Original Article

Received/Accepted Dates

27.09.2026/12.10.2026

DOI :10.52096/usbd.10.48.22

International Journal of Social Sciences

Uluslararası Sosyal Bilimler Dergisi

www.sobider.net e-ISSN: 2548-0685

IJSS

Epistemological Vigilance and Instructional Design in the Age of Generative AI: Aligning Evidence-Based Inquiry, Spatial-Historical Thinking, and Process Assessment in the Century of Türkiye Maarif Model and the C3 Framework

Halil Ersin AVCI

Ph.D., Canakkale Onsekiz Mart University Education Faculty, Social Science Education halilersinavci@comu.edu.tr, ORCID: 0000-0003-1580-7803 Researcher ID: N-8358-2018

Abstract

The proliferation of generative artificial intelligence (GenAI) and large language models (LLMs) in K–12 and teacher education poses a fundamental epistemological challenge to Social Studies instruction. Whereas disciplinary thinking in History, Geography, Civics, and Economics relies on provenance (sourcing), contextual friction between competing records, spatial scale analysis, and deliberative trade-offs, probabilistic language models synthesize fluent, single-voiced prose that erases source boundaries and invites cognitive offloading. Existing practitioner literature largely addresses classroom AI through generic prompt-engineering tips or binary bans. By contrast, this study develops a comprehensive, multi-disciplinary instructional design architecture that operationalizes AI governance across the 2024–2026 Century of Türkiye Maarif Model (TYMM) Social Studies Curriculum (Grades 4–7) and the United States National Council for the Social Studies (NCSS) C3 Framework Inquiry Arc.

Employing qualitative document analysis and normative instructional design synthesis across official curriculum and ethics mandates (MEB, 2025a, 2025b, 2026a, 2026b; NCSS, 2013), cognitive architecture theory (Mayer, 2009; Mayer & Moreno, 2003; Sweller, van Merriënboer, & Paas, 2019), and disciplinary literacy scholarship (Ashby &

Araştırma Makalesi

Makaleye Atıf Bilgisi

Avcı, H. E. (2026). Epistemological Vigilance and Instructional Design in the Age of Generative AI: Aligning Evidence-Based Inquiry, Spatial-Historical Thinking, and Process Assessment in the Century of Türkiye Maarif Model and the C3 Framework. International Journal of Social Science (IJJS Journal), *(e-ISSN:***2548-0685***)* Vol:10, Issue:48; s. 402- 419.

DOI: **10.52096/usbd.10.48.22**

Gönderim: 27.09.2026

Kabul: 12.10.2026

SOSYAL BİLİMLER DERGİSİ (ULUSLARARASI HAKEMLİ DERGİ)

Lee, 1987; Gersmehl, 2014; Hess, 2009; Wineburg, 1991, 2001; Wineburg & McGrew, 2019), the article advances five original contributions: (1) a Comparative Inquiry-Governance Matrix demonstrating how TYMM's sixth learning area ('Technology and Social Sciences'), Systems Thinking, nine literacies (OB1–OB9), and Virtue–Value–Action (EDE) framework extend the C3 Inquiry Arc for the AI era; (2) a Taxonomic Verb-Alignment and Scaffolding Withdrawal Model (across Revised Bloom's and SOLO taxonomies) that prevents 'cognitive verb drift' during the gradual release of responsibility (Modeling -> Guided Practice -> Independent Practice); (3) Domain-Specific AI Verification Protocols across all four core Social Studies disciplines—including contradiction-margin source packages in History (SBAB2, SBAB3, SBAB4), five-component spatial reasoning in Geography (SBAB7, SBAB8, SBAB10), institutional receptor verification and deliberative steel-manning in Civics (SBAB5), and the tripartite statistical rule plus assumption-testing in Economics/Entrepreneurship (SBAB6, SBAB11); (4) a Subtractive Material Engineering Protocol incorporating three rapid pre-classroom tests (Self-Completion, Subtraction, and Blind Peer tests); and (5) a Two-Tier Psychometric Rubric pairing comment-first formative feedback (Butler, 1988) with process-authentication criteria that replace biased AI detectors (Liang et al., 2023).

editorusbd@gmail.com

Keywords: *Social Studies instructional design, generative artificial intelligence, Century of Türkiye Maarif Model (TYMM), C3 Framework Inquiry Arc, constructive alignment, gradual release of responsibility, spatial and historical thinking.*

1. Introduction: The Epistemological Collision Between Probabilistic Language Models and Disciplinary Social Studies

Every instructional medium carries an implicit epistemology: it privileges certain ways of knowing while obscuring others. In Social Studies education—an interdisciplinary field forged from History, Geography, Civics, and Economics to cultivate reflective, evidence-minded citizens (Barr, Barth, & Shermis, 1977; Dewey, 1910, 1916; Evans, 2004; NCSS, 2023)—knowledge is neither a flat inventory of dates and definitions nor a frictionless consensus. Rather, disciplinary understanding in Social Studies emerges precisely from confronting epistemic friction: the discrepancy between a celebratory town newspaper report and a municipal budget ledger from the same year (Wineburg, 1991, 2001); the shift in causal patterns when an environmental or demographic phenomenon is examined at neighborhood, provincial, and national scales (Gersmehl, 2014; National Research Council [NRC], 2006); or the competing value priorities that citizens must negotiate when allocating scarce public resources (Hess, 2009; Westheimer & Kahne, 2004).

Generative artificial intelligence (GenAI) and large language models (LLMs) operate on an epistemological principle that is the exact inverse of this disciplinary friction. Trained to predict the most statistically probable sequence of tokens across vast text corpora, LLMs are engineered to smooth over contradictions, synthesize heterogeneous viewpoints into a single authoritative-sounding voice, and detach claims from their institutional and temporal provenance (künye).

When an elementary or middle school student queries an AI chat interface about a historical reform, a geographical region, or a civic problem, the model does not present a curated packet of primary sources with visible margins of uncertainty; it delivers a polished, pre-digested summary. Similarly, when a preservice or in-service teacher asks an LLM to generate a forty-minute lesson plan or worksheet, the model defaults to low-level recall questions and generic group activities that look visually complete on paper but strip away the specific cognitive labor required by the learning outcome.

Addressing this epistemological collision requires moving beyond generic technological enthusiasm or defensive bans toward rigorous instructional design. In Türkiye, the promulgation of the Century of Türkiye Maarif Model (Türkiye Yüzyılı Maarif Modeli, hereafter TYMM; Talim ve Terbiye Kurulu Başkanlığı, 2024; MEB, 2025b) and the updated Social Studies Curriculum for Grades 4–7 (MEB, 2026b), alongside the Artificial Intelligence in Education Policy Document (2025–2029) (MEB, 2025a) and the Ethical Guidelines for AI Applications in Education (MEB, 2026a), has created an unprecedented institutional framework for competency-based, evidence-driven Social Studies instruction. When analyzed in dialogue with the international NCSS College, Career, and Civic Life (C3) Framework (2013), the Maarif Model offers a rich taxonomy of Conceptual Skills (KB1–KB3), Social Sciences Field Skills (SBAB1–SBAB17), Systems Thinking literacies (OB1–OB9), and Virtue–Value–Action (EDE) commitments. This study constructs an integrated instructional design theory and operational methodology demonstrating how Social Studies educators can align learning outcomes, cognitive scaffolding withdrawal, domain-specific inquiry across History, Geography, Civics, and Economics, subtractive material engineering, and process-authenticated assessment in the age of generative AI.

2. Comparative Curricular Architecture: Bridging the C3 Inquiry Arc and the Maarif Model for the AI Era

A curriculum's resilience in the face of disruptive technologies depends on whether its standards are written as static content to be recalled or as disciplinary operations to be performed. Both the NCSS C3 Framework (2013) and the 2024–2026 TYMM Social Studies Curriculum (MEB, 2025b, 2026b) belong to the latter category. Inspired conceptually by the four-stage progression of the C3 Inquiry Arc—(1) Developing Questions and Planning Inquiries, (2) Applying Disciplinary Concepts and Tools, (3) Evaluating Sources and Using Evidence, and (4) Communicating Conclusions and Taking Informed Action (NCSS, 2013; American Historical Association, 2015)—the Maarif Model decomposes learning outcomes into step-by-step process components (süreç bileşenleri) drawn from Conceptual Skills (KB) and Social Sciences Field Skills (SBAB).

Yet comparative analysis reveals that the 2024–2026 Turkish architecture introduces four structural features that make it particularly well-equipped to govern generative AI and digital information flows in Grades 4–7:

• First, whereas the 2013 C3 Framework narrowed its Dimension 2 core to four traditional disciplines (Civics, Economics, Geography, History) and left the 2010 NCSS theme of 'Science, Technology, and Society' (NCSS, 2010) to state discretion, MEB (2026b) retains and modernizes

'Technology and Social Sciences' (Teknoloji ve Sosyal Bilimler) as a mandatory sixth spiral learning area across all four grade levels, explicitly pairing it with Evidence-Based Inquiry (SBAB2) and Perceiving Change and Continuity (SBAB4).

• Second, TYMM embeds nine cross-curricular literacies—Information Literacy (OB1), Digital Literacy (OB2), Financial Literacy (OB3), Visual Literacy (OB4), Cultural Literacy (OB5), Citizenship Literacy (OB6), Data Literacy (OB7), Sustainability Literacy (OB8), and Art Literacy (OB9)—within an overarching Systems Thinking (Sistem Düşüncesi) component structured across awareness (farkındalık), functionality (işlevsellik), and agency (eylemsellik) levels (MEB, 2025b, pp. 34, 37–40).

• Third, TYMM anchors epistemic inquiry in an explicit ethical ontology—the Virtue–Value–Action (Erdem–Değer–Eylem, EDE) framework organized around the three roof values of Respect, Responsibility, and Justice (MEB, 2025b, p. 36)—which transforms academic honesty, source attribution, and digital privacy from external rules into enacted virtues.

• Fourth, MEB (2025b, p. 15) explicitly stipulates that process components do not have to be executed in a rigid linear sequence, enabling teachers to launch inquiry by confronting students with a flawed or contradictory digital claim in order to provoke foundational questioning. Table 1 maps this comparative inquiry-governance architecture.

Table 1. Comparative Inquiry-Governance Matrix: C3 Inquiry Arc (NCSS, 2013), TYMM Skill Architecture (MEB, 2025b, 2026b), and AI Pedagogical Safeguards

Inquiry Phase & C3 Dimension	TYMM Conceptual & Field Skill Counterparts (MEB, 2025b, 2026b)	TYMM Cross-Curricular Literacies & EDE Values	Epistemological Vulnerability in GenAI & Required Safeguard
Phase I: Problem Framing & Questioning (C3 Dimension 1: D1.1–D1.5)	• SBAB2.1: Curiosity-based questioning • SBAB8.1: Asking geographical questions • KB2.8: Inquiry • KB3.2: Problem solving	• Dispositions: E1.1 Curiosity, E3.2 Open-mindedness, E3.4 Seeking truth • Literacy: OB1 Information Literacy	Vulnerability: AI answers prematurely foreclose student curiosity with closed summaries. Safeguard: Open every lesson with an unanswered inquiry question that requires primary evidence to resolve.
Phase II: Disciplinary Tool Application (C3 Dimension 2: D2.His, Geo, Civ, Eco)	• History: SBAB1 (Time), SBAB3 (Empathy), SBAB4 (Change/Continuity) • Geography: SBAB7 (Spatial), SBAB8, SBAB9, SBAB10 (Maps) • Civics/Econ: SBAB5, SBAB6, SBAB11	• Literacies: OB3 Financial, OB4 Visual, OB5 Cultural, OB7 Data, OB8 Sustainability • Values: Respect, Patriotism, Sensitivity	Vulnerability: AI flattens spatial scale, fabricates first-person historical quotes, and blurs denominational units. Safeguard: Enforce domain-specific verification rules (Section 5).

Inquiry Phase & C3 Dimension	TYMM Conceptual & Field Skill Counterparts (MEB, 2025b, 2026b)	TYMM Cross-Curricular Literacies & EDE Values	Epistemological Vulnerability in GenAI & Required Safeguard
Phase III: Source Interrogation & Corroboration (C3 Dimension 3: D3.1–D3.4)	• SBAB2.2: Gathering info • SBAB2.3: Examining source (provenance SBAB2.3.SB1) • SBAB2.4: Questioning source • SBAB2.5: Interpreting • KB2.4 Analysis, KB2.7 Comparison	• Literacies: OB1 Information, OB2 Digital, OB7 Data Literacy • Roof Value: Responsibility & Honesty (citing explicit warrants)	Vulnerability: LLMs generate uncited prose and hallucinated bibliographies. Safeguard: Treat uncited AI text strictly as a 'claim set' subjected to 4-move lateral verification (Wineburg & McGrew, 2019).
Phase IV: Warranted Products & Civic Action (C3 Dimension 4: D4.1–D4.8)	• SBAB2.6: Creating/sharing evidence-based products • SBAB5.3–5.5: Generating ideas, negotiating, acting • SBAB6.3: Finalizing ideas • KB2.18 Argumentation, KB3.1 Decision	• Social-Emotional: SDB2.2 Collaboration, SDB3.3 Responsible decision • Roof Values: Justice, Benevolence, Responsibility (EDE)	Vulnerability: Students submit AI-written products without cognitive ownership; AI suggests non-existent civic bodies. Safeguard: Require process artifacts (draft, decision log, disclosure) and verify official gov.tr receptors.

3. Methodology: Normative Instructional Design Modeling

This study employs qualitative document analysis (Bowen, 2009) coupled with normative instructional design modeling (Biggs & Tang, 2011; Taba, 1962; Tyler, 1949; Wiggins & McTighe, 2005). In educational research, normative instructional design modeling goes beyond describing current teacher habits; it synthesizes binding curricular standards, statutory constraints, and empirical learning sciences to derive operational design criteria capable of diagnosing and preventing instructional failure. The primary documentary corpus consists of (a) the TYMM Common Text (MEB, 2025b), (b) the 2026 TYMM Social Studies Curriculum for Grades 4–7 (MEB, 2026b), (c) the MEB AI Policy Document and Ethical Guidelines (MEB, 2025a, 2026a), (d) the NCSS C3 Framework (NCSS, 2013), and (e) statutory provisions governing personal data protection (Law No. 6698, 2016) and copyright in educational reproduction (Law No. 5846, 1951, Arts. 33–35, 38). These mandates were cross-analyzed against cognitive load and multimedia learning principles (Kirschner, Sweller, & Clark, 2006; Mayer, 2009; Mayer & Moreno, 2003; Sweller et al., 2011, 2019), cognitive apprenticeship and scaffolding theory (Collins, Brown, & Newman, 1989; Fisher & Frey, 2013; Pearson & Gallagher, 1983; Vygotsky, 1978; Wood, Bruner, & Ross, 1976), disciplinary thinking research across History, Geography, Civics, and Economics, and psychometric validity theory (Black & Wiliam, 1998; Brookhart, 2013; Butler, 1988; Liang et al., 2023; Messick, 1994; Sadler, 1989).

4. Cognitive Architecture, Taxonomic Alignment, and the Gradual Withdrawal of AI Scaffolding

4.1. Diagnosing 'Verb Drift' Across Revised Bloom's and SOLO Taxonomies

The foundational rule of instructional design is that a lesson is a coherent design only when it begins with a learning outcome (öğrenme çıktısı) and returns to that outcome through an aligned activity and an observable piece of learning evidence (öğrenme kanıtı; Biggs, 1996; Biggs & Tang, 2011; Mager, 1997; Webb, 1997; Wiggins & McTighe, 2005). In both Anderson and Krathwohl's (2001) revision of Bloom's taxonomy (Remember, Understand, Apply, Analyze, Evaluate, Create) and Biggs and Collis's (1982) SOLO (Structure of the Observed Learning Outcome) taxonomy (Prestructural, Unistructural, Multistructural, Relational, Extended Abstract), the cognitive level of a task is determined not by the difficulty of the topic or the sophistication of the technological tool, but by the relationship among three design variables: the input material given to the student, the mental operation performed, and the criterion used to judge the product.

When generative AI enters lesson preparation or classroom activity without taxonomic auditing, the most pervasive failure is cognitive verb drift (fiil kayması). In the 2026 TYMM Social Studies Curriculum (MEB, 2026b), the majority of learning outcomes operate at the Relational and Extended Abstract levels of SOLO—requiring students to 'analyze' (çözümleyebilme, KB2.4), 'compare' (karşılaştırabilme, KB2.7), 'interpret' (yorumlayabilme, KB2.14), 'question' (sorgulayabilme, KB2.8 / SBAB2), or 'evaluate' (değerlendirebilme, KB2.17). However, when an AI tool summarizes two historical documents for the student or generates a worksheet dominated by 'According to the text above...' questions, the relational cognitive operation is performed by the algorithm, dropping the student's actual work to the Unistructural or Multistructural (Remember/Understand) level. Constructive alignment is severed: the lesson plan header names an analytical outcome, but the classroom evidence collects only transcription.

4.2. Modeling, Guided Practice, Independent Practice, and Planned Scaffolding Withdrawal

Skills in Social Studies—whether reading a choropleth map, corroborating two conflicting testimonies, or identifying continuous elements across two photographs—cannot be imparted through declarative lecturing; they must be demonstrated, supported, and practiced through the three-stage architecture of Cognitive Apprenticeship and the Gradual Release of Responsibility (Collins et al., 1989; Fisher & Frey, 2013; Pearson & Gallagher, 1983; Rosenshine, 2012; Wood et al., 1976):

1. Modeling ('I Do' / Modelleme, ~8–10 minutes): The teacher projects a single source or spatial representation and thinks aloud (sesli düşünme), making visible the internal cognitive questions an expert asks (e.g., 'Before reading this 1938 newspaper clipping, I look at the masthead...'). Crucially, the teacher models only the first item or category and stops; if the teacher completes the entire source on the board, modeling swallows guided practice.
2. Guided Practice ('We Do' / Rehberli Uygulama, ~12–15 minutes): Students work in pairs (carrying SDB2.2 Collaboration) on a second source supported by a temporary instructional scaffold (yapı iskelesi)—such as a four-field sourcing card, an eight-item continuity category

list, or a structured comparison table. While circulating, the teacher asks probing questions ('Where in the document do you see the warrant for that claim?') without supplying the answer. 3. Independent Practice ('You Do' / Bağımsız Uygulama, ~12–15 minutes): Each student works individually on a third, unseen source or case, with the temporary scaffold intentionally withdrawn (yapı iskelesinin geri çekilmesi). Learning evidence (öğrenme kanıtı) is collected strictly from this third stage.

Integrating generative AI or digital tools into this three-stage architecture requires a strict rule of scaffolding withdrawal, detailed in Table 2: an AI or digital support may assist the teacher in preparing differentiated input materials prior to class or serve as an object of collective critique during Guided Practice, but it must be completely withdrawn during Independent Practice so that the collected learning evidence reflects the student's autonomous mastery of the target process component.

Table 2. Taxonomic Verb-Alignment and AI Scaffolding Withdrawal Across the Three-Stage Instructional Architecture

Instructional Stage & Duration (40-Min Lesson)	Target Cognitive Operation (Bloom / SOLO & TYMM Code)	Permissible Role of Digital / GenAI Tool	Mandatory Boundary & Collected Evidence
Pre-Lesson Design (Teacher Preparation)	Selecting the single target skill step (tek basamak seçimi) and curating the 3-source package.	Assisting the teacher in locating candidate texts, drafting distractor misconceptions, or checking reading length.	Teacher verifies every outcome code (MEB, 2026b), primary source quote, and statistical denominator; unverified AI text is deleted.
Stage 1: Modeling ('I Do' / Modelleme) 8–9 Minutes	Making expert heuristics visible (Unistructural -> Relational transition; KB2.4, SBAB2.3).	Projecting a single primary document, map layer, or an intentionally flawed AI paragraph on the interactive board.	Teacher thinks aloud on the first claim/item only; AI does not 'explain' the lesson in place of the teacher's cognitive modeling.
Stage 2: Guided Practice ('We Do' / Rehberli Uygulama) 12–13 Minutes	Applying the target step in pairs with a structured scaffold (Relational level; SBAB2.4, KB2.7, SDB2.2).	Pairs use a physical sourcing card/category checklist to audit a second source or test an AI claim set against textbook/archive data.	Scaffold (cue card / partially filled row) is active; teacher circulates asking 'How do you know?' without giving the answer.
Stage 3: Independent Practice ('You Do' / Bağımsız Uygulama) 13–14 Minutes	Autonomous transfer to a new source/context (Relational / Extended Abstract; SBAB2.5–2.6, KB2.14).	ZERO GenAI generation by the student: scaffold and digital assistance are withdrawn (yapı iskelesinin geri çekilmesi).	Individual student product (completed sourcing card, 2-column corroboration form, or warranted inference) collected as learning evidence.

5. Domain-Specific AI Pedagogies Across History, Geography, Civics, and Economics

Because each Social Studies discipline possesses its own evidentiary rules and representational tools, AI integration cannot be governed by a single generic formula. Drawing on the 17 Social Sciences Field Skills (SBAB1–SBAB17; MEB, 2025b, pp. 69–83) and their operationalization in the 2026 Grade 4–7 curriculum (MEB, 2026b), the following subsections establish discipline-specific methodologies across the four core pillars of Social Studies, summarized in Table 3.

5.1. History Education (SBAB1, SBAB2, SBAB3, SBAB4): Contradiction Margins, Lateral Verification, and the Simulation Boundary

In historical inquiry, a single isolated document—or a single AI-generated summary—cannot activate corroboration (çapraz doğrulama; Wineburg, 1991; VanSledright, 2002; Wineburg, Martin, & Monte-Sano, 2013). Effective historical source work requires a curated 'source package' (kaynak paketi) governed by six design criteria: (1) alignment with a single open-ended inquiry question, (2) a strict numerical limit of 2–4 sources per lesson (three being the upper ceiling in middle school), (3) accessible length (at most half a page per source so reading load does not crowd out analysis), (4) positional diversity across makers and genres, (5) complete provenance headers (künye bütünlüğü), and (6) a deliberate contradiction margin (çelişki payı) across four possible discrepancy types: factual/numerical discrepancy (olgusal çelişki), emphasis discrepancy (vurgu farkı), evaluative discrepancy (değerlendirme farkı: gain for one group vs. loss for another), or memory discrepancy (bellek farkı in oral history; Thompson, 2000).

When an uncited digital or AI text is introduced into this environment, students apply a four-move lateral verification procedure modeled on Wineburg and McGrew (2019): (Move 1: Claim) extracting a single testable factual assertion; (Move 2: Source) selecting an independent authoritative source (textbook, primary document, museum catalog, official statistic) rather than asking the AI for its own source; (Move 3: Test) comparing the claim against the independent source to record one of three outcomes—Verified (doğrulandı), Contradicted (çelişti), or No Verifying Source Found (sınayacak kaynak bulunamadı); and (Move 4: Decision) recording only the verified finding and its true source citation in the student notebook. Pedagogically, the third outcome ('no verifying source found') is often the most transformative discovery in the lesson, as students see firsthand that a grammatically confident AI sentence may rest on no documentary foundation whatsoever.

Equally critical in History is enforcing the simulation boundary (canlandırma sınırı) in Historical Empathy (SBAB3; Ashby & Lee, 1987; Barton & Levstik, 2004; Endacott & Brooks, 2013; Foster, 1999; Lévesque, 2008; Seixas & Morton, 2013). Historical empathy is an evidence-bound cognitive reconstruction comprising three sequential stages: (1) establishing historical context (tarihsel bağlamlandırma, SBAB3.1: what was known, what technologies and institutional constraints existed), (2) perspective taking (bakış açısı alma: mapping the actor's position and available option set via an 'options map' [seçenek haritası]), and (3) evidence-backed evaluation (where affective connection [duygusal bağ] emerges as a warranted result of reading primary testimony rather than ungrounded emotional sympathy). Prompting an LLM to role-play a named historical figure in the first person ('I am Sultan Mehmed II...' or 'Interview Mustafa Kemal Atatürk via AI') manufactures counterfeit quotations that students readily

confuse with authentic primary testimony. Consequently, across Grades 4–7, neither students nor AI tools should ever generate first-person text in the voice of a named historical figure; in sensitive topics (war, forced migration, epidemic, trauma), first-person simulation is strictly excluded on ethical grounds in favor of third-person documentary analysis. Instead, AI's legitimate role in SBAB3 is to act as an 'anachronism-hunting reader' that challenges student-drafted period narratives ('Find three details in this paragraph that did not exist in the 15th century').

5.2. Geography Education (SBAB7, SBAB8, SBAB9, SBAB10): Spatial Reasoning Over Verbal Summarization

In Geography learning areas (Evimiz Dünya and Hayatımızdaki Ekonomi), the primary risk of generative AI is verbal reductionism: turning spatial relationships into memorized text lists (e.g., listing a region's climate and crops in bullet points without examining a map). Both the National Research Council (2006) and the TYMM Common Text (MEB, 2025b, p. 74) define Spatial Thinking (Mekânsal Düşünme, SBAB7) as the constructive union of spatial concepts, representational tools (maps, cross-sections, block diagrams, satellite imagery, sketch maps), and spatial reasoning processes (Downs & Stea, 1977; Gersmehl, 2014; Liben & Downs, 1989; Lynch, 1960; Monmonier, 1996). While MEB (2026b) explicitly names only map reading (SBAB10.1) on its learning-area skill lines, higher-order map analysis (SBAB10.2) and inference (SBAB10.3) must be carried by teacher design across five operational spatial components derived from SBAB7.1–7.11: (1) Location (Konum, SBAB7.1: absolute vs. relative location), (2) Place (Yer, SBAB7.2/7.4/7.6: distinctive natural and human traits), (3) Connection/Interaction (İlişki, SBAB7.3/7.5/7.8: how human activity links to physical conditions), (4) Scale (Ölçek, SBAB7.7: how the answer changes across neighborhood, district, and national scales), and (5) Pattern (Örüntü, SBAB7.10/7.11: where phenomena cluster or thin out along linear or nodal features).

Aligned with Roberts' (2013) geographical enquiry framework and SBAB8 (Coğrafi Sorgulama, SBAB8.1–8.5), digital mapping and AI tools pass the five-question threshold only when they allow students to overlay spatial layers, compare multi-date satellite imagery, or test scale shifts that static paper cannot show—while silent sketch maps (dilsiz kroki) on which students plot their own data and search for spatial patterns remain superior to pre-labeled digital slides.

5.3. Civics and Public Deliberation (SBAB5, OB6): Institutional Verification and Steel-Manning

In Living Together (Birlikte Yaşamak) and Our Living Democracy (Yaşayan Demokrasimiz), Social Participation (Sosyal Katılım, SBAB5; MEB, 2025b, pp. 72–73) progresses through questioning (KB2.8), establishing social contact (SBAB5.1), maintaining group dynamics (SBAB5.2), generating ideas (SBAB5.3), negotiating (SBAB5.4), and transforming ideas into action (SBAB5.5). As Hart's (1992) ladder of participation, Kahne and Sporte (2008), and Levinson (2012) demonstrate, authentic civic learning requires moving beyond tokenistic slogans to real institutional problem-solving. When students design community-scale proposals, LLMs frequently hallucinate administrative jurisdictions—confusing municipal duties (belediye), provincial directorates (il müdürlükleri), district governorships (kaymakamlık), and

civil society organizations. Therefore, in civic inquiry, AI suggestions regarding public receptors must be cross-verified against official statutory portals (`gov.tr`, e-Devlet, CİMER). Furthermore, in structured classroom deliberations and debates on controversial issues (Hess, 2004, 2009; Hess & McAvoy, 2015), assigning sides by lot and requiring each group to formulate the strongest possible evidence-backed argument for the opposing view ('steel-manning') prevents confirmation bias and turns AI into a tool for auditing missing stakeholder perspectives rather than writing speeches.

5.4. Economics, Entrepreneurship, and Data Literacy (SBAB6, SBAB11, OB3, OB7): The Tripartite Data Rule

In Economy in Our Lives (Hayatımızdaki Ekonomi), Entrepreneurship (Girişimcilik, SBAB6; MEB, 2025b, pp. 73–74) is structured around inquiry (KB2.8), observation-based prediction (KB2.11), assessing risks to make decisions (SBAB6.1), managing resources (SBAB6.2), and finalizing ideas (SBAB6.3)—converging with the European EntreComp framework (Bacigalupo et al., 2016). In the classroom, entrepreneurship is operationalized across four elements: Opportunity (distinguishing a genuine unmet need from a mere inconvenience), Resource (recognizing time, labor, information, and permissions alongside money), Risk/Uncertainty (writing 'three assumptions in our plan that could turn out to be wrong' under SBAB6.1), and Value (naming the beneficiary and a single measurable indicator rather than vague claims like 'awareness will increase'). When AI outputs or digital texts introduce economic and demographic statistics (linking to SBAB11: Table, Graph, Figure, and/or Diagram Skills; MEB, 2025b, p. 79), students enforce the Tripartite Statistical Rule: every numerical claim must specify (1) the exact year, (2) the publishing institution (e.g., TÜİK, Central Bank, Provincial Directorate), and (3) the denominator (share of what total?). If any of the three is missing, the number is classified as an unverified impression rather than evidence.

Table 3. Domain-Specific AI Integration and Verification Protocols Across the Four Pillars of Social Studies

Disciplinary Domain & TYMM Field Skills	Primary Epistemological Risk of Unchecked GenAI Use	Authorized Pedagogical Role of GenAI / Digital Tool	Observable Student Verification & Learning Evidence
History (SBAB1, SBAB2, SBAB3, SBAB4)	Fabricated primary quotes; first-person role-play of named figures; erasure of contradiction margins across sources.	Generating intentionally anachronistic period stories for students to debug; serving as an uncited claim set for lateral reading.	4-Move Lateral Verification Table; 3-source corroboration form; options map (seçenek haritası) written in the 3rd person.
Geography (SBAB7, SBAB8, SBAB9, SBAB10)	Verbal reductionism (replacing spatial pattern analysis with memorized text bullet points); scale blindness.	Multi-layer GIS/satellite comparison and scale-shift visualization (neighborhood vs. province vs. country) when paper is insufficient.	Student-annotated sketch map (kroki) and 5-component spatial analysis form (Location, Place, Connection, Scale, Pattern).

Disciplinary Domain & TYMM Field Skills	Primary Epistemological Risk of Unchecked GenAI Use	Authorized Pedagogical Role of GenAI / Digital Tool	Observable Student Verification & Learning Evidence
Civics & Democracy (SBAB5, OB6)	Hallucinated administrative jurisdictions; one-sided confirmation bias in controversial issue discussions.	Testing whether a draft civic proposal overlooked affected community groups; contrasting AI jurisdictional claims with gov.tr.	Institutional Receptor Matrix (matching local problem to exact statutory authority) and dual-perspective deliberation brief.
Economics & Data (SBAB6, SBAB11, OB3, OB7)	Denominator-free statistics ('60% work in agriculture'); overly optimistic project plans lacking risk analysis.	Stress-testing student entrepreneurship plans by prompting: 'Identify 3 resource or regulatory assumptions that could fail.'	Tripartite Data Audit (Year, Institution, Denominator) and 4-element Entrepreneurship Plan including '3 falsifiable assumptions.'

6. Subtractive Material Engineering and Process-Authenticated Assessment

6.1. Subtractive Design ('Weeding') and Three Rapid Pre-Flight Material Tests

A core finding of multimedia learning research is the coherence effect: adding visually appealing but instructionally irrelevant graphics, borders, or text blocks ('seductive details') degrades student learning by consuming limited working memory (Mayer, 2009; Mayer & Moreno, 2003; Sweller et al., 2011, 2019). Consequently, whereas novice material designers —and generative AI layout tools—habitually ask 'What else can I add to this worksheet or slide deck?', expert instructional design reverses the question: 'What can I remove from this page without losing the cognitive operation?' Alongside the six-point defect audit for AI-drafted worksheets (checking for fabricated codes/quotes, verb drift, factual errors, generic content-free activities, loss of student thinking labor, and assessment gaps; MEB, 2025a, 2026a), every instructional material should undergo three rapid pre-classroom tests taking less than twenty minutes in total (Table 4).

Table 4. Three Rapid Pre-Classroom Tests for Auditing Instructional Materials and AI-Assisted Worksheets

Pre-Classroom Test	Execution Procedure (Under 20 Minutes Total)	Specific Design Defects Detected & Corrected
1. Self-Completion Timed Test (Kendin Doldur Testi)	The teacher completes the worksheet by hand, exactly as a student would, while running a stopwatch.	Exposes unrealistic time assumptions, cramped writing boxes where expected answer length exceeds physical space, ambiguous prompts, and unanswerable questions.

Pre-Classroom Test	Execution Procedure (Under 20 Minutes Total)	Specific Design Defects Detected & Corrected
2. Subtraction Test (Çıkarma Testi)	The teacher covers each visual, box, border, and question one by one and asks: 'If this element were removed, what would the lesson lose?'	Eliminates decorative clutter (süsleme), redundant recall questions (ensuring transcription items do not exceed one-third of the page), and seductive details (Mayer & Moreno, 2003).
3. Blind Peer Test (Akran Testi)	The teacher hands the worksheet to a colleague without any verbal explanation and asks them to state what the student is supposed to do.	Reveals hidden assumptions that exist only in the designer's mind, unclear task sequences, and missing purpose statements.

6.2. Psychometric Validity, Comment-First Feedback, and the Two-Tier Process-Product Rubric

Assessing Social Studies skills requires satisfying three conditions simultaneously: the product condition (the skill must be observed through a tangible student product rather than a definition test), the transfer condition (the task must present a new source or context not memorized in class), and the criterion condition (scoring must rely on pre-published, observable behavioral descriptions rather than vague adverbs like 'very well / adequately / poorly'; Brookhart, 2013; Messick, 1994; Sadler, 1989; Wiggins & McTighe, 2005). In the formative phases of inquiry, Butler's (1988) landmark experimental finding is decisive: when students receive a numerical grade alongside written comments on an intermediate draft, ego-involvement eclipses task-involvement and the formative effect of the comment disappears (see also Black & Wiliam, 1998; Hattie & Timperley, 2007; Wiliam, 2011). Therefore, intermediate inquiry products (ara ürünler) must receive criterion-referenced comments without numerical grades. Furthermore, because automated GPT/AI detectors exhibit severe false-positive bias against non-native and concisely written student texts (Liang et al., 2023), summative authenticity must be secured by combining disciplinary product criteria with content-independent process criteria, as modeled in Table 5.

Table 5. Integrated Two-Tier Analytical Rubric: Pairing Disciplinary Evidence Criteria (SBAB2) with Process-Authentication Criteria

Rubric Tier & Criterion	Level 1: Developing (Geliştirilmeli)	Level 2: Proficient (Yeterli)	Level 3: Advanced (İleri Düzey)
Tier I (Product) — 1: Sourcing & Provenance (Künyeleme, SBAB2.3)	Identifies only one source's maker/date or leaves provenance fields mostly blank.	Identifies maker, date, and purpose accurately for one source with at most one missing field in the second.	Identifies maker, date, audience, and purpose for both sources and explains how provenance shapes reliability.
Tier I (Product) — 2: Cross-Source Corroboration (Çapraz Doğrulama, SBAB2.4)	Fills only a single column of the comparison form; summarizes one text in isolation.	Fills both columns of the comparison table, noting at least one shared or diverging point.	Fills both columns with at least two specific elements from each source and names the exact type of discrepancy.

Rubric Tier & Criterion	Level 1: Developing (Geliştirilmeli)	Level 2: Proficient (Yeterli)	Level 3: Advanced (İleri Düzey)
Tier I (Product) — 3: Warranted Inference (Çıkarımın Dayanağı, SBAB2.5–2.6)	Writes a general summary or personal opinion instead of an evidence-backed inference.	Writes a relevant inference but cites the source only indirectly or generally.	Writes a clear inference and pairs it with a direct, explicit warrant quoted or cited from the verified source.
Tier II (Process) — 4: Contemporaneous Source Log & Decision Record	Attempts to reconstruct sources after finishing; states what was done without explaining why.	Records every source with full provenance at the moment of use and states the reason for key choices.	Records conflicting sources, documents rejected claims/paths, and justifies why they were discarded.
Tier II (Process) — 5: In-Class Intermediate Draft & Tool Disclosure (EDE)	Submits only a final product at the deadline with no prior draft or vague/missing tool disclosure.	Presents a verifiable intermediate draft mid-process and states which digital/AI tool was used at which step.	Revises the plan based on formative feedback on the intermediate draft and explicitly differentiates personal contribution from tool output.

7. Discussion and Conclusion: The Teacher as Curricular-Instructional Gatekeeper

Over three decades ago, Thornton (1991, 2005) conceptualized the Social Studies teacher as the 'curricular-instructional gatekeeper'—the reflective professional (Schön, 1983; Spillane, Reiser, & Reimer, 2002) whose day-to-day design decisions determine whether official curriculum standards translate into genuine student thinking or degenerate into mechanical routine. The rise of generative artificial intelligence does not diminish this gatekeeping role; it intensifies it. Because LLMs can produce fluent pseudo-historical role-plays, uncited factual summaries, and visually tidy recall worksheets in seconds, the boundary between an classroom that practices disciplinary inquiry and one that merely simulates it now rests entirely on the teacher's epistemological vigilance.

By aligning the four-dimensional Inquiry Arc of the NCSS C3 Framework (2013) with the skill, literacy, and value architecture of the 2024–2026 Century of Türkiye Maarif Model (MEB, 2025b, 2026b), this study has established an integrated design framework for Social Studies education in the age of AI. When teachers subject digital tools to a strict five-question pre-selection threshold, withdraw scaffolding systematically between Guided and Independent Practice, enforce domain-specific verification rules across History, Geography, Civics, and Economics, apply subtractive material testing, and evaluate process artifacts alongside final products, generative AI ceases to be a threat to cognitive rigor. Instead, it becomes a controllable laboratory material through which students learn the most enduring habit of Social Studies: never accepting a fluent claim until its provenance, context, scale, and evidence have been independently tested.

References

American Historical Association (AHA). (2015). The C3 Framework: Advocating for K–12 social studies education. Perspectives on History, 53(3). https://www.historians.org/perspectives-article/the-c3-framework-advocating-for-k-12-social-studies-education-march-2015/

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.

Ashby, R., & Lee, P. (1987). Children's concepts of empathy and understanding in history. In C. Portal (Ed.), The history curriculum for teachers (pp. 62–88). Falmer Press.

Bacigalupo, M., Kampylis, P., Punie, Y., & Van den Brande, G. (2016). EntreComp: The entrepreneurship competence framework. Publications Office of the European Union.

Barr, R. D., Barth, J. L., & Shermis, S. S. (1977). Defining the social studies (Bulletin 51). National Council for the Social Studies.

Barton, K. C. (2005). Primary sources in history: Breaking through the myths. Phi Delta Kappan, 86(10), 745–753.

Barton, K. C., & Levstik, L. S. (2004). Teaching history for the common good. Lawrence Erlbaum.

Biggs, J. (1996). Enhancing teaching through constructive alignment. Higher Education, 32(3), 347–364.

Biggs, J. B., & Collis, K. F. (1982). Evaluating the quality of learning: The SOLO taxonomy. Academic Press.

Biggs, J., & Tang, C. (2011). Teaching for quality learning at university (4th ed.). McGraw-Hill.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74.

Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40. https://doi.org/10.3316/QRJ0902027

Brookhart, S. M. (2013). How to create and use rubrics for formative assessment and grading. ASCD.

Bruner, J. S. (1960). The process of education. Harvard University Press.

Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. British Journal of Educational Psychology, 58(1), 1–14.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 453–494). Lawrence Erlbaum.

Dewey, J. (1910). How we think. D. C. Heath.

Dewey, J. (1916). Democracy and education: An introduction to the philosophy of education. Macmillan.

Downs, R. M., & Stea, D. (1977). Maps in minds: Reflections on cognitive mapping. Harper & Row.

Endacott, J., & Brooks, S. (2013). An updated theoretical and practical model for promoting historical empathy. Social Studies Research and Practice, 8(1), 41–58.

Evans, R. W. (2004). The social studies wars: What should we teach the children? Teachers College Press.

Fikir ve Sanat Eserleri Kanunu, Law No. 5846 (1951). Resmî Gazete, December 13, 1951, Issue: 7981.

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

Foster, S. J. (1999). Using historical empathy to excite students about the study of history: Can you empathize with Neville Chamberlain? The Social Studies, 90(1), 18–24.

Gersmehl, P. (2014). Teaching geography (3rd ed.). Guilford Press.

Hart, R. A. (1992). Children's participation: From tokenism to citizenship (Innocenti Essays No. 4). UNICEF International Child Development Centre.

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112.

Hess, D. E. (2004). Controversies about controversial issues in democratic education. PS: Political Science & Politics, 37(2), 257–261.

Hess, D. E. (2009). Controversy in the classroom: The democratic power of discussion. Routledge.

Hess, D. E., & McAvoy, P. (2015). The political classroom: Evidence and ethics in democratic education. Routledge.

Kahne, J., & Sporte, S. E. (2008). Developing citizens: The impact of civic learning opportunities on students' commitment to civic participation. American Educational Research Journal, 45(3), 738–766.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational Psychologist, 41(2), 75–86.

Kişisel Verilerin Korunması Kanunu, Law No. 6698 (2016). Resmî Gazete, April 7, 2016, Issue: 29677.

Lévesque, S. (2008). Thinking historically: Educating students for the twenty-first century. University of Toronto Press.

Levinson, M. (2012). No citizen left behind. Harvard University Press.

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. Patterns, 4(7), 100779. https://doi.org/10.1016/j.patter.2023.100779

Liben, L. S., & Downs, R. M. (1989). Understanding maps as symbols: The development of map concepts in children. Advances in Child Development and Behavior, 22, 145–201.

Lynch, K. (1960). The image of the city. MIT Press.

Mager, R. F. (1997). Preparing instructional objectives (3rd ed.). Center for Effective Performance.

Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist, 38(1), 43–52.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. Educational Researcher, 23(2), 13–23.

Millî Eğitim Bakanlığı (MEB). (2025a). Eğitimde yapay zekâ politika belgesi ve eylem planı (2025–2029). MEB Yayınları.

Millî Eğitim Bakanlığı (MEB). (2025b). Türkiye Yüzyılı Maarif Modeli öğretim programları ortak metni. Talim ve Terbiye Kurulu Başkanlığı. https://tymm.meb.gov.tr/upload/brosur/ortak_metin.pdf

Millî Eğitim Bakanlığı (MEB). (2026a). Eğitimde yapay zekâ uygulamaları etik kılavuzu. Talim ve Terbiye Kurulu Başkanlığı. https://ttkb.meb.gov.tr/meb_iys_dosyalar/2026_01/09152326_egitimdeyapayzekauygulamalari_etik_kilavuzu.pdf

Millî Eğitim Bakanlığı (MEB). (2026b). Türkiye Yüzyılı Maarif Modeli Sosyal Bilgiler dersi öğretim programı (4, 5, 6 ve 7. sınıflar). Talim ve Terbiye Kurulu Başkanlığı. https://tymm.meb.gov.tr/ogretim-programlari/ders/sosyal-bilgiler-dersi

Monmonier, M. (1996). How to lie with maps (2nd ed.). University of Chicago Press.

Monte-Sano, C. (2011). Beyond reading comprehension and summary: Learning to read and write in history by focusing on evidence, perspective, and interpretation. Curriculum Inquiry, 41(2), 212–249.

National Council for the Social Studies (NCSS). (2010). National curriculum standards for social studies: A framework for teaching, learning, and assessment. NCSS.

National Council for the Social Studies (NCSS). (2013). The College, Career, and Civic Life (C3) Framework for Social Studies State Standards: Guidance for enhancing the rigor of K–12 civics, economics, geography, and history. NCSS. https://www.socialstudies.org/standards/c3

National Council for the Social Studies (NCSS). (2023). Definition of social studies. NCSS.

National Research Council. (2006). Learning to think spatially: GIS as a support system in the K–12 curriculum. National Academies Press.

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. Contemporary Educational Psychology, 8(3), 317–344.

Roberts, M. (2013). Geography through enquiry: Approaches to teaching and learning in the secondary school. Geographical Association.

Rosenshine, B. (2012). Principles of instruction: Research-based strategies that all teachers should know. American Educator, 36(1), 12–19, 39.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144.

Schön, D. A. (1983). The reflective practitioner: How professionals think in action. Basic Books.

Seixas, P., & Morton, T. (2013). The big six historical thinking concepts. Nelson Education.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. Review of Educational Research, 72(3), 387–431.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive load theory. Springer.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. Educational Psychology Review, 31(2), 261–292.

Taba, H. (1962). Curriculum development: Theory and practice. Harcourt, Brace & World.

Talim ve Terbiye Kurulu Başkanlığı. (2024). Türkiye Yüzyılı Maarif Modeli öğretim programlarının kabulüne ilişkin 23/05/2024 tarih ve 20 sayılı Karar.

Thompson, P. (2000). The voice of the past: Oral history (3rd ed.). Oxford University Press.

Thornton, S. J. (1991). Teacher as curricular-instructional gatekeeper in social studies. In J. P. Shaver (Ed.), Handbook of research on social studies teaching and learning (pp. 237–248). Macmillan.

Thornton, S. J. (2005). Teaching social studies that matters: Curriculum for active learning. Teachers College Press.

Tyler, R. W. (1949). Basic principles of curriculum and instruction. University of Chicago Press.

VanSledright, B. A. (2002). In search of America's past: Learning to read history in elementary school. Teachers College Press.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Webb, N. L. (1997). Criteria for alignment of expectations and assessments in mathematics and science education (Research Monograph No. 6). Council of Chief State School Officers.

Westheimer, J., & Kahne, J. (2004). What kind of citizen? The politics of educating for democracy. American Educational Research Journal, 41(2), 237–269.

Wiggins, G., & McTighe, J. (2005). Understanding by design (2nd ed.). ASCD.

Wiliam, D. (2011). Embedded formative assessment. Solution Tree Press.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. Journal of Educational Psychology, 83(1), 73–87.

Wineburg, S. (2001). Historical thinking and other unnatural acts: Charting the future of teaching the past. Temple University Press.

Wineburg, S., Martin, D., & Monte-Sano, C. (2013). Reading like a historian: Teaching literacy in middle and high school history classrooms. Teachers College Press.

Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. Teachers College Record, 121(11), 1–40.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100.