor-provided outputs while demonstrating the same historical competence. Five transferable activity designs translate these principles into practical teaching models concerning evidence-bounded reconstruction, simulated historical voices, narratives of war and empire, visual representation, and source verification. The resulting framework does not seek to lower or relativise academic expectations. On the contrary, it depends upon defining more precisely what students must demonstrate in order for us to say that they have learned Ancient History.

KEYWORDS — Ancient History, artificial intelligence, neurodiversity, cognitive justice, inclusive teaching

1. INTRODUCTION: FORGOTTEN VOICES, DIVERSE LEARNERS, AND GENERATIVE NARRATIVES

The rapid incorporation of generative artificial intelligence into higher education has often been discussed primarily as a problem of technological adoption: whether students should use these systems, how their use should be declared, which forms of assistance constitute legitimate academic practice, and whether traditional modes of assessment remain viable when fluent written responses can be produced within seconds. These questions are unavoidable, but our experience in Ancient History suggests that generative AI is also valuable because it makes older pedagogical and historiographical problems unusually visible. Ancient History has always depended upon incomplete and unevenly preserved evidence, and historians have therefore always had to distinguish between what the sources allow us to establish, what can be proposed as a reasonable inference, and what remains beyond demonstration. Generative systems intensify this problem because they are particularly effective at suppressing the visible boundary between those categories: fragmentary documentation can rapidly become a fluent biography, an asymmetrical archive can become an apparently balanced first-person narrative, and uncertain or disputed evidence can be transformed into confident explanation. These risks are further compounded by algorithmic biases, which may reproduce or amplify patterns embedded in training data and model design, privileging particular perspectives, assumptions, and forms of representation while rendering others less visible. At the same time, increasing attention to neurodivergence and accessibility in universities has raised a parallel question about academic work itself. If the disciplinary objective is to contextualise evidence, compare interpretations, and construct an argument, to what extent must every student organise that process in an identical way? The encounter between these two issues—epistemic uncertainty in AI-generated historical narratives and cognitive diversity in the organisation of learning—became one of the central concerns of the teaching innovation project examined here.

The project emerged within an Ancient History teaching environment in which the question of historical visibility already possessed a substantial disciplinary genealogy. Within the Universitat Autònoma de Barcelona, scholarship since the 1980s has repeatedly questioned narratives excessively centred upon political institutions and dominant groups. Alberto Prieto Arciniega's analysis of the persistence of indigenous elements in Baetica challenged models in which Roman conquest was understood as the straightforward replacement of local societies, while the work of Encarna Sanahuja Yll and Prieto Arciniega on the *Bacchanalia* brought women into a broader analysis of social, economic, and political structures (Prieto, 1980, pp. 37–46; Sanahuja & Prieto, 1981, pp. 143–152). Subsequent research by Isafías Arrayás Morales or Joan Oller Guzmán continued to challenge

conventional narratives of Romanisation by examining the transformation of conquered territories while paying particular attention to the persistence and role of pre-Roman communities and structures within the processes generated by Roman conquest and provincial integration (Arrayás, 2005; Oller, 2016). The project extended that concern with visibility into teaching. Women, children, enslaved people, foreigners, subjected populations, and other actors conventionally marginalised within political and military narratives became not simply additional content to be inserted into existing courses, but a means of asking how historical narratives are constructed in the first place. The arrival of generative AI added a revealing new dimension: systems trained upon existing textual and visual cultures could reproduce precisely those hierarchies, stereotypes, and absences while presenting the resulting account as apparently neutral.

The project also developed in response to increasing attention to neurodivergence within higher education. Research on neurodivergent students has documented recurring challenges concerning disclosure, access to support, and the fit between institutional provision and individual needs, while also showing considerable diversity within and across diagnostic categories (Clouder et al., 2020). This diversity makes it difficult to translate diagnostic categories directly into predetermined pedagogical profiles. Yet recognition creates its own pedagogical danger when a diagnosis begins to operate as a ready-made learning profile. Neither neurodivergent nor neurotypical describes a pedagogically homogeneous population, and sharing a diagnostic category does not imply organising work, communicating, or processing information in the same way. The problem therefore became analogous to the historiographical problem at the centre of the project: categories can make previously neglected differences visible, but they can also become mechanisms through which diversity is simplified again. Our response gradually shifted from asking which activity might be appropriate for a particular student profile towards a different question: which disciplinary operations genuinely need to remain identical for every student, and which features of an assignment reflect conventions of organisation rather than the historical competence being evaluated?

This article reconstructs that development from classroom implementation to a transferable pedagogical framework. The broader teaching innovation project involved activities and adaptations across several courses in the Ancient History area during 2025–26, but the analysis developed here focuses primarily on two classroom experiences for which the implementation process and subsequent reflection can be reconstructed in greater detail. It also draws upon previous work on alternative Ancient History learning itineraries (Heredia, 2026), the broader research literature on neurodiversity, cognitive justice, and AI-supported teaching,

and Open Educational Resources produced during an earlier stage of the same project (Center for the Innovation in Ancient Worlds & Heredia, 2026). The first implementation involved introductory students constructing evidence-bounded fictional historical actors in Hammurabi's Babylon and the Roman Republic during the Second Punic War before critically examining AI-generated expansions of those reconstructions. The second applied the lessons of that experience to an advanced course on war and imperialism, where students were offered three organisational pathways while working with identical historical sources, scholarship, and assessment criteria. The two cases also exposed a contradiction in our own initial approach: while we had diversified the organisation of learning, we continued to require all students to participate directly in generative AI. The framework developed below therefore extends the same principle of choice to technological participation. Our argument is not that every aspect of learning should become negotiable, nor that diagnosis is irrelevant, nor that generative AI should be normalised as an obligatory component of higher education. Rather, we argue that inclusive teaching becomes academically defensible when it starts by defining the disciplinary core with unusual precision. Same academic standards do not necessarily require the same organisational pathway, and critical AI literacy does not necessarily require compulsory personal AI adoption.

2. FROM ADAPTATION TO DESIGN: NEURODIVERGENCE AND COGNITIVE JUSTICE

The concept of neurodiversity has become increasingly prominent in discussions of higher education, but its pedagogical use requires some precision. In its broadest sense, neurodiversity refers to variation in human cognitive functioning, whereas neurodivergent currently operates as an umbrella term associated with autism, ADHD, dyslexia, and other conditions and may also function as a personal or collective identity within a heterogeneous neurodiversity movement (Dwyer, 2022, pp. 73–92). The educational relevance of this distinction lies partly in resisting the assumption that diagnostic categories correspond neatly to preferred methods of learning. Research itself requires caution. Le Cunff et al. (2024, pp. 16–20), examining cognitive load in online learning, identified substantial overlap between difficulties reported by neurodivergent and neurotypical participants, even where particular forms of load appeared with different intensity. Howard and Sedgewick (2021, pp. 2265–2278) similarly found substantial variation in communication preferences among autistic adults across different communicative contexts, with written modes often preferred when interacting with unfamiliar people or organisations. Preferences for written communication, additional processing time, or greater task structure may therefore be relevant to some learners, but none can safely be treated as a diagnostic template. A diagnosis can alert the teacher to a

possible barrier; it cannot predict the learning behaviour of every person to whom it applies.

This becomes particularly important in course design. Our own earlier work illustrates the problem. Heredia (2026) proposed two routes under common historical expectations: a structured and linear pathway associated orientatively with autistic and dyslexic learners, and a more open, creative, and multimodal pathway associated with ADHD. The underlying principle—that equivalent disciplinary competence could be demonstrated through different routes—proved useful, but the diagnostic association was less defensible. Subsequent implementation persuaded us to preserve differentiated routes while removing diagnostic assignment. A structured pathway could be chosen by an autistic student, a student with ADHD, a neurotypical student unfamiliar with the topic, or anyone who simply worked better through explicit stages. The reason for the choice ceased to matter for assessment.

The institutional framework at the Universitat Autònoma de Barcelona reinforces rather than contradicts this move towards anticipatory design. UAB regulations allow curricular adjustments to affect teaching and learning methodologies, materials, and assessment while requiring that they do not reduce the academic level demanded (Universitat Autònoma de Barcelona, 2024). Individualised adjustments and specialist support therefore remain indispensable because no general framework can foresee every personal circumstance. This anticipatory logic is also compatible with Universal Design for Learning, which has been proposed as a framework for planning more accessible teaching and learning in higher education (Cumming & Rose, 2022). The framework developed here is narrower, however: it does not seek to reproduce UDL as a whole, but to distinguish disciplinary requirements from organisational features that may legitimately vary. The point is not to replace those mechanisms with a universal design that claims to eliminate the need for accommodations. It is to ask whether some predictable barriers can be reduced before students are required to demonstrate individually why they need an alternative. A complex assignment can contain explicit stages without requiring a formal accommodation; a comparative organiser can be available to everyone; a student can choose written rather than synchronous communication for an intermediate task where real-time participation is not itself a learning outcome. General accessibility and individual adjustment operate at different levels and should be complementary. This distinction also avoids treating neurodiversity as a reason to lower expectations: the question is how students reach and demonstrate a common disciplinary standard, not whether some students should be evaluated against a different intellectual threshold.

The concept of cognitive justice helps us articulate this distinction, although we use it analogically and in a deliberately restricted pedagogical sense. Visvanathan (1997) formulated cognitive justice in relation to inequalities between systems of knowledge and the power of dominant forms of knowledge to marginalise others; de Sousa Santos (2007) developed this concern within a broader critique of epistemic hierarchy. Leibowitz (2017, pp. 93–111) subsequently explored its possible implications for higher-education curricula while also acknowledging that further conceptual development is required to translate cognitive justice into teaching and learning. Our use therefore draws upon this literature rather than claiming equivalence with the historical and epistemic injustices it primarily addresses. In the narrower sense proposed here, the issue arises when a particular convention for producing or demonstrating academic knowledge becomes indistinguishable from the knowledge supposedly being evaluated. Ancient History provides a relatively clear test. The capacity to contextualise a source, distinguish evidence from interpretation, compare different forms of documentation, use scholarly literature, detect anachronism, and sustain an argument through evidence belongs to the disciplinary core. It is much less evident that every student must reach those objectives through exactly the same sequence, the same planning format, or the same relationship with a digital tool. One student may need to divide a research problem into explicitly delimited stages, another may make sense of it by comparison, while another may first establish relationships between evidence through diagrams or visual representation before converting those relationships into prose. None of those organisational differences determines the final academic level by itself.

The principle that emerged from this reflection is therefore more demanding than a generic appeal to flexibility. The first task of inclusive assessment is to identify what cannot vary because it constitutes the disciplinary competence; only then can we decide what may vary without changing the standard. This reverses a common sequence in which an existing activity is treated as the norm and alternatives are subsequently added for particular students. Instead, the activity is decomposed from the beginning into disciplinary constants and organisational variables. The distinction also creates a practical test for design: if changing a requirement would alter what the student is being asked to know or do historically, the requirement is likely to belong to the disciplinary core; if the historical operation remains intact, the requirement may be an inherited convention. Removing primary-source analysis from an assignment would alter its intellectual content. Allowing students to organise that source analysis through a linear worksheet, comparison table, or visual map need not (see Table 1).

Disciplinary constants	Organisational variables
Engagement with primary evidence	Order in which evidence is approached
Historical contextualisation	Planning method and intermediate structure
Distinction between evidence and interpretation	Linear, comparative, or visual organisation
Engagement with academic scholarship	Use of tables, maps, timelines, or diagrams
Identification of anachronism and unsupported reconstruction	Form of notes and working materials
Evidence-based argument	Route through which the argument is constructed
Critical analysis of AI-generated material where relevant	Direct generation or instructor-provided AI material
Transparency and academic integrity	Particular technological tools used in the process

Table 1. Disciplinary constants and organisational variables.

Methodologically, the article should be understood as a reflective pedagogical case study rather than as an experimental evaluation of effectiveness. The account draws upon aggregated assessment results, anonymous institutional student feedback, instructor observation during implementation, and the previously published report of the introductory activity (Arrayás & Heredia, 2026). These materials are used descriptively to reconstruct how the teaching design developed through and after the two implementations; they are not used to infer causal effects of generative AI use or pathway choice.

3. VEUS OBLIDADES A L'ANTIGUITAT: MAKING HISTORICAL WORK VISIBLE

The first implemented experience took place in *Introductory Ancient History*, a first-year Ancient History course intended to familiarise students not only with a very broad chronological and geographical range but also with some of the foundational operations of the discipline. The syllabus moved from the Ancient Near East and Egypt to the Greek and Roman worlds, while the course design emphasised the relationship between primary sources and modern historical interpretation and the idea that Ancient History should not be understood as a closed body of knowledge. This context mattered because first-year students were learning historical content while simultaneously learning how historical claims are constructed: how to locate an assertion within a source's context, how to distinguish an ancient text from a modern interpretation, and how to recognise that a historically plausible explanation is not necessarily an explanation that has been demonstrated. The continuous-assessment activity represented 30 per cent of the final grade and incorporated an ungraded intermediate submission before the final version, allowing the assignment to function as a process of revision rather than as a single terminal product. Two deliberately distant historical settings were selected: Hammurabi's Babylon and the Roman Republic during the

Second Punic War. Their documentary and cultural differences helped prevent the methodological lesson from becoming tied to one historical episode. Students were instead required to transfer the same operations of evidence assessment between different ancient societies.

The activity asked students to construct fictional but historically situated actors from primary evidence and academic scholarship. They had to define social, legal, economic, and personal variables that were plausible within the chosen society. The fact that the individual was fictional did not create unlimited imaginative freedom; in some respects, it created a stricter methodological demand. Since the person had never existed, no individual detail could be independently verified, and every aspect of the reconstruction therefore had to remain within the boundaries of what the sources allowed historians to infer about people occupying comparable positions. Only after this initial contextualisation did generative AI enter the activity. The system could expand the biography, suggest everyday experiences, explain relationships, or supply motivations, and its capacity to do so fluently was precisely what made it useful. The generated answer could combine documented information, reasonable inference, and entirely unsupported elements while projecting modern assumptions onto ancient societies or naturalising behaviour that the evidence could not establish. The exercise therefore began historically after generation rather than ending there. Students had to return to the documentation and decide what could be retained, what required qualification, and what should be rejected (Arrayás & Heredia, 2026, pp. 72–73).

This use of generative AI corresponded to a broader direction already visible in teaching within Ancient and Classical Studies. Díaz-Sánchez and Chapinal-Heras (2024, pp. 17–21), for example, introduced image generation after students had first worked with ancient textual evidence and then required them to assess the reliability and iconographic accuracy of the generated images against that evidence. De la Escosura Balbás and Duce Pastor (2025, pp. 28–32) explored guided integration of AI within epigraphic work, while Liu et al. (2026, pp. 28–30) combined assisted production with subsequent critical editing. In each case, the pedagogical value of the technology lies less in accelerating the arrival at a correct answer than in creating an object that can be interrogated. This distinction is particularly important for Ancient History because documentary incompleteness makes plausibility seductive. A coherent explanation may be compatible with everything we know while still exceeding everything we can demonstrate. Generative AI makes this gap extraordinarily visible because it has no inherent reason to stop speaking at the point at which the historian should begin to qualify.

The results of the first implementation were useful primarily because they exposed problems rather than because they demonstrated the effectiveness of the technology. Of 67 students enrolled, 55 submitted valid work and 48 passed the activity, with a mean of 6.46 and a median of 6.50 (Arrayás & Heredia, 2026, p. 74). These data are descriptive and cannot support a causal claim that AI or the activity improved performance. An earlier project infographic presented the proportion of passing submissions (48 of 55; 87.3 per cent) under the heading “success rate”. In the present article, the same figure is treated only as a descriptive pass proportion and not as evidence of causal effectiveness. This difference reflects the more cautious methodological interpretation adopted through subsequent reflection on the 2025–26 implementation.

More significant were the difficulties visible during the process. Some students began or completed the task too late to make effective use of formative feedback; others could identify a particular error in a generated answer but had greater difficulty transforming that detection into a historical argument; some weaker submissions delegated excessive analytical responsibility back to AI, asking the system itself to carry out part of the critique that was supposed to constitute the student’s work (Arrayás & Heredia, 2026, p. 75). Anonymous institutional feedback added another dimension. Students valued the emphasis placed on critical analysis of sources and academic rigour, but some also described difficulty following the structure of an extensive syllabus and noted the absence of textual organisers in presentations that relied heavily upon images. The apparently separate problems converged around organisation. Accurate historical content does not automatically make the relationships between pieces of information equally visible to every learner, just as identifying an AI mistake does not automatically make the reasoning needed to reject that mistake visible.

The most consequential lesson of this first experience was therefore less technological than initially expected. Detecting that an AI system is wrong is not equivalent to understanding why a statement is historically unsustainable, just as providing a sequence of support does not guarantee that the student can eventually transform that sequence into an argument. We needed to distinguish more explicitly between the operations every student had to perform—compare, contextualise, justify, revise—and the forms through which those operations could be organised. The question ceased to be which activity should be designed for which neurotype and became which elements of historical research must remain common because they constitute the competence being assessed. This shift would shape the second implementation much more directly.

4. FROM DIAGNOSTIC ASSOCIATION TO THREE LEARNING PATHWAYS

The second experience took place in *War and Imperialism in Antiquity*, an advanced course on war and imperialism for students already familiar with the principal periods, sources, and problems of Ancient History. Its academic context differed substantially from the introductory course. The course was not conceived as a narrative sequence of campaigns and battles. War was treated as a historical phenomenon affecting social structures, political relationships, and subjected populations, and the programme moved across the Near East, Egypt, Greece, and Rome while paying particular attention to imperial dynamics. The course design also explicitly incorporated what the project called “forgotten voices”—women, children, foreigners, enslaved people, and other groups conventionally marginalised within narratives of warfare and power. Interestingly, anonymous institutional feedback indicated that this orientation did not coincide with every student’s expectation of what a course on war should contain. Some respondents requested more battles, campaigns, units, and military development, while another perceived elements of the course as repetition of previously studied material. These comments were pedagogically useful because they showed that returning to a familiar historical process from a different historiographical question does not guarantee that students will automatically recognise the new intellectual operation being required.

The principal project, worth 40 per cent of the final grade, therefore combined the thematic concern with forgotten voices, critical use of AI, and a more explicit reconsideration of the architecture of assessment. Seven case studies were offered: power and inequality in Middle Kingdom Egypt; Neo-Assyrian violence and propaganda; Achaemenid imperial order and repression; Athenian imperialism and the contradictions of freedom discourse; war and social disintegration during the Peloponnesian War; conquest, resistance, and violence in Roman Hispania; and war, empire, and women in the Roman world. The chronological range was not intended to create a general history of ancient warfare. The common operation was instead to observe how the explanation of conflict changes when the position from which the narrative is organised changes. Students first worked with a narrative centred upon imperial power and examined how conquest or domination could be presented as necessary, beneficial, or inevitable. Attention then shifted towards territories, subjected communities, and civilian populations, and finally towards actors easily diluted within a narrative organised around power: women, children, enslaved people, defeated populations, and other poorly represented groups. Generated answers were never treated as sufficient interpretations. They had to be compared with ancient evidence and scholarship, with particular attention to moralising or teleological language, historiographical commonplaces, structures of

legitimation, and absences that persisted even when the prompt explicitly requested a change of perspective.

This sequence also clarified the historiographical meaning of “forgotten voices”. The objective was not to replace an imperial narrative with a supposedly authentic generated voice of the oppressed. Changing a prompt cannot recover evidence that ancient documentation has lost, and AI cannot legitimately manufacture absent experiences. The exercise was valuable precisely because it showed how easily narrative hierarchy could be mistaken for neutrality. A generated account could contain correct dates, persons, and events while still assuming without justification that the perspective of the conquering state constituted the natural centre of the story. Conversely, asking the system to foreground women, enslaved people, or defeated communities could produce narratives filled with unsupported psychological detail in an effort to satisfy the request. The methodological task therefore remained identical across perspectives: distinguish what the evidence makes visible from what the narrative machinery of the model has supplied.

The organisational innovation of the second course was more directly connected with neurodiversity. Instead of two routes carrying even an indirect diagnostic association, three possible forms of organisation were made available to the entire class. The first was guided and sequential, moving progressively through contextualisation, the imperial narrative, social consequences, relegated actors, and critical comparison. The second was systematic and comparative, allowing students to organise relationships among ancient sources, generated outputs, and modern historiography explicitly through tables or matrices. The third was synthetic and visual, enabling maps, timelines, conceptual schemes, or representations of power relationships to form part of the construction of the argument. None was identified as an autistic, ADHD, dyslexic, or neurotypical route, and no student had to justify the choice. What mattered was not why a student preferred a particular organisation but what all three routes shared: primary sources, academic scholarship, contextualisation, critical examination of generated responses, identification of mechanisms of legitimation and exclusion, distinction between evidence and argument, and production of an independent historical synthesis. The architecture changed; the historical standard did not.

The available data must again be interpreted descriptively. The register contained 52 students, of whom 45 had a numerical grade for the principal project at the point analysed. The mean was 7.50, the median 8, and the range 4–10, with 41 of 45 grades reaching or exceeding 5. No causal comparison with the introductory experience is possible because the courses differed in level, composition, content, and assessment weighting, and pathway choice was not

systematically recorded alongside each grade. For the reflective purpose of this study, instructor observation was more useful for identifying design issues than the aggregate numerical results. Most students continued to use structures close to conventional academic formats even where visual or more markedly multimodal possibilities existed. Rather than indicating that the innovation had failed, this result clarified what meaningful choice should look like. Providing several routes does not require those routes to be selected in equal proportions, and it should certainly not transform formal innovation into another requirement. For many students, the conventional format continued to work well; the difference was that it ceased to be the only legitimate route available.

5. CRITICAL AI LITERACY WITHOUT COMPULSORY AI ADOPTION

The second implementation also exposed an inconsistency that had survived our attempt to diversify the organisation of learning. Anonymous institutional survey responses revealed at least two divergent responses to AI: one student explicitly rejected assignments involving the technology, while another identified the combination of artificial intelligence and attention to historically marginalised groups as a positive feature of the course. The number of responses was far too small for generalisation, but the contrast was sufficient to prompt reconsideration of a design assumption. We had accepted that students might organise the same historical work differently, yet continued to require the same answer to another question: whether they would interact personally with a generative system. Students had to generate and document prompts, responses, dates, and models. Choice applied to the architecture of the assignment but not to technological participation.

Existing accounts of AI literacy commonly distinguish knowledge and understanding, use and application, evaluation and creation, and ethical dimensions (Ng et al., 2021). Our implementation suggests, however, that critical competence should not be equated with compulsory personal use. Critical knowledge of a technology and personal adoption of that technology are not the same thing. Direct interaction undoubtedly provides experiences that are difficult to communicate entirely through description, and generative systems already intervene in information retrieval, writing, research, visual production, and public representations of the past. A historian should therefore understand what such systems can do and where their outputs become methodologically problematic. Yet this does not mean that every student should be required to incorporate a commercial generative platform into their personal study practice. Objections may concern ethical, environmental, institutional, or personal considerations; alternatively, a student may simply not wish to use a chatbot for academic work.

Universities cannot remove curricular content whenever a student dislikes a topic or method, but they can distinguish between the competence of analysing a technology and the act of adopting it. This distinction is especially important in an inclusive framework because an otherwise flexible activity can become unnecessarily prescriptive at precisely the technological point where students may have legitimate and substantially different preferences.

For this reason, future implementation should allow several modes of participation while preserving a common analytical requirement. Students who wish to generate material directly can continue to do so, documenting prompt, tool, model or version, date, and relevant settings. Students who do not wish to generate material personally can receive a corpus prepared by the instructor, including the same metadata. A third possibility is a comparative corpus containing responses produced with different prompts or systems, which can make model variability itself part of the analysis. The instructor-provided corpus has an additional methodological benefit: when every student produces an independent response, differences in the final analysis are entangled with differences in prompt wording, model version, and generated output. Giving all students the same documented material controls part of that variation and makes it easier to distinguish differences in historical judgement from differences in what the machine happened to produce. Table 2 summarises these modes of participation.

Mode	Student activity	Material provided	Common competence
Direct interaction	Generate, document, and analyse an output	Student records prompt, model/version, and date	Evidence-based critical evaluation
Instructor-provided corpus	Analyse a documented generated output	Instructor supplies complete output and metadata	Evidence-based critical evaluation
Comparative corpus	Compare multiple generated responses	Instructor supplies controlled outputs with metadata	Evidence-based critical evaluation

Table 2. Models of participation in critical AI activities.

This distinction should not be transformed into a new hierarchy in which refusing AI becomes an automatic sign of academic authenticity or ethical superiority. Criticality is an operation, not a predetermined conclusion. A student may reasonably decide that a system is useful for a particular stage of work, another may decide that it adds little, and another may accept some uses but reject others. The academic competence lies in the capacity to justify that judgement. A student who chooses generative AI must verify and document it; a student who chooses not to use it personally remains responsible for understanding and analysing generated material where that analysis forms part of the course. Critical

AI literacy should therefore not be defined as the production of more AI users, but as the development of historians capable of deciding when a tool contributes to historical work, when it distorts that work, and when it is unnecessary.

6. FROM IMPLEMENTATION TO A TRANSFERABLE FRAMEWORK

The two experiences allow the project to be translated into a framework that can operate beyond the specific courses in which it was developed. Transferability does not mean reproducing the original assignment word for word. Hammurabi’s Babylon, the Second Punic War, Athenian imperialism, or Middle Kingdom Egypt can all be replaced. What must survive is the historical operation that made the activity pedagogically useful. The same applies to AI. An exercise is not transferable merely because students can use a chatbot in another course. The transferable element lies in what students are required to do with a generated object: identify the evidentiary basis of a reconstruction, examine the narrative consequences of changing perspective, interrogate a visual representation, verify references, or distinguish evidence from plausible fabrication. Five activity designs can therefore be systematised from the implemented work and the broader project. The first and third were directly tested in the courses discussed above; the remaining designs develop activities already envisaged during the project and partially represented in its earlier Open Educational Resources (Center for the Innovation in Ancient Worlds & Heredia, 2026). Those resources document an intermediate stage of the project rather than the final framework presented here. Further analysis and reflection on the 2025–26 classroom implementations subsequently led us to revise several assumptions retained in those resources. Table 3 gives an overview of the five designs before they are discussed in turn.

Activity	Central historical operation	Possible AI role
Evidence-bounded historical reconstruction	Separate evidence, inference, and invention	Extend a historically constrained biography for subsequent audit
Simulated historical voices	Examine the limits of reconstructing poorly documented experience	Produce a fictional first-person voice to be interrogated
Reframing war and empire	Compare narrative perspective and structures of legitimization	Generate or provide alternative accounts from changing perspectives
Visual representation audit	Treat visual representation as a set of historical claims	Generate or provide historical images for evidence-based analysis
Source verification lab	Distinguish plausibility from traceability	Verify references and claims in generated material

Table 3. Five transferable activity designs.

6.1. EVIDENCE-BOUNDED HISTORICAL RECONSTRUCTION

The first model generalises the introductory activity. Students construct a fictional individual within a historical context using a restricted evidence base before encountering an AI-generated expansion. The key point is that the generated biography is not assessed for creativity or realism but decomposed into claims. Those claims can be classified as documented, plausible but inferential, unsupported, or anachronistic/contradicted. Students must then provide the relevant primary evidence or scholarship and explain why a detail is retained, qualified, or rejected. This activity is especially appropriate for societies in which surviving documentation creates strong asymmetries between what historians know about institutions or elites and what they can reconstruct about everyday experience. Its methodological strength lies in forcing students to recognise that possible and demonstrable are not synonyms. A vivid story can be historically weak; a restrained reconstruction may demonstrate far better control of evidence. The three pathways can all be used without changing this standard: a guided student may complete each claim in a fixed sequence, a comparative student may place every generated assertion beside its source and scholarly interpretation, and a synthetic student may map the actor's legal, economic, and social relationships before converting the visual structure into a written or oral argument.

6.2. SIMULATED HISTORICAL VOICES AND THE LIMITS OF RECOVERY

The role-playing activity developed within the project can be reframed more critically as an exercise in the limits of historical voice. Generative AI can produce a first-person account from a Babylonian merchant, Roman enslaved woman, child in wartime, or indigenous subject of imperial rule with remarkable fluency, but fluency risks giving documentary absence an artificial sense of recovery. Students should therefore begin by identifying what is known, what is reasonably inferable, and what remains inaccessible before the simulated interview is introduced. They can then classify elements of the generated answers as evidence-based information, plausible inference, unsupported interiority, modern stereotype, contradiction, or unverifiable detail. Particular attention should be directed towards emotion and motivation because these are precisely the elements a model supplies readily in order to make a narrative persuasive, even where historical evidence offers little access to an individual's internal experience. The purpose of the exercise is therefore deliberately contrary to the rhetoric of immersive simulation: AI does not allow the absent person to speak; it allows students to analyse how easily an apparently historical voice can be manufactured from the gaps in the archive. Students can further test whether explicit prompting to mark uncertainty actually changes the model's behaviour, thereby making prompt design

itself an object of historical-methodological critique rather than a technical skill rewarded in isolation.

6.3. REFRAMING WAR AND EMPIRE

The third model retains the structure already tested in the advanced course. The same conflict or imperial process is examined successively from the organising perspective of power, from the consequences for subjected communities, and from the position of actors traditionally marginalised within the narrative. Students compare how necessity, legitimacy, violence, resistance, and social disruption are framed in each version and then return to the sources to determine what has genuinely changed. The methodological safeguard is essential: shifting perspective does not automatically make a generated account more accurate. A narrative centred on conquered populations may be ethically attractive while remaining historically speculative. The exercise therefore teaches two things simultaneously: that conventional narratives frequently normalise the perspective of dominant actors, and that correcting this hierarchy cannot mean inventing evidence for those whose experiences are poorly preserved. Because the historical operation is comparative rather than merely generative, the activity can also be carried out entirely with an instructor-provided corpus, enabling students who prefer not to interact directly with an AI platform to demonstrate the same competence.

6.4. THE POWER OF REPRESENTATION

The fourth model turns generated images into historical arguments. A depiction of an ancient household, marketplace, battlefield, or ritual communicates propositions about architecture, clothing, gender, ethnicity, labour, hierarchy, bodily appearance, and spatial organisation before any explanatory text is read. Students receive a heterogeneous evidence dossier and either generate or analyse a small set of images, identifying the claims embedded in each one and classifying them as supported, possible, contradicted, unverifiable, or stereotyped. The exercise is particularly useful because visual plausibility is powerful: an image can make a modern convention feel materially present. Studies of text-to-image systems have documented the amplification of demographic stereotypes in generated images (Bianchi et al., 2023). In historical teaching, representational diversity must therefore remain an object of verification rather than being treated as a guarantee of historical adequacy. Asking explicitly for women, enslaved people, or ethnic diversity does not solve the problem automatically: the resulting representation must still be tested against the available ancient evidence. The objective is not necessarily to produce a final “accurate” image, since accuracy itself may imply certainty unavailable to the historian, but to produce an annotated reconstruction or visual brief in which uncertainty is made visible rather than

hidden. In this sense the activity extends the same method identified by Díaz-Sánchez and Chapinal-Heras (2024, pp. 17–21): the generated image becomes useful when it sends the student back to the sources rather than when it replaces them.

6.5. TRUST, VERIFY, AND CONTEXTUALISE

The fifth model extends the same disciplinary principle to source verification. Empirical research on ChatGPT-generated bibliographies has documented both wholly fabricated citations and substantive errors within citations to real works (Walters & Wilder, 2023). Students receive a generated passage containing a mixture of accurate statements, weakly supported claims, bibliographical distortions, and, where appropriate, nonexistent references. They divide the text into discrete claims and connect each one with a verification route: primary source, scholarly publication, library catalogue, journal platform, DOI record, or other reliable evidence. Each item can then be classified as verified, partially supported, misleading, false, or unverifiable. The exercise should not stop at locating an obviously fabricated citation. A real publication may fail to support the proposition attributed to it, while a bibliographical error does not automatically make the underlying historical claim false if independent evidence sustains it. What students learn is therefore a broader practice of evidentiary anchoring: formal plausibility must be converted into traceable support. Although generative AI makes the problem particularly visible, the underlying scholarly operation is not new. Historical method has always depended upon the distinction between something that sounds convincing and something that can be demonstrated. A further methodological distinction is important in educational settings: technical error, poor scholarly practice, and deliberate misconduct are not equivalent categories. The teaching principle required here is narrower and fundamental: conclusions about knowledge require evidence proportionate to the claim.

7. ASSESSMENT: EVALUATING HISTORICAL OPERATIONS RATHER THAN SURFACE FORM

A model based upon multiple pathways succeeds only if its assessment criteria genuinely remain common. Otherwise, flexibility becomes cosmetic. A lecturer may formally permit a visual synthesis while continuing to allocate most marks according to conventions of continuous prose, or may allow direct and indirect engagement with AI while implicitly rewarding sophisticated prompting rather than historical analysis. The assessment framework must therefore follow the disciplinary constants rather than the surface characteristics of a particular format. Historical contextualisation, primary-source analysis, engagement with scholarship, distinction between evidence and inference, critical evaluation of generated claims, argument, and transparency can all be demonstrated through

more than one organisational route, provided the required reasoning remains visible. This does not imply that every medium is automatically equivalent. If a visual representation conceals rather than explains the relationship between evidence and interpretation, supplementary commentary may be necessary. Equivalence refers to the learning outcome, not to the assumption that every form of expression reveals every competence equally well. Table 4 presents a common analytical assessment framework organised around these operations.

Criterion	Core question
Historical contextualisation	Are claims, actors, and sources accurately situated within the relevant context?
Primary evidence	Is appropriate ancient evidence used critically, with its limits recognised?
Academic scholarship	Are modern interpretations used accurately and distinguished from primary evidence?
Evidence and inference	Does the student distinguish what can be demonstrated from what is reconstructed?
Critical AI analysis	Is generated material verified and interpreted rather than merely accepted or rejected?
Historical argument	Does the final interpretation follow from evidence and critical reasoning?
Transparency	Are sources, AI involvement where applicable, and the student's own contribution traceable?

Table 4. Common analytical assessment framework

The experience of the introductory course makes the wording of such criteria especially important. An assessment criterion such as “identifies AI errors” risks making detection the endpoint of the task. The relevant historical operation is stronger: the student must demonstrate, through evidence and scholarship, why a generated claim should be retained, qualified, or rejected. Similarly, multimodality should not create hidden technical requirements unrelated to Ancient History. If video editing, graphic design, or audio production is not a learning outcome, a highly polished digital product should not receive additional academic credit merely for production quality. A plain comparative table may communicate historical judgement more effectively than an elaborate infographic. Inclusive assessment should broaden the routes through which competence can become visible without confusing novelty with quality. The use of AI should likewise attract neither positive nor negative marks by itself: what matters is the quality of verification, reasoning, and methodological transparency.

8. WHAT CHANGED THROUGH IMPLEMENTATION

The two cases did not produce a closed pedagogical model; rather, each forced the framework to change. The first implementation showed that the internal operations of historical criticism need to be made explicit. Students may recognise that an AI-generated statement looks wrong without knowing how to construct the evidentiary chain required to reject it. Future versions should therefore externalise that sequence: claim → evidence → comparison → evaluation → justified decision. The same implementation showed that a flexible final product does little to address difficulties in planning if the project itself remains temporally opaque. Milestones, intermediate submissions, planning sheets, and continuous low-pressure feedback can therefore function as accessibility measures independently of any diagnostic status. One of the project's earlier visual resources described this kind of process through a "learning backpack", a system in which progress, feedback, and incomplete stages remain visible during the activity rather than appearing only in final evaluation (Center for the Innovation in Ancient Worlds & Heredia, 2026). Although other aspects of those early resources were subsequently revised, this principle remained useful: complex assignments should show students where they are within the process, not merely where they must eventually arrive.

The second implementation changed a different assumption. Making alternative forms available did not lead most students to abandon traditional academic structures, and there is no pedagogical reason why it should. The purpose of inclusive design is not to maximise formal diversity but to remove the monopoly of a single route where that monopoly is unnecessary. A student who chooses a conventional essay should not be treated as having missed the innovation, just as a student who uses diagrams to plan an argument should not be regarded as completing an easier task. The meaningful outcome is that students can select among bounded, equivalent forms without having to justify the choice through diagnosis or deficit. This observation also cautions against assuming that the mere presence of multiple options demonstrates inclusion: alternatives must be intelligible, genuinely equivalent, and supported by assessment criteria that do not quietly privilege the conventional form.

Most importantly, technological participation emerged as another dimension of learner diversity. The project had initially treated critical use of AI as if direct interaction were inherently required, yet the students' divergent responses showed that this assumption deserved the same scrutiny as organisational uniformity. Once the competence is identified as the critical evaluation of generated historical material, an instructor-provided corpus becomes a legitimate alternative. This

insight is particularly useful because it prevents AI literacy from becoming technologically prescriptive. Universities have strong reasons to teach students how generative systems affect contemporary research and knowledge production, but the educational objective should remain judgement rather than compliance with a particular tool culture. A critical approach should allow a student to conclude that a tool is useful, unnecessary, or objectionable in a given context, provided that the conclusion is reasoned and the relevant disciplinary competence has been demonstrated.

Taken together, these changes suggest a broader conclusion about inclusion. Greater flexibility requires greater disciplinary precision. “Complete the task however you prefer” is not an inclusive methodology; it may create additional ambiguity and cognitive load. A stronger model tells students which operations are common, which standards will be applied, and where genuine choice exists. On the basis of these implementations, we therefore adopt bounded choice—a small number of clearly articulated alternatives under common criteria—as a design principle, rather than assuming that a larger number of options is intrinsically more inclusive. The same principle protects academic standards because variation occurs only after the instructor has determined what must remain stable. The framework consequently shifts attention away from the search for an ideal format and towards a more fundamental task: defining the evidence of learning with sufficient clarity that more than one route can reach it.

9. OPEN EDUCATIONAL RESOURCES AND ITERATIVE DESIGN

Open Educational Resources were produced during the project as part of its original commitment to transfer and remain openly available through Zenodo (Center for the Innovation in Ancient Worlds & Heredia, 2026). They should, however, be understood as documenting an earlier stage in the development of the pedagogical model rather than as a definitive visual representation of the framework presented in this article. This distinction became necessary because subsequent analysis and reflection on the 2025–26 implementations led us to revise some of the assumptions retained in the original resources.

Most importantly, some of the earlier visual materials retained an orientative association between particular learning pathways and diagnostic profiles, and presented direct interaction with generative AI more centrally than the model eventually developed. Subsequent analysis of the classroom experiences discussed above led us to revise both assumptions. The present framework makes all organisational pathways available to all students without diagnostic assignment, and distinguishes critical AI literacy from compulsory personal adoption of generative AI. The earlier resources therefore remain useful as records of the

project's development, but these aspects should not be treated as recommendations of the revised framework.

This discrepancy also reveals a broader problem in the transfer of teaching innovation. Educational resources can stabilise a model visually at a particular moment even while subsequent analysis and reflection continue to change the assumptions beneath it. Transferability should therefore concern the underlying disciplinary operation rather than the preservation of a fixed activity or diagram. "Interview an ancient person with AI" is technologically reproducible but pedagogically weak if the instructor does not explain why the simulation matters. "Use a simulated voice to teach the limits of reconstructing poorly preserved experience" is a transferable historical operation. Likewise, "generate a picture of antiquity" is not in itself an educational objective, whereas "identify and verify the historical propositions embedded in a visual reconstruction" can be adapted across periods and disciplines. What transfers is the distinction between disciplinary requirements and organisational conventions, and with it the sequence implementation → reflection → revision → adaptation.

10. DISCUSSION: INCLUSION AS DISCIPLINARY CLARITY

The framework developed through the project complicates the familiar opposition between academic rigour and educational flexibility. If rigour is identified with uniform procedure, alternative pathways inevitably appear to weaken standards. Yet a university assignment normally combines several different elements: genuine disciplinary requirements, institutional habits, practical preferences, and conventions inherited from earlier forms of teaching. Treating all of them as equally essential makes the standard appear clearer than it really is. Separating them requires instructors to articulate what their discipline actually demands. A historian should be able to explain why an interpretation follows from evidence; whether the student first saw that relationship in linear notes, a source matrix, or an annotated diagram is secondary unless mastery of a particular scholarly genre is itself an explicit learning outcome. A historian should distinguish plausible reconstruction from evidence; whether the material under criticism was generated personally through a chatbot or supplied by the instructor is secondary if the competence being evaluated is the historical analysis of that material. A historian studying marginalised ancient actors must recognise both their historical significance and the limitations of an archive that often preserves them unevenly; a fluent first-person simulation cannot solve that documentary problem, but it can make the problem unusually visible.

Generative AI has been particularly revealing because it acts as a stress test for these distinctions. Its fluency exposes the difference between plausibility and

evidence; its simulations expose the temptation to replace documentary absence with invented presence; its capacity to generate images exposes the historical claims hidden inside apparently immediate visual representation; its fabricated or distorted references expose the difference between bibliographical form and source traceability; and its uneven reception among students exposes the assumption that technological participation itself is neutral. These problems do not originate with AI. Historians have always dealt with fragmentary evidence, narrative hierarchy, visual convention, unreliable citation, and the need to justify methodological choices. What AI changes is the speed and persuasiveness with which the underlying problems can be reproduced. Its pedagogical value therefore depends upon resisting the temptation to turn the technology itself into the learning outcome. The strongest activities use AI to make disciplinary judgement necessary, not to make historical work disappear behind a technically impressive product.

The same argument applies to neurodiversity. The value of neurodiversity-aware teaching does not lie in constructing increasingly detailed typologies of what an autistic, ADHD, dyslexic, or neurotypical student supposedly needs. Such a project risks repeating the homogenisation that the concept of neurodiversity originally helped challenge. Its more productive contribution is to make instructors question unnecessary uniformity. If a predictable barrier can be removed without changing the disciplinary standard, there is little reason to wait until a student obtains permission to avoid it. If a particular convention is genuinely essential, the instructor should be able to explain why. Individual accommodations remain necessary because no general design can anticipate every circumstance, but a more accessible baseline can reduce the number of occasions on which students must convert difference into a formally documented deficit before accessing an alternative. This may be one of the most productive implications of neurodiversity-aware teaching: not the construction of increasingly precise categories of learner, but a more precise description of what the discipline itself requires.

This approach also helps avoid romanticising flexibility. The framework therefore does not treat a larger number of options as inherently more inclusive. An unrestricted menu can shift additional organisational decisions onto the learner, whereas bounded choice limits variation to alternatives whose relationship to the common learning outcomes has been established in advance. The three pathways are useful not because three is inherently optimal, but because each embodies a recognisable mode of organising complex historical work while keeping the core requirements stable. Instructors can adapt the number and form of routes, but the underlying sequence should remain the same: define the

competence first; identify unnecessary uniformity second; provide alternatives only where equivalence can be demonstrated. In this sense, inclusive design is best understood not as the relaxation of disciplinary expectations but as a method for separating those expectations from the habits that have accumulated around them.

11. CONCLUSION

The principal outcome of the project has become less technological as it has developed. Generative AI initially appeared to be one of the most conspicuous innovations because it offered a new way to produce narratives, simulations, and images concerning the ancient world. In practice, its greatest value was methodological. It generated coherent reconstructions where the evidence was incomplete, supplied certainty where historians should hesitate, reproduced narrative hierarchies while appearing neutral, and transformed bibliographical plausibility into apparent authority. Each of these behaviours returned students to an old historical question: what does the evidence actually allow us to say? The answer could not be delegated to the same system that had produced the problem. Students had to contextualise, compare, verify, justify, and revise. The attempt to make those operations more visible simultaneously changed our approach to inclusion. Once we identified which parts of the historical task genuinely constituted the competence, it became more difficult to defend the assumption that every student had to organise those operations through an identical sequence. Earlier learning itineraries therefore evolved into three pathways available without diagnostic assignment, and the same logic eventually extended to AI itself. If the academic objective is critical analysis of generated material, direct personal generation may be pedagogically valuable but cannot automatically be treated as the only legitimate way to demonstrate the competence. A documented common corpus can permit the same historical operation while respecting different relationships with technology.

The framework presented here should not be understood as a completed experimental model. The numerical data from the two implementations are descriptive, the courses are not directly comparable, and pathway choice was not systematically recorded alongside individual performance in the advanced course. Future implementation should therefore track pathway selection more carefully, examine whether students move between organisational routes during a semester, and assess how direct and corpus-based forms of AI participation affect both engagement and analytical performance. These limitations do not invalidate the design development described here, because the central claim is not that one pathway produces better grades than another. Rather, the project generated a design principle: same standards do not require the same pathway. This principle

does not relativise Ancient History teaching; it requires greater clarity about what Ancient History teaching is actually evaluating. Contextualising evidence, distinguishing source from interpretation, identifying anachronism, confronting generated claims with documentation, engaging scholarship, and constructing a defensible argument remain common expectations. The order in which those operations are organised, the intermediate structures through which students make them visible, and—in appropriate cases—the mode through which generated material is accessed need not always remain identical.

The same principle provides a useful definition of critical AI literacy. Its objective should not be measured by how frequently students use generative systems or by how sophisticated their prompts become. A historically trained student should be capable of interrogating a generated narrative, identifying what it assumes, tracing its claims to evidence, recognising unsupported certainty, understanding what the archive cannot recover, and deciding when technological assistance contributes to historical work, when it distorts that work, and when it is simply unnecessary. That judgement, rather than technological adoption itself, is the competence we seek to teach.

DECLARATIONS

Funding. This article forms part of the Universitat Autònoma de Barcelona Teaching Innovation Project GI517492, *Veus oblidades a l'Antiguitat: docència inclusiva i pensament crític en temps d'intel·ligència artificial*. Earlier Open Educational Resources produced within the same project remain available independently through Zenodo. The present article develops and revises that work through implementation evidence, five adaptable activity models, and a common analytical assessment framework. The project supported the development, implementation, evaluation, reflection, and transfer of these teaching resources.

REFERENCES

- Arrayás, I. (2005). *Morfología histórica del territorio de Tarraco (ss. III–I a.C.)*. Barcelona: Publicacions i Edicions de la Universitat de Barcelona.
- Arrayás, I., & Heredia, C. (2026). 'Repensar la Història Antiga: veus oblidades, docència inclusiva i intel·ligència artificial'. In J. L. Muñoz Moreno, C. Ruiz Bueno, & S. Sabaté Diaz (Eds.), *Innovació docent: de la idea a la pràctica en Educació Superior* (pp. 69–75). Servei de Publicacions de la Universitat Autònoma de Barcelona. <https://doi.org/10.5565/lib/9791388031298>
- Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2023). 'Easily accessible text-to-image generation amplifies demographic stereotypes at large scale'. In *AccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1493–1504). Association for Computing Machinery. <https://doi.org/10.1145/3593013.3594095>

- Center for the Innovation in Ancient Worlds, & Heredia, C. (2026). *CIAW Teaching Infographics: Inclusive Learning, Critical AI Literacy, and Responsible Knowledge Practices* [Open Educational Resources]. Zenodo. <https://doi.org/10.5281/zenodo.21763472>
- Clouder, D., Karakus, M., Cinotti, A., Ferreyra, M. V., Amador Fierros, G., & Rojo, P. (2020). 'Neurodiversity in higher education: A narrative synthesis'. *Higher Education*, 80(4), 757–778. <https://doi.org/10.1007/s10734-020-00513-6>
- Cumming, T. M., & Rose, M. C. (2022). 'Exploring universal design for learning as an accessibility tool in higher education: A review of the current literature'. *The Australian Educational Researcher*, 49, 1025–1043. <https://doi.org/10.1007/s13384-021-00471-7>
- De la Escosura Balbás, C., & Duce Pastor, E. (2025). 'Epigrafía digital e inteligencia artificial: preguntas y respuestas de una experiencia docente'. *AI & Antiquity: Journal of Teaching and Technology in Ancient Studies*, 1(1), 15–34. <https://doi.org/10.64946/aiantiquity.v1i1.001>
- de Sousa Santos, B. (Ed.). (2007). *Cognitive Justice in a Global World: Prudent Knowledges for a Decent Life*. Lexington Books.
- Díaz-Sánchez, C., & Chapinal-Heras, D. (2024). 'Use of open access AI in teaching Classical Antiquity: A methodological proposal'. *Journal of Classics Teaching*, 25(49), 17–21. <https://doi.org/10.1017/S2058631023000739>
- Dwyer, P. (2022). 'The neurodiversity approach(es): What are they and what do they mean for researchers?' *Human Development*, 66(2), 73–92. <https://doi.org/10.1159/000523723>
- Heredia, C. (2026). 'Flipped learning and AI-supported itineraries for neurodivergent learners: A case study in Ancient History'. *Journal of Classics Teaching*, 1–9. <https://doi.org/10.1017/S2058631026100956>
- Howard, P. L., & Sedgewick, F. (2021). "Anything but the phone!": Communication mode preferences in the autism community'. *Autism*, 25(8), 2265–2278. <https://doi.org/10.1177/13623613211014995>
- Le Cunff, A.-L., Giampietro, V., & Dommett, E. J. (2024). 'Neurodiversity and cognitive load in online learning: A focus group study'. *PLoS ONE*, 19(4), e0301932. <https://doi.org/10.1371/journal.pone.0301932>
- Leibowitz, B. (2017). 'Cognitive justice and the higher education curriculum'. *Journal of Education*, 68, 93–111. <https://doi.org/10.17159/2520-9868/i68a03>
- Liu, Y., D'Elia, D., & Palermo, R. (2026). 'Let's chat about archaeology: Responsible and thoughtful use of AI tools in the classroom, a case study'. *AI & Antiquity: Journal of Teaching and Technology in Ancient Studies*, 2(1), 25–36. <https://doi.org/10.64946/aiantiquity.v2i1.002>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). 'Conceptualizing AI literacy: An exploratory review'. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Oller, J. (2016). 'La Layetania interior en época augústea: un ejemplo de reordenación territorial en la Hispania Citerior Tarraconense'. *Dialogues d'histoire ancienne*, 42(1), 137–175. <https://doi.org/10.3917/dha.421.0137>
- Prieto, A. (1980). 'La pervivencia del elemento indígena en la Bética'. *Faventia*, 2(1), 37–46.
- Sanahuja, M. E., & Prieto, A. (1981). 'El papel de la mujer en las Bacanales romanas'. *Memorias de Historia Antigua*, 5, 143–152.
- Universitat Autònoma de Barcelona. (2024). *Protocol d'ajustos curriculars per a l'estudiantat amb necessitats específiques de suport educatiu* [Institutional protocol]. Approved by the Consell de Govern, 11 July 2024.
- Visvanathan, S. (1997). *A Carnival for Science: Essays on Science, Technology and Development*. Oxford University Press.
- Walters, W. H., & Wilder, E. I. (2023). 'Fabrication and errors in the bibliographic citations generated by ChatGPT'. *Scientific Reports*, 13, 14045. <https://doi.org/10.1038/s41598-023-41032-5>